CENTER FOR ARMY ANALYSIS
6001 GOETHALS ROAD
FORT BELVOIR, VA 22060-5230

CAA-2020055

COVID-19 DATA ANALYSIS PROJECT

JUNE 2021

DISTRIBUTION A. Approved for public release: distribution unlimited.

DISCLAIMER

The findings of this report are not to be construed as an official Department of the Army position, policy, or decision unless so designated by other official documentation. Comments or suggestions should be addressed to:

Director
Center for Army Analysis
ATTN: CSCA-OA
6001 Goethals Road
Fort Belvoir, VA 22060-5230

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED *(From –To)***
June 2021	Final	01APR20 – 04JUN21

4. TITLE AND SUBTITLE

COVID-19 Data Analysis Project
CDAP

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
MAJ Chad Chapman, MAJ Maxine Drake, MAJ Trey Harvey, Mr. Collin Henley, MAJ Sandra Jackson, MAJ Scott Lynch, Mr. Kyle Minor, LTC Matthew Pacheco, MAJ Dusty Turner, Mr. Robert Ward, Mr. Michael Warme, Ms. Michaela Zuber

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

Center for Army Analysis
6001 Goethals Road
Fort Belvoir, VA 22060-5230

8. PERFORMING ORGANIZATION REPORT NUMBER

CAA-2020055

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

HQDA DCS-G-3/5/7 DAMO SS
200 Army Pentagon Washington, DC 2310-0200

10. SPONSOR/MONITOR'S ACRONYM(S)

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

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
CAA analytic support that addresses Army and Department of Defense leadership's highest priority issues concerning the global COVID-19 pandemic. Enables senior leaders across the Army enterprise to see, understand, and respond to rapidly changing conditions resulting from the COVID-19 pandemic by providing timely, credible, and actionable analytics to help anticipate and mitigate risks to force and mission.

15. SUBJECT TERMS
COVID-19, XGBoost, cPROBE,

16. SECURITY CLASSIFICATION OF: UNCLASSIFIED			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Mr. Russell Pritchard
U	U	U	UU	**56**	**19b. TELEPHONE NUMBER *(include area code)*** (703) 806-5421

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

(THIS PAGE INTENTIONALLY LEFT BLANK)

COVID-19 DATA ANALYSIS PROJECT

SUMMARY

THE STUDY PURPOSE AND SPONSOR

Severe acute respiratory syndrome associated Coronavirus-2 (SARS-CoV-2), which causes novel Coronavirus Disease 2019 (COVID-19), emerged in late 2019. By February 2020, increasing numbers of confirmed COVID-19 cases in the United States accelerated national- and state-level concerns regarding risk mitigation. As national- and state-level leadership began implementing strategies to slow the spread of disease within the United States, Army senior leaders issued guidance to reduce risk to force and mission while preparing to assist civil authorities. While reducing risk was a necessary goal, it challenged what was then the Army's number one priority—Readiness. Therefore, the Army identified the need for objective and rigorous analysis to assist in making difficult decisions pitting readiness against COVID-19 risk. It was this—assisting leaders across the breadth of the Army, Joint, and Department of Defense (DoD) enterprises in understanding the impact of COVID-19 and making well-informed decisions—that was CAA's chief study purpose.

The pandemic presented the Army with three unique challenges. First, few Army organizations had the expertise or tools to quickly assemble large amounts of disparately generated data and synthesize them to enable senior leader awareness, understanding, and action. Second, although many non-military organizations produced notable COVID-19 analyses, they did not tailor analyses to Army leaders nor were they responsive to leader input or requests. Third, and most importantly, Army leaders did not have COVID-19 studies at the level of granularity they needed, hindering them from managing resources effectively. The study team focused on alleviating these problems by producing tools, estimates, reports, and briefings to enable senior leadership to see, understand, and act in a complex and volatile environment.

The Center for Army Analysis (CAA) COVID-19 Analysis Study team addressed these three challenges through decision support analysis and informational briefings. The Vice Chief of Staff of the Army (VCSA) charged CAA with synchronizing the Army's COVID-19 modeling effort. He directed CAA to (1) provide nationwide COVID-19 analysis tailored to the Army's needs and (2) synthesize data and information from multiple analytical organizations to inform Army senior leaders of the operating environment. CAA supported Army and Joint leaders, including (but not limited to) GEN Joseph Martin, 37th VCSA; LTG Charles Flynn, Deputy Chief of Staff of the Army, G-3/5/7; LTG Laura Richardson, Commander of U.S. Army North; LTG Lee Quintas, Deputy Commanding General, U.S. Army Forces Command (FORSCOM); MG Stephen Sklenka, Director of Strategic Planning and Policy Directorate, U.S. Indo-Pacific Command (USINDOPACOM); and MG Douglas Crissman, Deputy Commander of U.S. Army Central (USARCENT).

THE STUDY OBJECTIVES were to:

(1) Create an SEIR forecast model to forecast future cases and hospitalizations for every U.S. county and 53 INDOPACOM countries.

(2) Create SEIR projection models to answer focused Army policy and resource management questions.

(3) Create an agent-based model to answer questions related to unit readiness.

THE MAIN ASSUMPTION

One overarching assumption was that a compartmental model was appropriate for forecasting COVID-19 cases.

METHODS

The study team's core analytical functions were to (1) develop automated techniques to gather, consolidate, analyze, and visualize data related to public health COVID-19 and (2) develop original analysis and modeling on COVID-19. To accomplish these core functions, the study team employed a wide range of methods. Throughout the duration of the study, customer needs evolved, requiring the study team to continuously improvise and develop unique modeling solutions to satisfy demand and solve problems. These analytical tasks ranged from estimating medical supply orders for USARCENT, to creating heat maps for the Army Geospatial Center illustrating the amount of time until county hospitals were under stress due to demand, to studying COVID-19 cases across over 3,000 distinct geographic regions in the United States and Indo-Pacific region, to leading collaboration efforts among fellow analytical organizations. The study team consistently produced original analyses and models. Its two most lasting analytical efforts were its development of a nationwide COVID-19 forecast model and collaborating with two other agencies to develop an agent-based COVID-19 stochastic simulation.

The study team's COVID-19 forecast model estimated future active COVID-19 cases and their resulting hospitalizations nationwide at the U.S. county level. The model also forecasted cases for 53 countries in the USINDOPACOM area of responsibility. The model was, at its core, an SEIR (Susceptible, Exposed, Infectious, Removed) compartmental model. The SEIR model relied on static, deterministic parameters, with the exception of the contact rate, β. The model used a separate gradient boosted machine-learning algorithm to estimate the future βs for each county or country. The gradient boosted model was trained on 38 variables from data sets covering nationwide mobility, seasonality, geographic attributes, state policies, COVID-19 testing results, and population demographics. This model produced valuable results on its own; but, more importantly, the model served as a driving analytical force, informing additional analytical methods.

The study team also collaborated with the Army Public Health Center (APHC) and Lawrence Livermore Labs (LLNL) to develop an agent-based COVID-19 stochastic simulation. The model was a robust set of Monte Carlo simulations that modeled an Army brigade deploying to and

training at a Combat Training Center. The CAA study team and its partners built the model using the framework of the Susceptible, Infected, Removed (SIR) model, and integrated thoroughly researched epidemiological characteristics and individual and unit interaction dynamics.
All of the study team's analysis and modeling hinged on a critical component—a wealth of open-source data. The study team used the R programming language to scrape and clean data—a task that could take hours using common DoD tools—in seconds, every day. These data allowed the study team to pursue multiple analytical methods efficiently.

IMPACT ON CUSTOMERS
The study's most notable customers were Army senior leaders in Headquarters, Department of the Army (HQDA); installation commanders and their staffs; USINDOPACOM; USARCENT; FORSCOM; and U.S. Special Operations Command. CAA rendered its analysis through two primary mechanisms: inclusion in senior-leader-focused dashboards and the development and delivery of briefings and reports. The impact of these mechanisms ranged from directly informing discrete resource management decisions to improving general situational awareness for organizational senior leaders.

Study leadership briefed Army senior leaders during crisis-focused Army synchronization meetings chaired by the VCSA or the HQDA Deputy Chief of Staff, G-3/5/7. These briefings provided recent COVID-19 trends, highlighted areas where the study team projected stress on hospital infrastructure, and highlighted installations in areas where the study team projected significant COVID-19 surges. These briefings helped establish situational awareness among Army senior leaders.

The study team supported Army installation commanders and their staffs across the Army with daily reports published by Army Vantage, the Army's senior leader dashboard, on their COVID-19 dashboard. These reports provided COVID-19-case projections around Army installations. Army Vantage reported 250-300 unique users view this dashboard daily and 700-800 users weekly. Additionally, the study team provided daily projections directly to the Army Materiel Command, U.S. Army Corps of Engineers, and U.S. Army North at various points during the pandemic.

For USINDOPACOM, the study team produced a weekly report that included analysis of recent data in all countries in their area of responsibility (AOR) as well as 45 days of projections for each country in the AOR. USINDOPACOM used this report to brief ambassadors and flag officers. Planners also used the report to identify strategic, operational, and tactical risk. Finally, it informed decisions regarding COVID-19-response policies.

The study supported USARCENT resource-management decisions through tailored modeling and analysis. Throughout the pandemic, USARCENT has ordered supplies based on the Centers for Disease Control and Prevention (CDC) estimates of U.S. COVID-19 supply requirements. The study team built multiple models to refine these estimates, allowing USARCENT to scale down excess medical supplies without incurring risk to ill Soldiers.

Supporting FORSCOM was a collaborative effort among multiple organizations. The CAA study team led the collaboration among FORSCOM, APHC, and LLNL, to develop a robust simulation that determines the optimal COVID-19 testing protocol within Army units. LLNL built the agent-based simulation in the Python programming language to show the spread of a virus given various testing scenarios. The CAA study team provided (1) statistical expertise and an understanding of COVID-19 data to tune the model's many parameters and (2) operational experience, helping to make LLNL's model realistic. The results of this model shaped decisions regarding the best way to conduct surveillance testing for units conducting major training events at the Army's Combined Training Centers.

THE STUDY EFFORT was conducted by MAJ Chad Chapman, MAJ Maxine Drake, Mr. Collin Henley, MAJ Sandra Jackson, MAJ Scott Lynch, Mr. Kyle Minor, LTC Matthew Pacheco, MAJ Harvey Smith, MAJ Dusty Turner, Mr. Robert Ward, Mr. Michael Warme, and Ms. Michaela Zuber.

COMMENTS AND QUESTIONS may be sent to the Director, Center for Army Analysis, ATTN: CSCA-OA, 6001 Goethals Road, Suite 102, Fort Belvoir, VA 22060-5230.

CONTENTS **Page**

1 INTRODUCTION 1
1.1 Literature Review 2
1.2 Laying the Groundwork 3

2 THE SEIR MODEL 5
2.1 Modeling Methodology 5
2.2 Inferring β by Modelling Reff 9
2.3 Source Data 11
2.4 Model Validation and Model Improvement Strategy 12
2.5 Impacts of the Model 15
2.6 Using the SEIR Model for Focused Studies 19

3 AGENT-BASED MODEL 23
3.1 Introduction 23
3.2 Agent-based Model: Motivation and Challenges 23
3.3 Creating Disease Profiles 23
3.4 Creating an Interaction Network 27
3.5 Initializing the Model 28
3.6 Running the Model 29
3.7 Impact 29

4 CONCLUSION 33

APPENDIX A STUDY CONTRIBUTORS 35

APPENDIX B REQUEST FOR ANALYTICAL SUPPORT 37

APPENDIX C BIBLIOGRAPHY 39

(THIS PAGE INTENTIONALLY LEFT BLANK)

FIGURES

Figure 1. Graphical Depiction of an SEIR Model 5
Figure 2. A heat map describing projected hospital stress in each county in the continental US according to the study team's SEIR model. The study team provided this product to USARNORTH to inform their decisions regarding field hospitals and troop mobilization. 16
Figure 3. New case predictions according to the study team SEIR model provided to USINDOPACOM as part of the 2 February 2021 report. 17
Figure 4. A snapshot of the Army Vantage COVID-19 dashboard, which displayed this projection of new cases in the Fort Bragg area on 5 April 2021. 18
Figure 5. Results of analysis regarding how closures of base facilities would impact the most likely and dangerous COVID-19 outcome 20
Figure 6. A sample disease profile *Note: The illness duration represents how long a person presents with symptoms and has no bearing on the spread of the disease under the study team's assumptions and so was not modeled 24
Figure 7 Empirical CDFs of the time based elements disease profile 26
Figure 8. A sample unit hierarchy used to describe the process of building a unit interaction matrix 27
Figure 9. The matrix (left) and graph (right) form of the interaction of an infantry brigade combat team. In the matrix, each person is both a row and a column. The values of the matrix represent the contact rate between two individuals. In the graph, each person is a node and each connection between nodes is weighted based on the contact rate between the two individuals. 28
Figure 10. Simulation results of experiments considering different testing protocols for units prior to a CTC rotation. Adding a rapid antigen surveillance test (green dots) after arrival to the training center reduced new exposures without increasing quarantines when compared to the current protocol (orange dots); the numbers and colors correspond to the various testing protocols examined 30

(THIS PAGE INTENTIONALLY LEFT BLANK)

TABLES

Table 1. SEIR Ordinary Differential Equations System Formulation 6
Table 2. Initial Conditions of the SEIR model 7
Table 3. Medical Infrastructure Demand Rates by Age Cohort 9

(THIS PAGE INTENTIONALLY LEFT BLANK)

1 INTRODUCTION

The severe acute respiratory syndrome associated coronavirus-2 (SARS-CoV-2), which causes novel coronavirus disease 2019 (COVID-19), emerged toward the end of 2019 in Wuhan, China. The severity of COVID-19 and the apparent high transmissibility caused international concern for the potential of a global pandemic. During February 2020, increasing numbers of confirmed COVID-19 cases in the United States accelerated national- and state-level concerns regarding risk mitigation.

As national- and state-level leadership began implementing strategies to slow the spread of disease within the United States, Army senior leaders issued guidance to reduce risk to force and mission while preparing to assist civil authorities. A central theme in all guidance from senior leaders was readiness. Highly restrictive policies mitigating risk to Soldiers and communities restricted personnel availability and, thus, tactical unit readiness. Conversely, failing to act in response to the pandemic would have resulted in negative health consequences for Soldiers and communities, also degrading unit readiness. Senior leaders had to design a suite of policies to support the national COVID-19 response to ensure no long-term drop in readiness while also responding to the short-term health threat to readiness. Therefore, the Army identified the need for objective and rigorous analysis to assist in maximizing readiness in a constrained operating environment with little operational intelligence. It was this—assisting leaders across the breadth of the Army, Joint, and Department of Defense (DoD) enterprises in understanding the impact of COVID-19 and making well-informed decisions—that was CAA's chief study purpose.

CAA's study team of operations researchers was junior in grade, had no epidemiological experience, and had no high-performance computing capability; these limitations, however, did not prevent them from developing the most influential COVID-19 model in the U.S. Army and becoming its lead analytical agency with respect to modeling the impacts of this pandemic. The study team accomplished this through extensive academic research, collaboration with epidemiologists, and cutting-edge data science techniques.

Development of this model began in March 2020, when CAA created a large-scale Susceptible, Exposed, Infectious, Removed (SEIR) model to forecast COVID-19 cases and hospitalizations for each county in the United States. CAA then began presenting SEIR model insights and results to Army senior leaders, which built increasing appetite for CAA's analysis at several commands. Section 2 will address the SEIR model and these relationships.

In the summer of 2020, the study team began addressing more focused analytical problems, using an agent-based model to answer training- and deployment-specific COVID-19 questions from various Army commands and Army service component commands. Section 3 addresses this model and its associated sponsor relationships.

Throughout both of these phases, the study team constantly adapted to changing senior leader priorities and requests, as well as the rapidly changing nature of the pandemic, to help Army leaders understand the impact of COVID-19 and make well-informed decisions. Ultimately,

CAA's work enabled the Army to respond to the Nation in crisis and sustain the readiness of the Army to respond when needed.

1.1 Literature Review

Due to the ever-evolving nature of the pandemic and analytical needs of the project's sponsors, the study team conducted the literature review in parallel with the modeling and analytical efforts. The study team constantly absorbed the most up-to-date information about COVID-19 from academic research and collaboration with epidemiologists, and implemented it as quickly as was feasible in models and analysis—even when that meant changing or undoing work that had been informed by the best available information just days or weeks before.

At the initial stages of modeling, the study team sought to understand Susceptible, Infectious, Removed (SIR) and other compartmental models.[1] [2] [3] Once information became available about the lag between exposure to COVID-19 and full infectiousness, the study team researched compartmental models that were more elaborate, and chose to use a SEIR instead of a SIR model.[4] The study team conducted further research to inform implementation of the SEIR model, including gathering estimates for key parameters, and then continued to refine these estimates based on emerging research.[5] [6] Simultaneously, the study team drew on research that guided the transition from 56 state/territory models to over 3,000 county models.[7]

A key moment during model development was the realization that the transmissibility of COVID-19 could be broken down into behavioral and biological components. A number of authors provided insight regarding how to incorporate these findings into the SEIR model to ensure that it took account of variance in human behavior, prompting the creation of a sub-model to predict future values of the effective reproductive number (R_{eff}).[8] [9] [10]

Even after the model was in a more stable condition, the epidemiological community continued to publish new research that the study team incorporated into its models, thus maximizing the models' accuracy and relevance to Army decision makers. For instance, the study team incorporated non-pharmaceutical interventions[11] into the R_{eff} model as a proxy for actual

[1] Massad et al. "Forecasting versus projection models in epidemiology" *Medical hypotheses.* 17-22.

[2] Lipsitch et al. "Estimating Case-Fatality Risks during Outbreaks." *PLoS Neglected Tropical Diseases*.

[3] Hethcote, "The Mathematics of Infectious Diseases." *SIAM* Review. 599-653.

[4] Nishiura, Linton, and Akhmetzhanov. "Serial interval of (COVID-19) infections." *International Journal of Infectious Diseases.* 284-286.

[5] Lauer et al. "The Incubation Period of (COVID-19)." *Annals of internal medicine.* 577-582.

[6] Bar-On et al. "A quantitative compendium of COVID-19 epidemiology." 1-10.

[7] Hu, Nigmatulina, and Eckhoff. "The scaling of contact rates for infectious disease models." *Mathematical biosciences,* 125–134.

[8] *Jones. "Notes on R_0."*

[9] Nishiura and Chowell. "The Effective Reproduction." *Mathematical and Statistical Estimation Approaches in Epidemiology.* 103–121.

[10] Delamater et al. "Complexity of the Basic Reproduction Number (R0)." *Emerging Infectious Diseases.* 1-4.

[11] Kucharski et al. "Early dynamics of transmission and control of COVID-19." *The Lancet Infectious Diseases. 553 - 558*.

human behavior and then added weekly flu patterns as a proxy for the seasonal effects on virus transmission.[12] The study team also reviewed research on other dynamics of COVID-19 transmissibility, such as heterogeneous mixing,[13] but found that modeling these dynamics did not improve the model's predictive accuracy.

When adapting to more focused analytical questions in chapter 3 of the study, the study team conducted additional research on building higher-fidelity, smaller-scale agent-based models, first learning the basics of how to implement a SIR model in this framework,[14] and then understanding how to represent various SIR parameters as random variables.[15] [16] [17]

Even more than 9 months after the project began, the study team continued to improve the model using emerging research. Specifically, the study team reviewed work that informed their inclusion of vaccines into the SEIR model,[18] updated methods to recalibrate the initial conditions,[19] and assessed impact of different strains of COVID-19.[20]

1.2 Laying the Groundwork

CAA's COVID-19 analysis was a response to a rapidly evolving global crisis. Because of this, the study team had little time to plan its internal organization, knowledge management, tools, and project management. Despite the lack of preparation time, study team leadership made a number of decisions in the initial stages of the project that enabled them to continue to expand their influence over the course of 14 months. These key decisions included the adoption of an Agile project management style, the division of coding and research into two sub-teams, the choice of a powerful open source coding language to power the models, the use of Git and GitLab to manage code, and the use of cloud-based systems to automate model runs.

Typical Army operations research projects start with a clearly defined objective and due date. Study teams then create a glide path to deliver the project on the assigned due date. Agile project management, instead, focuses on continuous interaction with customers and delivering minimum viable products as rapidly as possible while improving them over time. This management style enabled the study team to meet the ever-changing requirements of Army senior leaders, who, at the start of the project, did not know what kind of analytical support they needed or what CAA could do for them. The study team leaders spent a large amount of time interfacing with various customers, communicating the modeling capabilities of the study team to find new Army organizations that could benefit from CAA's modeling and analysis, and quickly tailoring products and analysis for new customers. The study team would not have been

[12] Smit et al. "Winter Is Coming" *International journal of environmental research and public health, 5634.*

[13] Cui, Zhang, Feng. "Influence of non-homogeneous mixing on final epidemic size." *Journal of biological dynamics 31–46.*

[14] Kiskowski. "A three-scale network model."

[15] Bar-On et al. "A quantitative compendium of COVID-19 epidemiology." 1-10.

[16] He et al. "Temporal dynamics in viral shedding and transmissibility of COVID-19." *Nature Medicine.* 672–675.

[17] Wikramaratna et al. "Estimating the false-negative test probability of SARS-CoV-2 by RT-PCR." *Euro Surveillance.*

[18] IHME, "COVID-19 vaccine efficacy summary"

[19] IHME, "COVID-19 Results Briefing, The United States of America"

[20] Galloway et al. "Emergence of SARS-CoV-2 B.1.1.7 Lineage"

able to respond rapidly to such a wide range of stakeholders without Agile project management.

Early in the project, the study team split into two efforts, research and model development. Members of the research team focused on understanding the rapidly emerging research by reading published journal articles and engaging with epidemiologists in the Army and in academia. The research team provided explicit direction to the model development team on what features should be included in the model and explained the mathematics required to develop those features. The model development team could then focus all of their efforts on building the model code.

The study team chose to use the R programming language to implement its COVID-19 forecast model. This open-source programming language offers the advantages of being available on unclassified government laptops and in cloud-based environments. The study team used functional programming techniques to organize the over 10,000 lines of code required to run the model so that it could be efficiently maintained and expanded over time. Additionally, the study team was able to leverage parallelization, which performs distributing functions to multiple computer cores in parallel. This parallelization significantly reduced model run time from about 4 hours to 1 hour.

The study team also chose to house the model in a sub-component of Army G-8's Army Resource Cloud called cloud PROgrogramming, Budgeting, and Execution (cPROBE). Housing the model in the cloud meant further freeing up analyst time because running the model did not consume a study team members' local computing power and because daily model runs in the cloud could be automated. This allowed the study team to focus on customer engagement and model improvement instead of generating daily outputs as study team membership fluctuated. Cloud computing is common in private industry, but it is seldom done on the Department of Defense (DoD) network due to network restrictions, which proved a continuous challenge throughout the project.

The study team also used GitLab, hosted by the National Geospatial-Intelligence Agency (NGA), to manage the model code. This is a standard practice in software engineering, but is not common in the DoD. GitLab provides a suite of version control tools to allow multiple people to improve model code simultaneously, incorporate or discard model excursions rapidly, and prevent user created version issues. This allowed the study team to develop model features during all stages of the project rapidly.

These five key decision made the study team much more efficient and responsive, which made rapid response to senior leader requests possible. With all of these tools and systems in place, the study team was typically able to answer a senior leader request in days or weeks, instead of months or years.

2 THE SEIR MODEL

2.1 Modeling Methodology

In support of the Army's need to understand the likely path of the pandemic, mitigate the most severe threats to the force's health and readiness, and prepare to support civil authorities, CAA developed a SEIR model to project future COVID-19 cases. The model runs for each U.S. county and 53 countries in the U.S Indo-Pacific Command area of responsibility. The model also has an option to consolidate multiple counties that span a single metropolitan area together to produce consolidated output for large cities like New York City. SEIR models are part of a set of standard epidemiological modeling approaches called compartmental population models. Compartmental population models decompose the totality of the populace into discrete compartments that characterize an individual's progress through disease phases. A set of ordinary differential equations dictates movement from one compartment to the next over time Figure 1 shows a representation of movement between compartments in an SEIR model. The compartments are:

Susceptible: Proportion of population at risk of becoming infected.
Exposed: Proportion of population incubating COVID-19, but not yet contagious.
Infectious: Proportion of contagious carriers of COVID-19, whether symptomatic or not.
Removed: Proportion of population for whom infection has ended; made up of fatalities and recoveries.

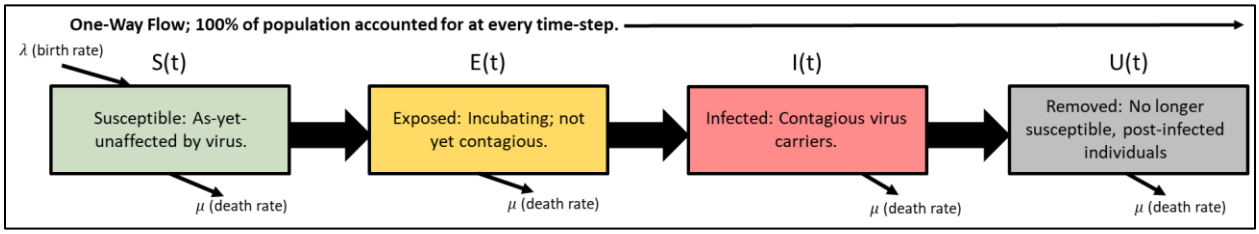


Figure 1. Graphical Depiction of an SEIR Model

CAA's model formulation obeyed assumptions standard to compartmental epidemiological models:

1. The population within the discrete compartments accounts for the totality of a geographically bounded population.
2. The total population intermixes uniformly regardless of compartment; geographically distinct populations do not intermix.
3. The size of every population is static, meaning there are zero births, deaths, and cross-border movements.
4. All members of the population are equally susceptible to infection.
5. The distribution of the infected compartment population among age cohorts is proportional to that of the total population.
6. Demand for hospitalization manifests only while an individual is infected.
7. SARS-CoV-2 is equally infectious for any given geographic population.

8. Infected individuals exhibit uniform behavior regardless of whether an individual is symptomatic or hospitalized.
9. Recovered individuals are immune and no longer susceptible.
10. The populace lacks widespread pre-existing immunity. The study team initially also assumed that there were no available effective vaccines, but later incorporated vaccine immunity.[21]
11. The disease has a 6-day-long asymptomatic, non-contagious incubation period and communal spread patterns, indicating the need for an exposed compartment.
12. COVID-19's contagious period is 14.3 days, concluded by either recovery or death.

The study team assessed the SEIR approach as suitable to reflect COVID-19 dynamics because of disease attributes the medical community identified during the first several months of the pandemic. Some of these assumptions weakened or even changed during the course of the project as the study team gathered updated information. The SEIR construct, however, proved adaptable enough that the study team could keep delivering useful forecasts even as new dynamics, such as mass vaccination, came into play.

2.1.1 Model Formulation

A system of ordinary differential equations governs a SEIR model. The derivatives represent how quickly the population moves through the SEIR compartments. Since the study team assumed the total population was static, the sum of the derivatives is zero. Since each variable represents a proportion of the population, the sum of the variables must always be one. Table 1 shows the ordinary differential equations for each compartment. Note, the variable for the removed compartment is $U(t)$; this is to avoid confusion with one of the key parameters, R.

Compartment	Equations		Simplified Equations
Susceptible	$\frac{ds}{dt} = \mu - \beta I(t)S(t) - \mu S(t)$		$\frac{ds}{dt} = -\beta I(t)S(t)$
Exposed	$\frac{de}{dt} = \beta I(t)S(t) - (\sigma + \mu)E(t)$	$\boldsymbol{\mu}$ **removed** →	$\frac{de}{dt} = \beta I(t)S(t) - \sigma E(t)$
Infectious	$\frac{di}{dt} = \sigma E(t) - (\gamma + \mu)I(t)$		$\frac{di}{dt} = \sigma E(t) - \gamma I(t)$
Removed	$\frac{du}{dt} = 1 - \left(\frac{ds}{dt} + \frac{de}{dt} + \frac{di}{dt}\right)$		$\frac{du}{dt} = 1 - \left(\frac{ds}{dt} + \frac{de}{dt} + \frac{di}{dt}\right)$

Table 1. SEIR Ordinary Differential Equations System Formulation

[21] CDC, "COVID-19 Frequently asked questions"

2.1.2 Model Initialization

To initialize a SEIR model, one must first assign parameter values (i.e., σ, γ, and β) and second assign the initial proportions in each of the four SEIR compartments (i.e., S_0, E_0, I_0, and U_0).

The parameters are defined as:

- Incubation factor σ: the inverse of the length of time (in days) before an exposed individual becomes infectious.
- Infectious factor γ: the inverse of the length of time (in days) before an infectious individual is removed.
- Transmission Factor β: the inverse of the mean time (in days) between successive exposures of susceptible individuals by a single infectious individual. This can be thought of as how contagious a disease is at a given time. The more commonly known version of this parameter is the effective reproduction rate, R_{eff}. $R_{eff} = \frac{\beta}{\gamma}$, which means β and R_{eff} are directly correlated and, for the purpose of communication, interchangeable.

Of these parameters, β—or R_0—is the most complex and has the most impact on the SEIR model results. For σ and γ, the study team used the Centers for Disease Control and Prevention (CDC) estimates. The study team continued to experiment with various values of σ and γ, only to find that the other values did not result in significantly better case-projection accuracy than the CDC's estimates. To estimate β, the study team relied on machine learning techniques, which are addressed in section 2.2.

To assign initial proportions to each of the four compartments (at $time = 0$), the study team interpolated values from the Johns Hopkins University (JHU) COVID-19 case data. JHU case data only provides daily numbers of total cumulative cases. Thus, all of the initial variable values ($S(t=0), E(t=0), I(t=0)$ and $U(t=0)$) originate with the daily cumulative cases for each county. Table 2 provides the equations for each initial variable.

Variable	Calculations
$S_j(t=0)$	$1 - \frac{Confirmed\ Cases_{t=0} + Vaccinated_{t=0}}{Region\ Population_j}$
$I_j(t=0)$	$\frac{(Confirmed\ Cases_{t=0} - Confirmed\ Cases_{t=-14})}{Region\ Population_j}$
$E_j(t=0)$	$\frac{\beta * I(0) * s(0)/\sigma}{Region\ Population_j}$
$U_j(t=0)$	$1 - S(0) - E(0) - I(0)$
Where j is the county from 1 to 3,155	

Table 2. Initial Conditions of the SEIR model

2.1.3 Time Horizon

The study team initially made predictions a full year (365 days) into the future, but reduced the modeled period to 120 days after it became clear that the pandemic was evolving too quickly to put any trust in long-term forecasts. Eventually, the study team focused on releasing predictions 30-45 days into the future, based on the needs of customers and the confidence of the study team.

2.1.4 Forecasting Stress on Hospital Infrastructure

One of CAA's more influential analytical outputs for Army senior leaders was its forecasts of requirements for medical infrastructure, which included the dates CAA expected demand for hospital and/or intensive care unit (ICU) beds to exceed supply in a given county. CAA produced these forecasts by applying age-weighted hospitalization rates to the predicted number of infectious individuals in each region over time. The study team considered integrating an additional "hospitalized" compartment and the associated parameters into the SEIR model itself, but chose not to. Such a method would have substantially increased the complexity of the model, but would likely not have resulted in significantly more accurate predictions, due to a lack of high-quality data on the parameters needed to represent a hospital compartment and the heightened difficulty of tuning a more complex model (particularly without access to high-performance computing). CAA made three simplifying assumptions to enable this methodological choice:

1. Hospitalization status does not affect the rate of interaction between the populations inside and outside the infectious compartment.[22]
2. Demand for hospitalization manifests only while an individual is in the infectious compartment.[23]
3. The distribution of the infectious compartment population among age cohorts is proportional to that of the total population.[24]

CAA used age cohort-aligned rates for non-invasive inpatient hospitalization and ICU hospitalization provided by the CDC, shown in Table 3.

[22] Enables fidelity of the SEIR model as previously characterized – absent this assumption, the model would require differentiated values for β for the populations at-large and hospitalized, respectively.

[23] Simplifies the overall problem – absent this assumption, the model would need to account for (e.g.,) time dispersion between the manifestation of the disease in an individual, his/her intent to seek medical care, and any complicating factors that might extend hospitalization after the acute infection has resolved.

[24] Enables the consideration of the total population as a whole – absent this assumption, the model would require a means to measure interaction rates between (e.g.,) demographic age cohorts; the limiting case of such a breakdown is an agent-based instantiation.

Age Group	Hospitalization Rate	ICU Rate
0-44	1.7%	0.37%
45-64	4.5%	1.3%
65+	7.4%	1.98%

Table 3. Medical Infrastructure Demand Rates by Age Cohort[25]

The remaining $I(t)$ proportion reflects individuals who do not require hospitalization. Using assumption 3, the study team applies these rates against the respective age cohort population in $I(t)$ to forecast demand for medical infrastructure. This produces a forecast for hospital and ICU bed demand by day for each region, which the study team then compared to the actual medical infrastructure available in the region.

2.2 Inferring β by Modelling R_{eff}

The study team found that β (closely linked to R_{eff}, the effective reproduction rate) was the most influential parameter in the SEIR model, and modeling it accurately was key to producing accurate forecasts of the spread of COVID-19. While the study team initially used a static R_{eff} value of 2.5, provided by the CDC, this produced uniform and unrealistically dire forecasts across the country; given that the rate at which the disease spread varied greatly by time and place, the study team quickly abandoned this approach. Instead, the study team developed a sub-model that predicted R_{eff} by day for each region of interest, using a wide range of data representing demographics, environmental factors, human behavior, and data on the pandemic itself.

This R_{eff} sub-model predicted future values of R_{eff} for each county. The study team used these predictions in the SEIR ordinary differential equations to create predictions of the size of various compartments at any time in the future. To find the best model, the study team focused on (1) error metrics, such as Poisson Deviance and R^2 from the various models, and (2) more importantly, the SEIR model performance with the β produced from the R_{eff} model. Section 2.4 discusses the study team's extensive validation and verification efforts on the output of the SEIR model.

The R_{eff} prediction model had two steps. First, the study team inferred the value of R_{eff} for every previous day of the pandemic in each modeled region. Second, the study team fit a regression model that predicted R_{eff} using past values and a wide range of other predictors.

2.2.1 Estimating Historical R_{eff} Values

To calculate historical R_{eff} the study team derived the equation for R_{eff} from the set of ordinary differential equations used in the SIR model—a model similar to SEIR, but without an exposed compartment. In equation (1) below, the study team took all variable values directly from JHU case data.

[25] CDC, "Planning Parameters for COVID-19 Outbreak Scenarios"

$$R_{eff}{}_{tj} = -\frac{\frac{dS_j}{dt}}{S_{tj} * I_{tj}} * \frac{1}{\gamma} \quad (1)$$

where j is the county from 1 to 3,155 and
t is the time from March 22, 2020 to the current day.

2.2.2 Modeling Future R_{eff} Values with the eXtreme Gradient Boosted Tree (XGBoost) Algorithm

After determining the values of R_{eff} for previous days, the study team used these values as a response variable to train a model that predicted values of R_{eff} using a cloud-based automated machine-learning tool called DataRobot. DataRobot allows users to input cleaned, pre-processed data with indicated response and predictor variables, and then trains hundreds of different models using dozens of different machine learning techniques. DataRobot then recommended the model with the lowest cross-validation error, and provided tools for examining, visualizing, and understanding it. The study team used this program to predict future values of R_{eff}.

During the development phase, the CAA study team found that models created using XGBoost Regression with Early Stopping consistently performed the best. Though the study team would periodically revisit this process to verify that the XGBoost models remained the most effective, the study team was able to save hours of time by requesting DataRobot only create three different specifications of XGBoost and selecting the best one for each model run.

XGBoost is a tree ensemble model, which means it builds many trees that all come together to produce one prediction for each row of data. Tree ensemble models are effective because a prediction from many trees is typically better than a prediction from one single tree, and many trees reduce the possibility of overfitting. Within tree ensemble models, XGBoost falls into a subclass of models called Gradient Boosting Machines (GBMs).

GBMs are a generalization of Freund and Schapire's AdaBoost (Adaptive Boosting) algorithm[26] to handle arbitrary loss functions. Similar to their better-known sister algorithm, Random Forest, GBMs fit individual trees to random samples of the rows and columns of the input data. While Random Forest uses the bootstrap aggregation (or bagging) technique, building all trees at once, GBMs build trees in series, fitting each successive tree to the residual errors from all the previous trees combined.

Another unique feature of GBMs is that they can use one of several different loss functions to build each tree. That is, for each tree, the GBM algorithm builds branches and leaves to minimize the residuals of the selected loss function. Because the R_{eff} data were derived from case counts, the study team chose the Poisson loss function.

$$L_i = R_{eff_i} - e^{\hat{y}_i}, \quad (2)$$

where $\hat{y}$ is the result of the tree regression, L_i is the Poisson loss function, and i is the record in the data set that corresponds to a unique county and day of pandemic combination.

[26] Freund and Schapire. "A decision-theoretic generalization of on-line learning and an application to boosting." *Journal of computer and system sciences*. 119-139.

Compared to other GBMs, the XGBoost tree algorithm is more efficient and more accurate. One of these optimizations is "early stopping," in which the model stops training earlier than usual if test-set predictive accuracy does not improve as it builds successive trees; beyond this point, the model would likely begin overfitting. In DataRobot's implementation of XGBoost Tree Regression, there are 24 parameters. DataRobot handles the computationally intensive task of tuning the parameters, although users can set alternative values. The study team used all of DataRobot's recommended parameter values tuned during the cross-validation model building process. Example parameters included maximum number of terminal leaves, maximum tree depth, number of trees, and learning rate. The number of trees and the learning rate are notable because they are key to preventing the GBM from overfitting.

2.2.3 R_{eff} Model Assumptions

The study team made the following assumptions when designing and implementing the XGBoost model.

1. Each randomly sampled row and column used to create each tree are representative of the entire data set.
2. R_{eff} follows a Poisson distribution.
3. Data that change from day to day (e.g., Google mobility data, state policy data, case trends) carries on last known values into the future. An alternative to this assumption would have been to develop a sub model to forecast all of these variable values into the future; however, this was too computationally intensive and would have baked in more uncertainty that would have been extremely complicated to dissect and communicate.

2.3 Source Data

The study team collected data from 19 public sources; all but three were open-source (i.e., publicly available). Thirteen of the 19 data sets were static and, thus, only required downloading once. These 13 data sets included U.S. Census Bureau tables covering county populations and age distributions[27], county geographic information from the National Oceanic and Atmospheric Association[28], state historical flu patterns from the CDC[29], numbers of hospital beds in each county from the World Bank[30], etc. The other six data sets were republished daily. These were:

- case data from JHU;
- case data from USAFacts;
- vaccination data from Our World in Data;
- state policy data from the COVID-19 U.S. State Policy Database;
- community mobility reports from Google; and

[27] United States Census Data; Customized Table. https://data.census.gov

[28] NOAA Integrated Surface Database Station History. https://www1.ncdc.noaa.gov/pub/data/noaa/isd-history.txt

[29] CDC Flu Incidence Data. https://www.cdc.gov/flu/weekly/weeklyarchives2020-2021/data/NCHSData09.csv

[30] ERSI No longer publishes this data set. https://coronavirus-disasterresponse.hub.arcgis.com/datasets/definitivehc::definitive-healthcare-usa-hospital-beds/

- testing data from the COVID-19 Tracking Project.

In order to ensure the model was using the most up-to-date version of the aforementioned data, the study team used various application programming interfaces (APIs) to automate the process of checking for updates and downloading the latest data from each source. Several data sources were hosted on GitHub, whose API does not work reliably on the Non-classified Internet Protocol Routing Network (NIPRNET), so the study team created mirror repositories on the NGA-hosted GitLab instance and used GitLab's extensive file API to check for and download updates.

2.3.1 Data Limitations

The study team had to address a few key limitations to the data. The COVID-19 case data were inconsistent across counties because of the heterogeneity in approaches to testing, treating, and reporting COVID-19 cases. JHU and USAFacts frequently revise historical observations to reflect emergent information and/or reclassification; on occasion, these revisions resulted in data showing negative new cases for a day in a particular county. There is also an observed cyclic component to the historical data in that reports corresponding to Mondays tend to show higher-than-anticipated new case counts; the study team assessed this phenomenon to be more likely due to workflow than to a characteristic of COVID-19. These limitations are all similar in that humans generally cause them and the study team was able to execute arithmetic corrections to these issues. For instance, to address the cyclical reporting patterns, the study team implemented a rolling average function to diminish the spikes and dips each week.

The model was also limited because many COVID-19 cases were simply not recognized because the carriers were asymptomatic. This affected case data and testing data more than other data sets. To address this issue, the study team explicitly forecasted reported cases—not all cases. This made hospitalization forecasts more accurate because (1) forecasting all cases would have added more uncertainty and error into the decomposition of the infectious compartment into hospital projections, thus making hospitalization forecasts less accurate; and (2) hospitalizations are more tightly correlated with reported cases than they are with all cases.

2.4 Model Validation and Model Improvement Strategy

Acknowledging the need to rapidly improve the model and to incorporate the latest research into the modeling strategy, the study team created a Development-Operations workflow. This allowed the study team to produce daily modeling projections while simultaneously working to improve model accuracy. Whenever the study team gained new insights from emerging research or colleagues shared new methods, the model developers would incorporate the ideas into the SEIR model. GitLab enabled the study team to keep a working version of the model producing output daily, while model developers could build and test modifications to the model on separate "branches" of code. These modifications would rapidly, but deliberately, go through a formal review process and would either be discarded or adopted as the working version of the model.

Part of this process was the development of a model registry and tools for automated error comparison. The study team built a set of functions that made it easy to track model

experiments: analysts simply had to choose a set of model parameters, and the code would: assign a unique model identifier (or retrieve an existing one from the model registry); generate and store predictions from previous weeks or months, using only the data that would have been available on those days; and calculate error metrics against past actual confirmed-case data. This made it easy to compare the past performance of different model specifications, including the study team's current production model. The study team eventually created 88 different versions of the SEIR model, accepting only 8 as official production models. Improvements were generally categorized as an attempt to improve the R_{eff} model—either through the inclusion of new predictor variables, the use of a new predictive model type, or the way that historical values of R_{eff} were calculated—or an attempt to improve the initialization of the model, either through new methods to calculate the initial conditions or changes in parameter values.

2.4.1 Error Metrics

The study team used four key error metrics to track and compare the performance of different model specifications. These four metrics were Mean Absolute Error, Mean Percentage Error, Mean Absolute Percentage Error, and Root Mean Squared Error. Each metric has various strengths and weaknesses; the study team used them in combination to determine the accuracy of a given model.

Mean Absolute Error (MAE) calculates residuals at various time horizons, without accounting for over or under predictions.

$$MAE_{i,j} = \left|predicted_{i,t+j} - actual_{i,t+j}\right| \quad (3)$$

$$MAE_j = \frac{\sum_{i=1}^{n} MAE_{i,j}}{n} \quad (4)$$

Where i is a given county or geographic region, j is a time horizon to evaluate measured in days, t is the current time since the start of the pandemic measured in days, and n is the number of counties or geographic regions modeled. This metric is simple to understand and calculate, but it could be difficult to interpret when comparing model performance across the entire country. This is because regions with very large populations tended to have high case counts and high absolute error, while regions with low populations tended to have low case counts and low absolute error. Thus, a few very large absolute error values could greatly influence the overall MAE, making the actual value somewhat difficult to interpret by itself, although it was still useful for comparing models.

Mean Absolute Percentage Error (MAPE) calculates the absolute percentage error in each prediction at various time horizons. Since this metric is in absolute terms, it is not possible to have negative $MAPE$, and lower $MAPE$ indicates higher accuracy. Unlike MAE, this metric is proportional, so high-population/high-case-count regions are not unduly influential, and the value is more intuitively interpretable. However, MAPE has the opposite problem: regions with very small case counts may have surprisingly high MAPE even with very reasonable predictions. For instance, a county with one real case and three predicted cases would have 200% MAPE, despite its very low MAE of 2. In addition, MAPE exaggerates the effect of over prediction: under predicting by a factor of two produces 50% $MAPE$, while over predicting by a factor of two produces 100% $MAPE$.

$$MAPE_{i,j} = \frac{|predicted_{i,t+j} - actual_{i,t+j}|}{actual_{i,t+j}} \quad (5)$$

$$MAPE_j = \frac{\sum_{i=1}^{n} MAPE_{i,j}}{n} \quad (6)$$

Mean Percentage Error (MPE) calculates the percentage error in each prediction at various time horizons. This metric shows if there is systemic over or under prediction, but because all errors are averaged, the MPE could be 0% despite the absolute error being extremely high. In addition, MPE suffers from the same asymmetry as MAPE in how it calculates over prediction and under prediction.

$$MPE_{i,j} = \frac{(predicted_{i,t+j} - actual_{i,t+j})}{actual_{i,t+j}} \quad (7)$$

$$MPE_j = \frac{\sum_{i=1}^{n} MPE_{i,j}}{n} \quad (8)$$

Root Mean Squared Error (RMSE) is another absolute, non-proportional measure of predictive error, but is difficult to interpret as a raw number, because it emphasizes large absolute error values to a greater degree than MAE does. The study team used this mostly for comparative purposes.

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{n} (predicted_{i,t+j} - actual_{i,t+j})^2}{n}} \quad (9)$$

2.4.2 Criteria for model adoption

The study team developed a standardized model performance comparison report that it used upon completion of a promising developmental model specification. For each model specification, the report calculated the four error metrics above at 4, 7, 14, 21, and 28 days out from the prediction date, using the automated prediction history generation tools discussed in section 2.4. The study team generated multiple sets of prediction history for the developmental model and averaged its performance at each time horizon, to account for day-to-day variability in accuracy. The study team could then compare the accuracy of the developmental model to the current production model in a consistent manner. There was no fixed standard for accuracy improvement that the study team used to accept the developmental model as the new production model, because the wide variety of error metrics, time horizons, and dates of prediction meant that—after the major modeling innovations made early in the study—developmental models were rarely, if ever, strictly better on all possible metrics. Instead, the study team considered all available metrics and made a decision based on the totality of the information.

2.4.3 Other Model Development Experiments

The study team was constantly developing model improvements; many of them specifically intended to address the limitations of the SEIR model. However, in most cases, making the model more complicated in an attempt to better reflect reality did not improve predictive accuracy. For example, the study team knew that not all COVID-19 infections were being reported, so the method of calculating R_0 and S_0 did not reflect the true number of removed or

susceptible people in a region at a given date. Reported data always reflected more people categorized as susceptible and fewer people categorized as recovered due to the non-reported infections. This was a negligible problem for the first few months of the project, but in the fall of 2020, the study team explored the idea of increasing the number of initial recovered to account for the missing infections. While the idea seemed promising in theory, it did not systematically improve model accuracy; thus, the study team only briefly implemented it in the production model. An imbalance in the amount of non-reported infections by geographic region and a lack of regionally specific data regarding non-reported infections may have caused this result.

Other improvements considered but not implemented included using data on mass public protest attendance in the R_{eff} model, accounting for non-homogenous mixing of different population groups inside a geographic region, accounting for potential re-infection, adding a hospitalized or quarantined compartment with reduced contact frequency for very sick individuals, accounting for multiple strains of COVID-19, and accounting for infections across county or region borders.

2.4.4 Comparison to other models

In addition to calculating error metrics to understand the absolute error in the SEIR model, the study team compared its work to leading publicly available models. The CDC ensemble model was one of the best-known models that produced forecasts for new COVID-19 cases at the county level. This model produced weekly forecasts instead of daily forecasts, so the study team aggregated their own model results at the weekly level for comparison. The study team's SEIR model had higher error based on all error metrics, but the differences in the error were relatively minor. The study team's average MAE at 21 days from prediction was approximately 35, while the average MAE for the CDC ensemble model was 30 cases per day. Similarly, the MAPE at 21 days from prediction, averaged across every county in the country for the study team's model, was typically 75%–80%, while the CDC ensemble model's MAPE was typically 60%–70% for these same predictions. Despite this slight disadvantage in accuracy, the study team's responsiveness to Army Senior Leaders, willingness to explain outputs and modelling decisions to them, and ability to adjust model inputs, outputs, visualizations, and reports to suit their desires made this product valuable across the Army and Joint enterprise.

2.5 Impacts of the Model

CAA's SEIR model had a major impact on the Army's response to the pandemic; it enabled the Army to respond to the national crisis while maintaining readiness in units across all components of the Army. The model informed and influenced approximately 10 different decisions and ongoing processes in organizations spanning the Joint and Army enterprise. The model also provided a common operating picture for senior Army leaders, up to and including the Vice Chief of Staff of the Army, helping them understand and respond to the COVID-19 crisis as it emerged and developed. The study team remained relevant in an ever-changing environment by responding rapidly to changing customer demands, being able to explain model results succinctly, and being honest about the limitations of the model. The large number of consumers of CAA's model output meant that there was not enough time for the study team to

document all the decisions the model informed, but even the impacts of which they were aware are quite substantial.

2.5.1 Impacting U.S. Army North (USARNORTH) Crisis Response Decisions

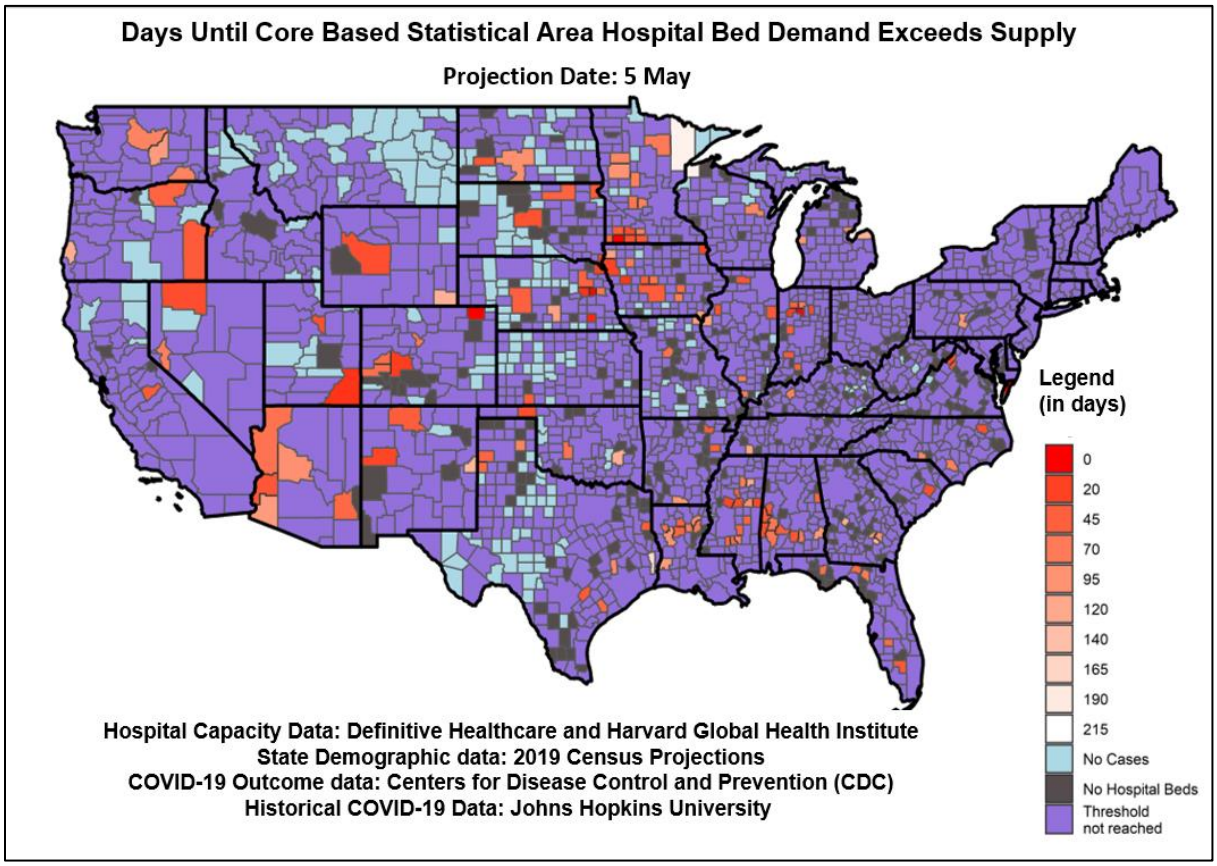


Figure 2. A heat map describing projected hospital stress in each county in the continental US according to the study team's SEIR model. The study team provided this product to USARNORTH to inform their decisions regarding field hospitals and troop mobilization.

One of the first tangible impacts of the SEIR model was helping USARNORTH and its subordinate units, the U.S. Army Corps of Engineers (USACE), and the Federal Emergency Management Agency (FEMA) decide where, when, and how to support areas of the United States experiencing hospital stress due to COVID-19. USARNORTH was in charge of the Army's defense support of civil authorities (DSCA) mission in responding to the pandemic, and regularly made decisions regarding the activation of National Guard units to support COVID-19 testing and/or the construction of field hospitals in pandemic hotspots. The study team added the hospital projections based on a request from USARNORTH and then provided custom visualizations that enabled leaders to quickly understand where there was a need for Army support for a local COVID-19 outbreak. Figure 2 shows an example of output provided to USARNORTH.

2.5.2 Impacting U.S. Indo-Pacific Command (USINDOPACOM) Strategic and Multi-national Decisions

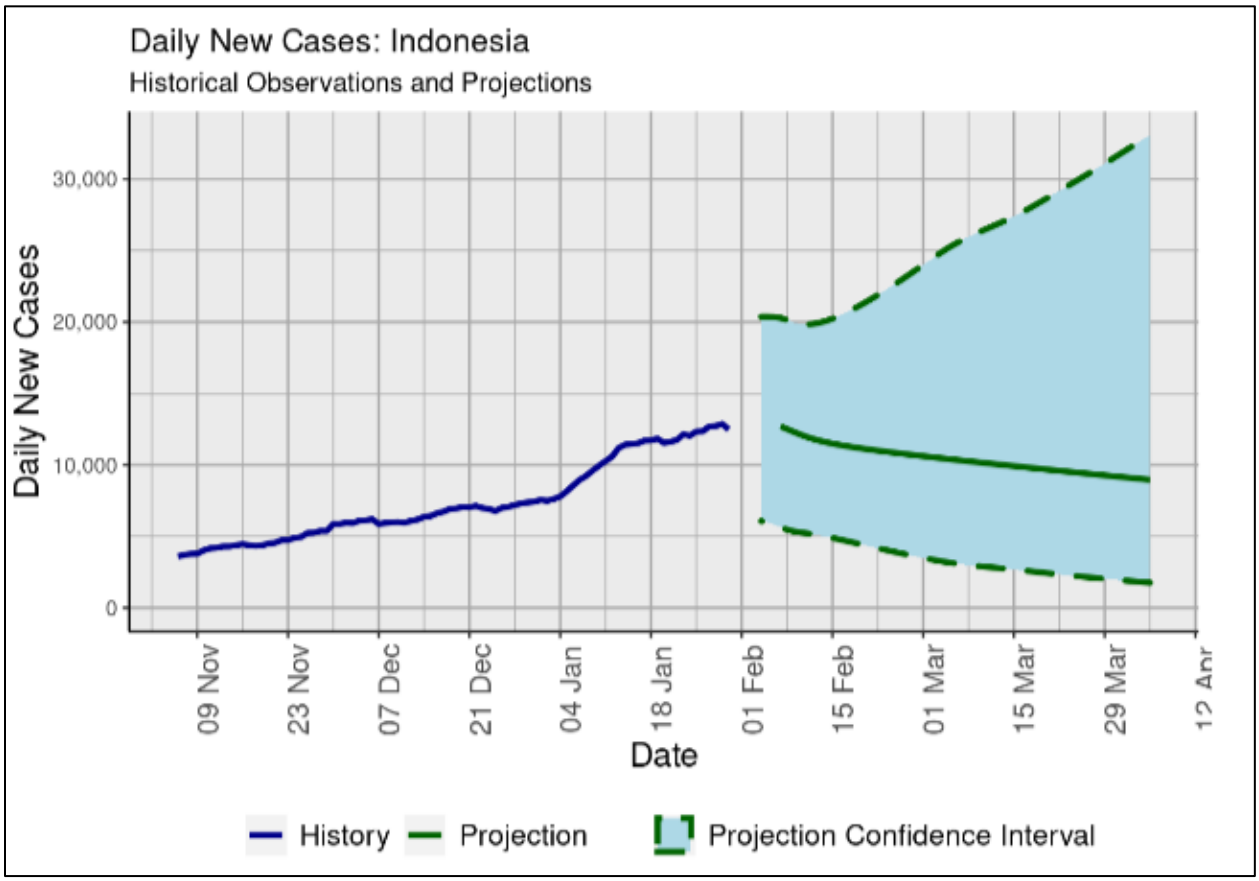


Figure 3. New case predictions according to the study team SEIR model provided to USINDOPACOM as part of the 2 February 2021 report.

USINDOPACOM became aware of the study team's modeling effort through CAA's Strategic Partner embedded at the command. They were immediately interested in the SEIR projection model and wanted to know if the study team could use it to forecast COVID-19 cases in each of the countries in the USINDOPACOM area of responsibility (AOR). The study team found new data sources and built them into the existing data processing and modeling code; due to the flexible nature of the model, it took just 3 weeks to modify the SEIR model to produce country-level projections for the USINDOPACOM AOR. The study team packaged these projections in a weekly report for the command's plans directorate, who used it to inform decisions about canceling or holding planned multi-national training exercises and key leader engagements at the flag officer and ambassador level. Figure 3 shows a sample of the output the study team provided to USINDOPACOM. The study team also shared these projections with U.S. Army Pacific (USARPAC), who used them to inform decisions about DSCA missions in the U.S. Territories of Guam and the Marianas Islands. According to MG Stephen Sklenka, USINDOPACOM J5, The study team was "instrumental in shaping the health, safety, and

security of the joint force operating across Indo-Pacific."[31] CAA contributed directly to USINDOPACOM's ability to accomplish its mission, working with its partners to promote development, enhance security, and provide humanitarian assistance.

2.5.3 Providing a Common Operating Picture across the Army Enterprise

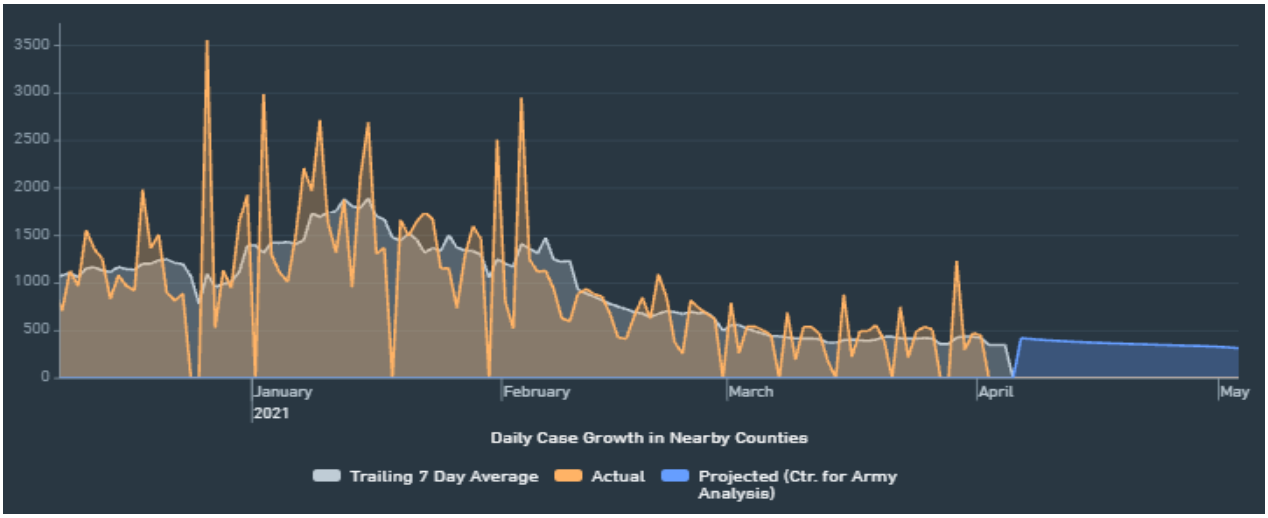


Figure 4. A snapshot of the Army Vantage COVID-19 dashboard, which displayed this projection of new cases in the Fort Bragg area on 5 April 2021.

The study team's initial customer relationship with the Deputy Chief of Staff, G-3/5/7 resulted in involvement with the Army COVID Campaign Plan (AC2P), which organized periodic briefs to all senior leaders in the Army, chaired by the Vice Chief of Staff and the Undersecretary of the Army. The study team routinely presented at these briefs to provide insight into trends, areas of concern, and the larger modeling community's thoughts on how COVID-19 would evolve in the future. Because of the study team's participation in these briefings, the Vice Chief of Staff mandated that the study team's SEIR model be the forecast model of record for the entire Army. This created an opportunity to bring CAA's model to Vantage, the Army's leader dashboard, which is accessible directly by all mid-grade leaders and above. The Vantage team created a COVID-19-specific dashboard that enabled users to view the forecast for future cases in the area around each installation, along with other information about the state of the pandemic. This could inform decisions at the battalion, brigade, and installation levels, such as operating status of base quality of life facilities, mitigation protocols to use during daily activities, and scheduling of in-person social events. These decisions, made every day across the Army, had significant impact on unit readiness. CAA's analysis enabled leaders to make informed decisions to balance the health and training aspects of readiness. Figure 4 shows an example of the output displayed in Vantage. Because the study team did not interact directly with Vantage users, there is no definitive record of what specific leader made what specific decision based on the SEIR forecast hosted on Vantage. However, user statistics showed that there were typically 750 unique users of the COVID-19 installation dashboard per week, indicating a strong usage of the SEIR forecast model. In addition, numerous other agencies such

[31] Major General Stephen Sklenka. Letter of thanks to Center for Army Analysis, April 6 2020.

as Army Materiel Command, the Defense Threat Reduction Agency, and the White House COVID-19 Task Force used the study team's SEIR forecast in conjunction with other forecast models to build situational awareness for their leaders.

2.6 Using the SEIR Model for Focused Studies

In the summer of 2020, there were regional surges of COVID-19, but many people began to believe that the worst of the pandemic was over. Senior leaders were still very interested in COVID-19 but were beginning to think beyond how they could support the national response. Instead, they were focusing on more on maintaining unit readiness in a COVID-19 environment. The connections the study team had made through its initial COVID-19 work, in combination with informal networking by other CAA personnel, led to new types of questions from organizations such as U.S. Army Central (USARCENT) and U.S. Special Operations Command (USSOCOM). The study team was able to use the SEIR model, with altered parameters or initial conditions, to answer specific policy and resource management questions from these organizations.

2.6.1 USARCENT

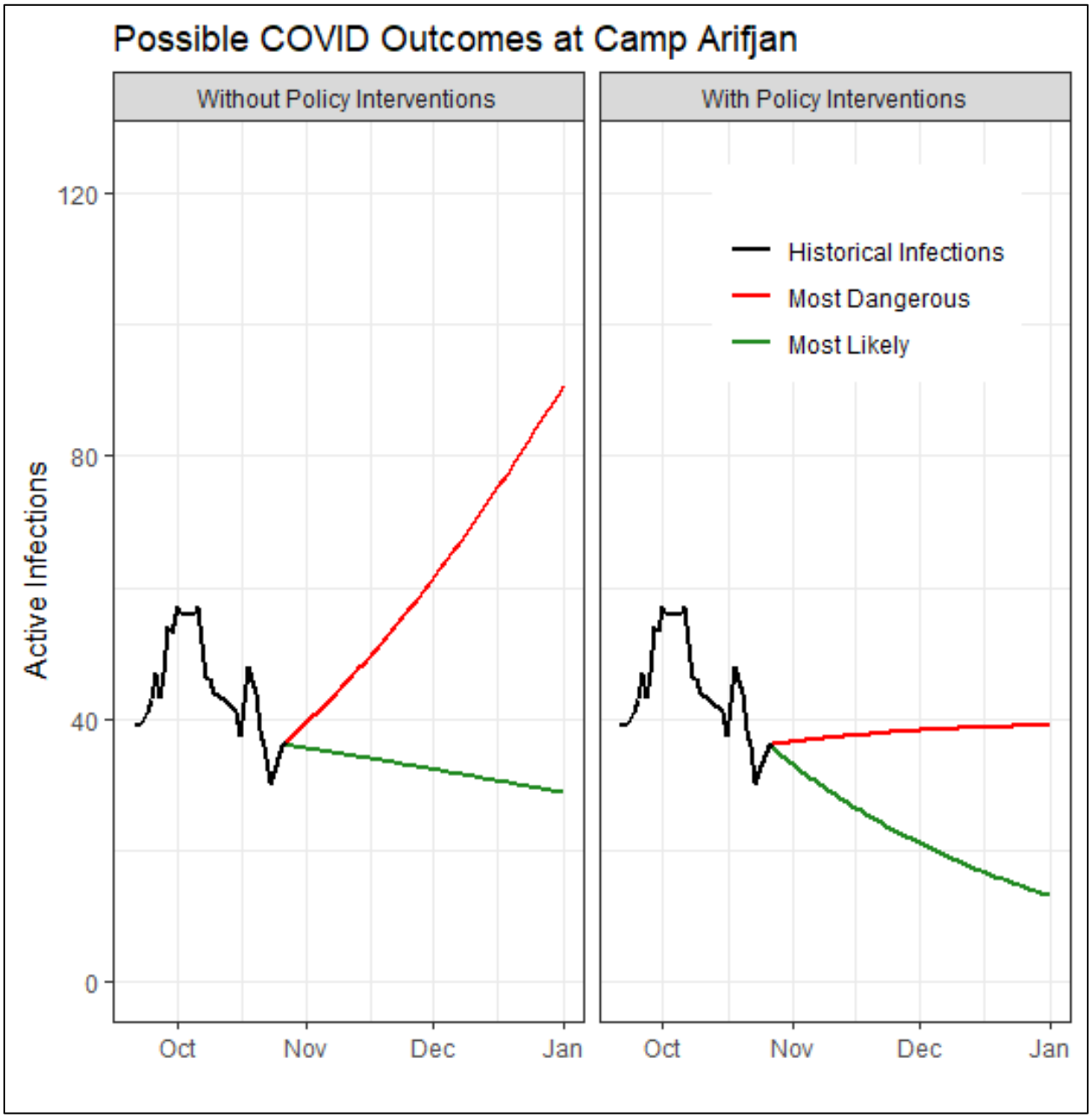


Figure 5. Results of analysis regarding how closures of base facilities would impact the most likely and dangerous COVID-19 outcome

Two different staff sections from USARCENT approached the study team looking for decision support analysis regarding different problems. In the summer of 2020, the Surgeon's office was looking to try to quantify the risk of COVID-19 spread if the base commander at Camp Arifjan were to re-open quality of life facilities such as the gym and coffee shops. The study team used information from the R_{eff} model discussed in section 2.2 to estimate the effects of specific interventions on the spread of COVID-19, and then built a probabilistic model that produced possible distributions of future COVID-19 caseloads on Camp Arifjan under different policies.

The study team concluded that re-opening closed facilities would not greatly increase cases in the "most likely," or median, set of potential outcomes, but it would expose the installation to a small possibility of a major surge in cases, which would be nearly impossible if the base commander kept COVID-19 restrictions in place. This analysis helped the base commander make decisions, weighing increased Soldier quality of life against potentially decreased short-term readiness. Figure 5 shows the topline results the study team shared as part of this analysis.

In the late fall of 2020, USARCENT G-4 requested CAA's assistance with a supply inventory problem. USARCENT had built an internal model to project the number of COVID-19 cases in its AOR that the Medical Treatment Facilities (MTFs) would treat. This model resulted an over-predicted future number of cases and, therefore, an extremely large oversupply of Personal Protective Equipment (PPE) in warehouses and at MTFs. The study team was able to use the SEIR model to create a probabilistic projection of future cases over a 4-month time horizon; combined this with DoD hospitalization data to determine the anticipated demand for care in terms of hospital days, ICU days, and isolation days; and then be translated this into projected PPE usage over a 4-month period, given a worst-case scenario for the spread of COVID-19. Even using this worst-case scenario, the study team recommended reducing the ordering rate of medical supplies in USARCENT by 95%–98%. The G-4 section used this information to brief the Deputy Commander of USARCENT, which led him to greatly reduce medical supply orders, freeing up resources and reducing waste.

2.6.2 USSOCOM

CAA offers continuous deployed support to USSOCOM. Because of the key placement of a forward deployed analyst from the organization, and that analyst's familiarity with the COVID-19 project, the study team had an opportunity to assist the deployed command with analysis regarding COVID-19 protocols. The command was interested in shortening the time that new arrivals spent in quarantine; depending on the country of origin and arrival method to the multi-national forward operating base, personnel would have to undergo a lengthy quarantine, which could account for almost 10% of a Soldier's deployment. However, the command did not want to increase the risk of importing an active case onto the base, decreasing unit readiness. The study team was able to use the SEIR model, in combination with information about the time lag between COVID-19 exposure and likelihood of a positive test, to quantify the risk of importing an active case onto the base under various quarantine lengths. The study team then repeated this experiment to account for vaccinated arriving personnel. Implementation of the study team's recommendations resulted in shortening the longest quarantines in all cases and shortening the quarantines for all vaccinated arrivals.

(THIS PAGE INTENTIONALLY LEFT BLANK)

3 AGENT-BASED MODEL

3.1 Introduction

In the summer of 2020, the SEIR model was stabilizing. The study team continued improving the model, but found it increasingly difficult to improve accuracy. Simultaneously, the study team began to receive more questions that focused on Army policy and resources and fewer questions about how the Army could support the national response to the pandemic. In addition to answering those specific questions, the study team began to think of new ways to help inform leader decisions regarding COVID-19. One area that seemed ripe for study was COVID-19 spread in smaller Army units, specifically during training or deployment. The study team anticipated that the Army would return to normal training and deployment cycles, and leaders would need to understand the COVID-19 risk associated with higher operational tempo and develop mitigation strategies to maximize unit readiness by balancing training requirements with health risks. To be prepared to answer these types of questions, the study team set out to build a new model of COVID-19 spread in training or deployment scenarios. The study team leveraged an existing collaboration with Lawrence Livermore National Laboratory (LLNL) and the Army Public Health Center (APHC) to combine modeling expertise, medical expertise, and high performance computing skills in the construction of this model.

3.2 Agent-based Model: Motivation and Challenges

The study team chose to create a new model to understand how disease would spread in a military training environment because the existing SEIR model had limitations that made it less useful for this purpose. Specifically, the SEIR model represents large, homogeneous populations and assumes that each individual behaves and experiences the disease identically. When modeling at a large scale, this is computationally necessary. However, to model Army units, the study team set out to incorporate individual attributes and behaviors. The study team knew that most Soldiers in training interact exclusively, but very frequently, with a small number of other Soldiers, while leaders typically interact with a very high number of other Soldiers in the training unit, spanning multiple small, tight-knit groups. Furthermore, clinical research has shown that the disease course—how long it takes to become contagious and how long contagiousness lasts—varies greatly from person to person.[32] A SEIR model cannot reflect either of these sets of dynamics, so the study team created an agent-base model, which would be able to incorporate individual disease and interaction profiles.

3.3 Creating Disease Profiles

Because of the small number of overall cases in a small- to medium-sized Army unit over the short time span of a training exercise, there is a wide range of possible outcomes, depending on random factors such as initial infections and the potential occurrence of a super-spreader event early in the exercise. To capture this randomness, the study team gave each agent in the agent-

[32] Walsh et al. "SARS-CoV-2 detection, viral load and infectivity over the course of an infection." *The Journal of infection. 357–371.*

based model a specific disease profile. A disease profile includes six different factors, five time-based and one binary, which create a timeline of each persons' illness.

1. Positive Generation Period: The amount of time from exposure until a person would test positive for COVID-19
2. Positive Duration: The amount of time that from when a person will first test positive for COVID-19 until when they will stop testing positive for COVID-19
3. Incubation Period: The amount of time from exposure until symptoms of COVID-19 present
4. Latent Period: The amount of time from exposure until a person is contagious with COVID-19
5. Infectious Duration: The amount of time from when a person become contagious with COVID-19 until when they stop being contagious.
6. Symptomatic Rate: The percentage of people with COVID-19 who present with symptoms.

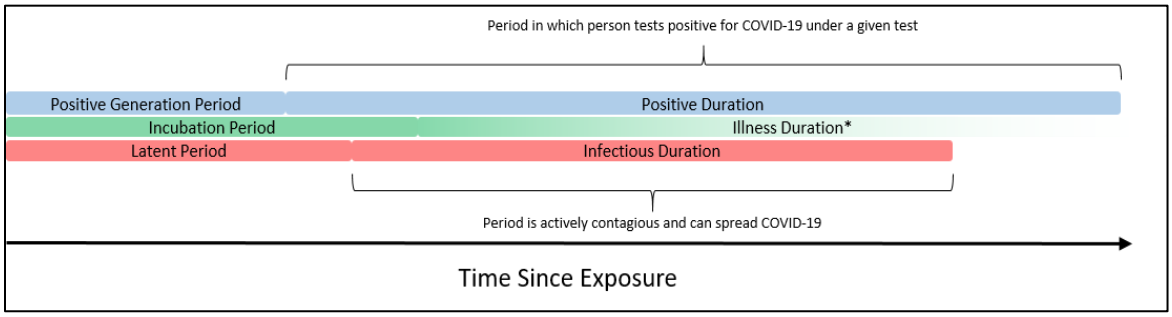


Figure 6. A sample disease profile *Note: The illness duration represents how long a person presents with symptoms and has no bearing on the spread of the disease under the study team's assumptions and so was not modeled

The study team gathered information from peer-reviewed scientific journals to determine distributions for each of the five periods described in Figure 6 as well as the symptomatic rate, or the percentage of cases that would present with any symptoms at all. The study team took some of these parameter values directly from clinical data published as part of a peer reviewed journal article; in other cases, they used raw data from published reports to create empirical cumulative distribution functions for the distributions. The study team treated the positive generation period and latent period as a joint distribution. Research showed that the viral load, a measure of how many virus molecules were in a given volume of bodily fluid, was strongly tied contagiousness. This same viral load is exactly what reverse transcription polymerase chain reaction (RT-PCR) tests measure to determine if a person has COVID-19. The study team, therefore, chose to model the viral load over time since exposure. They then used that output to determine what percentage of people would test positive and/or be contagious at a given time. The study team treated all parameters related to the physical illness associated with COVID-19 separately from the parameters related to returning a positive COVID-19 test or being contagious. This is because research showed that timing, presence, and severity were independent from how quickly a person became ill and not related to viral load.

3.3.1 Infectious Duration and Positive Duration

The infectious duration and positive duration are both strongly linked to viral load, so the study team chose to estimate these durations based on a model of viral load over time. To create a model of viral load since infectiousness, the study team created synthetic data based on the trends and variance in the clinical data from He and Ferreti.[33] [34] The study team considered a synthetic data point to test positive if the viral load was above a threshold amount. The same point was assigned a probability of being contagious based on a finding from Scola et al.[35] that showed a strong correlation between viral load and the probability of being contagious. The study team assigned points a status of either contagious or not based on a Monte Carlo simulation and their probability of being contagious. Summary statistics of how may synthetic data points were classified as testing positive and being contagious on each day after infectiousness were used to create a cumulative distribution for the infectious duration and positive duration. Figure 7shows the graph of the cumulative distribution functions (CDFs).

3.3.2 Latent Period and Positive Generation Period

The latent and positive generation periods have been less studied because it is rare to collect clinical data in the first few days after a COVID-19 exposure. Most people are not even aware that they have been exposed until some days after exposure. There has been little research that gives point estimates regarding the latent period,[36] [37] [38] [39] and even less regarding the positive generation period.[40] The study team used the information available from the latent period research and knowledge of the variance of individual disease profiles from the author, He.[41] The study team created a second synthetic data set of viral load values at discrete time intervals since exposure to COVID-19—note that this data set recorded viral load versus time from exposure, not the beginning of infectiousness. The study team then used cumulative summary statistics of that synthetic data set to determine empirical CDFs for the latent period and positive generation period in the same way that the distribution of the infectious duration and positive duration.

[33] He et al. “Temporal dynamics.” 672–675.

[34] Ferretti et al. “Quantifying SARS-CoV-2 transmission” *Science.*

[35] Scola et al. “Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards.” *European Journal of Clinical Microbiology & Infectious Diseases*. 1059-1061

[36] He et al. “Temporal dynamics.” 672–675.

[37] Li, et al. “Substantial undocumented infection.” *Science*

[38] Tian et al. “An Investigation of Transmission Control Measures.” *Science*

[39] Ma et al. “Epidemiological parameters of coronavirus disease 2019.” *medRxiv.* https://doi.org/10.1101/2020.03.21.20040329

[40] Kucirka et al. “Variation in False-Negative Rate SARS-CoV-2 Tests.” *Annals of Internal Medicine 173*. 262-267

[41] He et al. “Temporal dynamics.” 672–675.

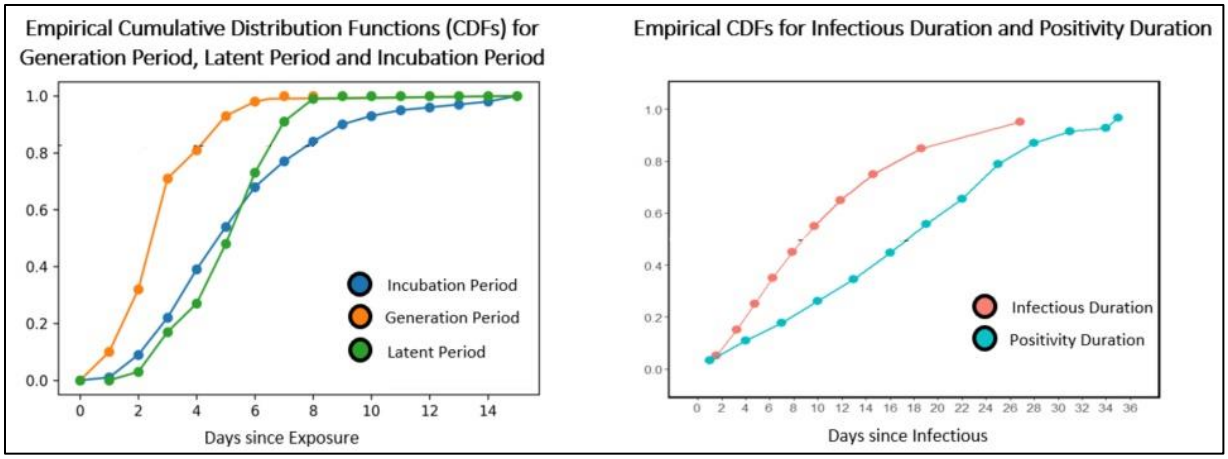


Figure 7 Empirical CDFs of the time based elements disease profile

3.3.3 Symptomatic Rate

The symptomatic rate is difficult to discern because of reporting bias. Equation (10) shows the most straightforward way to calculate a symptomatic rate of COVID-19.

$$symptomatic\ rate = \frac{reported\ positive\ cases\ with\ symptoms}{reported\ total\ positive\ cases} (10)$$

However, because people with symptoms are more likely to be tested, any symptomatic rate calculated using equation (10) would be artificially high. An additional problem is pre-symptomatic tests: many individuals report no symptoms at the time of their test, but some may report symptoms days or weeks later. Clinical records are unlikely to count these people as symptomatic because of the reporting time lag. Furthermore, symptoms do not present evenly across all demographics. Younger, healthier people typically tend to have fewer and less severe symptoms than those with underlying conditions. This led the study team to hypothesize that the symptomatic rate among the relatively young, healthy members of brigade combat teams (BCTs) would be quite low, which peer-reviewed academic research has confirmed.[42] [43] These sources shaped the study team's decision to set the symptomatic rate at 10% in their simulations.

3.3.4 Creating an Individual Disease Profile Using Monte Carlo Simulation

Once all the various parameter values and distributions have been set, the agent-based model uses three independent Monte Carlo simulations and inverse transform methods to construct a disease profile. The first simulation draw determines if the new COVID-19 infected person will become symptomatic based on the symptomatic rate. The second random draw uses the inverse transform method to determine the incubation (if applicable) positive generation and latent periods. The final random draw again uses the inverse transform method to determine

[42] Letizia et al. "SARS-CoV-2 Transmission among Marine Recruits during Quarantine." *New England Journal of Medicine*. 2407-2416.

[43] Kasper et al. "An Outbreak of COVID-19 on an Aircraft Carrier." *New England Journal of Medicine.* 2417-2426.

the infectious and positivity durations. These values can be stitched together to form a realistic disease profile for each agent in the simulation.

3.4 Creating an Interaction Network

One of the defining characteristics of any agent-based model is how the agents interact. The modeler can choose an entirely random interaction pattern, or a strictly structured one. In this model, the study team attempted to replicate the interactions of a combat unit. To do this, the study team used the unit Modified Table of Organization and Equipment (MTOE) to create an interaction network of all soldiers in the unit. An MTOE for a unit shows its organizational hierarchy, as well as the number of personnel at each level of command. The study team designated interaction levels between individual agents as high, low, or none, based on their place in the MTOE organizational hierarchy. High interaction generally refers to daily close contact. Low interaction generally refers to weekly close contact. The study team treated any contact rate that occurs less than weekly as no interaction.

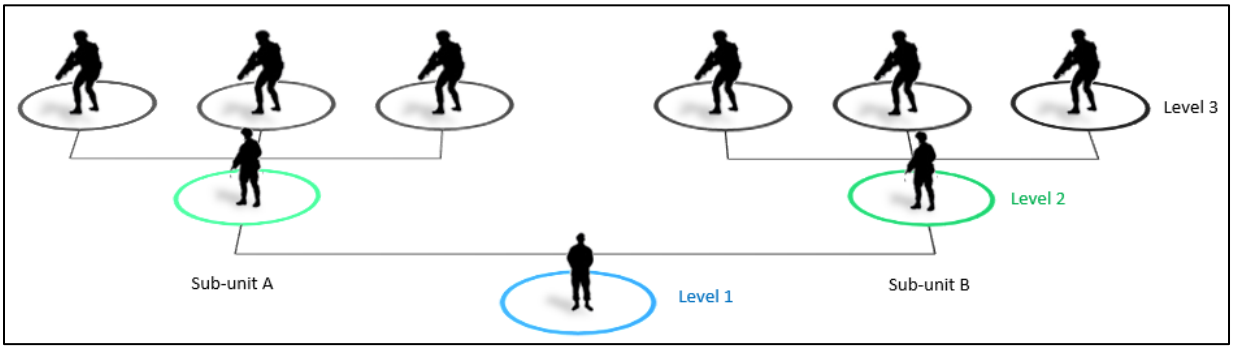


Figure 8. A sample unit hierarchy used to describe the process of building a unit interaction matrix

Figure 8 shows a small example of a unit hierarchy to illustrate the rules the study team used to form the interaction matrix. Take the level-2 Soldier in sub-unit A, for instance. The rules assign this individual a high rate of interaction with his/her parents (level 1), children (level 3), and siblings (sub-unit B, level-2 Soldier). The rules assign this level-2 Soldier in sub-unit A a low rate of interaction with his/her nieces and nephews (level 3 of sub-unit B) as well as his/her cousins, aunts, uncles, grandchildren, and grandparents, none of which appear in Figure 8. Although the exact set of rules that formed the actual interaction matrix is more sophisticated, they generally follow this construct. Figure 9 displays the study team's modeled individual interactions in an infantry BCT in matrix form and graph form.

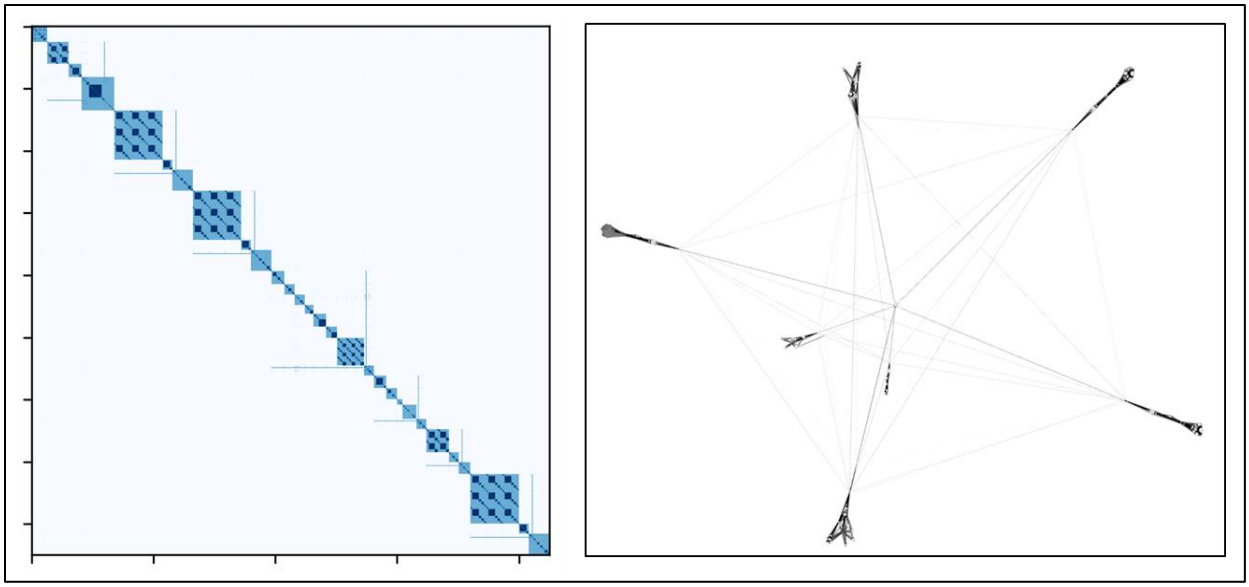


Figure 9. The matrix (left) and graph (right) form of the interaction of an infantry brigade combat team. In the matrix, each person is both a row and a column. The values of the matrix represent the contact rate between two individuals. In the graph, each person is a node and each connection between nodes is weighted based on the contact rate between the two individuals.

3.5 Initializing the Model

The final elements required to initialize the model are the initial COVID-19 infection status of each agent and the actual interaction rates that correspond to the high/low/none values discussed in section 3.4.

3.5.1 Initial COVID-19 Infection Status

Because each infected person has a known infection age and a unique disease profile, there is no need for an exposed classification. It is then sufficient to designate each agent as susceptible, infectious, or removed. The first step was to estimate the true "removed" percentage of any given Army unit—not just the cases detected through testing. To do this, the study team scaled data on reported cases among uniformed DoD personnel using the Institute for Health Metrics and Evaluation (IHME) estimates for the infection detection rate; the study team calculated this value by dividing reported infections by an estimate of true infections, derived from seropositivity surveys and hospitalization data. Next, the study team calculated the number of infected individuals using an instantaneous measurement of the infection detection rate along with the recently reported number of uniformed DoD personnel. The number of susceptible individuals in the model is the remainder. Equations (11), (12), and (13) show the specific equations used to calculate the percentage of agents assigned as susceptible, infectious, and removed.

$$S_0 = 1 - I_0 - R_0 (11)$$

$$I_0 = \frac{Total\ Uniformed\ DoD\ Cases\ in\ last\ two\ weeks}{Total\ Uniformed\ DoD\ Population} * \frac{1}{instantaneous\ infection\ detection\ rate} (12)$$

$$R_0 = \frac{Total\ Uniformed\ DoD\ Cases}{Total\ Uniformed\ DoD\ Population} * \frac{1}{cumulative\ infection\ detection\ rate} (13)$$

3.5.2 Interaction Rate

To initialize the model, the study team had to quantify the "high" and "low" interactions by assigning values that denote the probability of a contact between any pair of agents during each time step. The study team used the idea of an effective reproductive number to tune this parameter, using Kasper's study regarding the spread of COVID-19 aboard an aircraft carrier to determine a reasonable effective reproductive number to target.[44] The effective reproductive number during that outbreak was 1.4. The study team set the interaction rate on the agent-based model such that, aggregated over repeated simulations, the initial seed cases each produced between 1.25 and 1.50 cases in the population.

3.6 Running the Model

Once initialized, the model steps through time using Monte Carlo methods to determine if certain events such as infection transmissions and testing positive occur. In general, the model first uses Monte Carlo simulations to determine if an interaction between two people occurs during a given time step. If one of the people in the interaction is infectious and the other is susceptible, then an effective transmission occurs. Second, the model assigns each newly infected person a disease profile according to the above parameters and inverse transform methods. Thirdly, the model simulates COVID-19 tests based on one of two criteria. If an infected person becomes symptomatic according to their personal disease profile during a given time step, they will voluntarily submit for testing. They will be identified as COVID-19 positive after the time required to return a test result. If the mitigation policy in question designates that a surveillance test occur during that given time step, then personnel that are infected and would test positive at that time step according to their disease profile will be identified as a COVID-19 positive after the time required to return a test result. Finally, personnel are removed from the simulation and placed in a quarantine based on being a newly identified infected person, or having had a known interaction with a newly infected person in the past 7 days of the simulation. These people are place into quarantine for 14 days. These four primary tasks occur in each time step of the simulation and enable the extraction of the total number of infections and quarantines during the simulation.

3.7 Impact

The study team first used this model to answer questions for U.S. Army Forces Command (FORSCOM) related to the Combat Training Centers (CTCs). Brigade-sized units deploy to the CTCs to conduct a monthlong training exercise. This training event is typically the culmination of 1–3 years of training. The CTCs were closed for the first 5 months of the pandemic, but in the

[44] Ibid.

late summer of 2020, they re-opened for training. Although units continued to train at the CTCs for the fall and winter of 2020, there were significant issues with COVID-19 outbreaks.[45] Data from the FORSCOM Surgeon's office showed that, in the initial rotations, approximately 1% of Soldiers developed a symptomatic infection, suggesting a much higher number developed asymptomatic cases. Additionally, between 3%–10% of Soldiers had to quarantine due to being a close contact of a confirmed positive case during the rotation. These outcomes occurred even with all training units undergoing a 14-day restriction of movement (ROM) prior to travel and a 100% surveillance RT-PCR test 5 days prior to travel to CTC. The study team's analysis helped the FORSCOM surgeon's office and Deputy Commander understand the tradeoff between training benefits and health risks in their effort to provide trained and ready units to geographic combatant commanders.

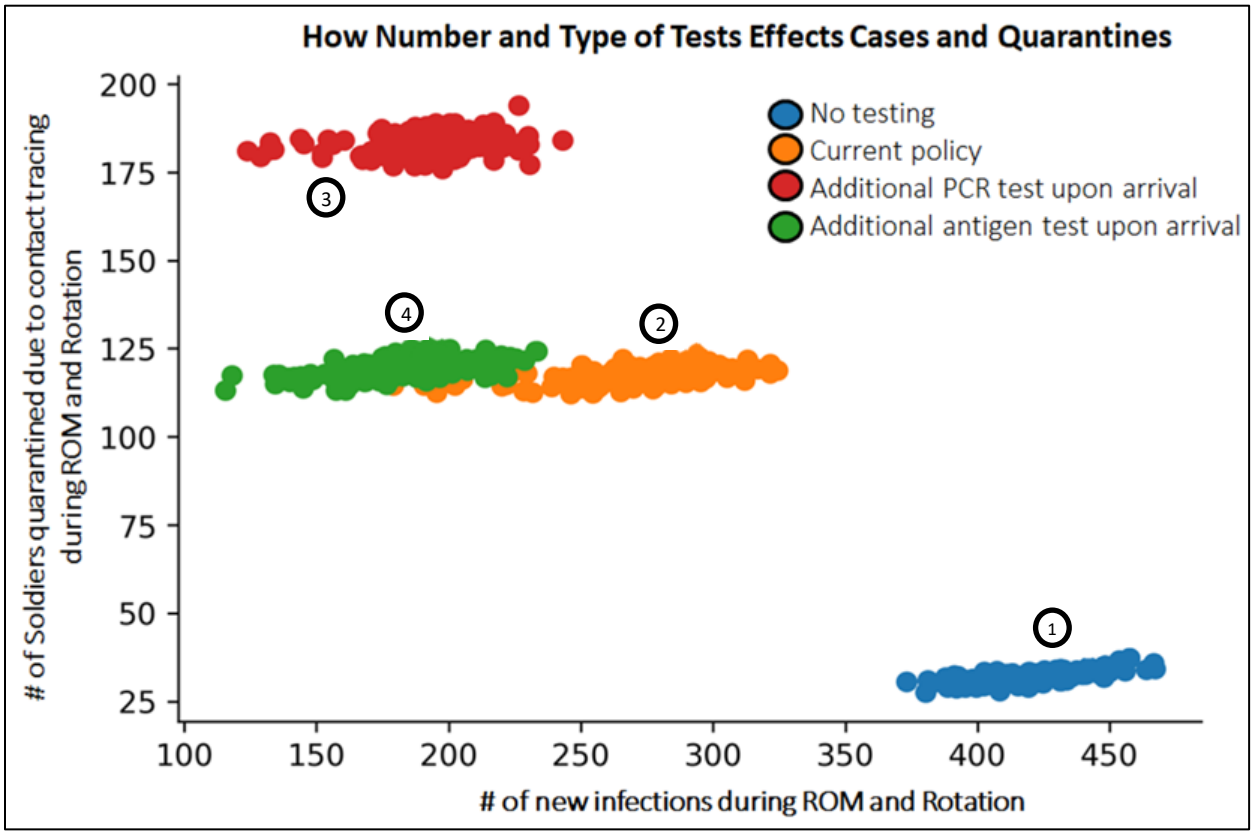


Figure 10. Simulation results of experiments considering different testing protocols for units prior to a CTC rotation. Adding a rapid antigen surveillance test (green dots) after arrival to the training center reduced new exposures without increasing quarantines when compared to the current protocol (orange dots); the numbers and colors correspond to the various testing protocols examined

The study team used the agent-based model to determine if changes to this testing protocol could reduce the number of infections without significantly increasing lost training time due to quarantines. APHC provided crucial information regarding different COVID-19 tests and Army

[45] Cox, "82nd Disputes Claims of COVID-19 at Training Center."

laboratory capabilities. The study team used this information to build two new mitigation strategies in the agent-based model and run experiments to compare their effectiveness to the current policy and a policy with reduced mitigation measures with regard to total COVID-19 infections and total quarantined Soldiers. The protocols were:

1. 14-day ROM, no surveillance testing (reduced mitigation policy);
2. 14-day ROM, one 100% RT-PCR surveillance test 5 days prior to traveling to the CTC (current FORSCOM policy);
3. Protocol 2, plus a second 100% RT-PCR surveillance test upon arrival at CTC;
4. Protocol 2, plus 100% rapid antigen test (less sensitive than RT-PCR test and faster return time) upon arrival at CTC.

LLNL analysts used a supercomputer to run hundreds of simulations of each protocol to generate a distribution of outcomes and estimate best case, worst case, and expected scenarios. The results of the simulation are in Figure 10. The study team recommended implementing a second rapid antigen test upon arrival (Protocol 4, green dots in Figure 10), which reduced COVID-19 exposure nearly as well as Protocol 3, but was much less costly in terms of training days lost to quarantine. Following the study team's brief, FORSCOM decided to implement the use of rapid antigen tests as part of their mitigation efforts.

Once the DoD started making COVID-19 vaccines available to Soldiers in large numbers, the study team also used the agent-based model to recommend a level of vaccinations in a training unit at which FORSCOM could safely discontinue surveillance testing without expecting an increase in total COVID-19 cases during the rotation.

In addition to helping FORSCOM develop improved COVID-19 mitigation policies, the study team used the agent-based model to help USSOCOM understand the risk of an outbreak on their base, given an initial case imported from deployers or local national contractors. This helped them decide on the level of mitigation efforts, such as requiring masks and social distancing in headquarters buildings, to use at their base.

Finally, the APHC leadership is in the process of becoming the owner of this model. They plan to use it as a framework to study larger and longer-term problems related to epidemics, such as the effects of various housing strategies on the spread of an infectious disease in a garrison environment.

(THIS PAGE INTENTIONALLY LEFT BLANK)

4 CONCLUSION

The study team faced an immense challenge in February of 2020: providing analytic grounding to senior leaders to use in making decisions in an extremely uncertain and ever-changing environment. The study team accomplished this by building two models: one that predicted how COVID-19 would spread through the general population at a large scale, and another that predicted how COVID-19 would spread through a military unit on a small scale. These models, their outputs, and the analysis communication done by the study team influenced countless decisions across Army and Joint commands at all echelons, on topics such as training, deployment, policy, supply management, and operations. This included small decisions, such as "should a battalion conduct, cancel, or postpone a unit social?" as well as much larger ones, such as "should a National Guard unit be activated and sent to operate a mobile field hospital in an area where hospital demand is expected to exceed supply in the next month?" Readiness was central to nearly all of these decisions. Leaders weighed the positive effects on readiness of training or morale against the negative effects on readiness of increased health risks in the short term, while planning defense support to civil authority missions to defeat the virus. With CAA's analysis, these leaders made informed decisions that maximized the overall, long-term readiness of the Army and the health of the nation.

To do all this, the study team developed predictive models that could be constantly improved and adapted to new information without pausing for major overhauls, and developed a workflow efficient enough that modeling capabilities could expand even as the study team staffing reduced. The study team did this with no prior experience in epidemiological modeling, no formal organizational structure at the outset, and a rotating cast of leaders and developers who had to balance competing priorities and missions. The study team used cutting edge tools and advanced analytic methods to do work that would have been impossible given the time and personnel constraints just a short time ago. The phrase "built the airplane in flight" is a common one in the Army, but it is very appropriate here. This study team not only built the airplane in flight, it developed a new and improved airplane in flight while delivering it ahead of schedule and under budget.

(THIS PAGE INTENTIONALLY LEFT BLANK)

APPENDIX A STUDY CONTRIBUTORS

A-1 STUDY TEAM

Study Director: Mr. Kyle Minor, MAJ Scott Lynch

Team Members: MAJ Chad Chapman, MAJ Maxine Drake, Mr. Collin Henley, MAJ Sandra Jackson, LTC Matthew Pacheco, MAJ Harvey Smith, MAJ Dusty Turner, Mr. Robert Ward, Mr. Michael Warme, and Ms. Michaela Zuber

Other Contributors: COL Josh Helms, LTC Kelly Ryan, Ms. Sandra Hatch, Mr. Dallas Kuchel, Mr. Duane Schilling, Ms. Lisa Kim, and MAJ Chris Ehlers

A-2 PRODUCT REVIEWERS

Mr. Russell Pritchard, Quality Assurance

A-3 EXTERNAL CONTRIBUTORS

Lawrence Livermore National Laboratory: Mr. Brian Spears, Mr. Jim Gaffney, and Mr. Luc Peterson

Army Public Health Center: Mr. Jake Ball and Ms. Stephanie Cinkovich

(THIS PAGE INTENTIONALLY LEFT BLANK)

APPENDIX B REQUEST FOR ANALYTICAL SUPPORT

REQUEST FOR ANALYTICAL SUPPORT

Performing Division: OA		***Account Number:*** 2020055	***FY:*** 2020
Acronym: CDAP		***Start Date:*** 10-Jun-20	***Est Compl Date:*** 30-Sep-20
Title: COVID-19 Data Analysis Project			
Category: Support for Strategic Analysis			***Method:*** In-house
Sponsor (e.g., DCS-G3) Name: DCS-G-35/SS			***Office Symbol:*** DAMO-SS
Phone:	***E-Mail:*** bradley.t.gericke.mil@mail.mil		***POC:***
Resource Estimates:	***a. Estimated Hrs:***		***b. Estimated Funds:***
Models to be Used:		***Product:*** Analytical Tool	

Description/Abstract:
CAA analytic support that addresses Army and Department of Defense leadership's highest priority issues concerning the global COVID-19 pandemic. Enables senior leaders across the Army enterprise to see, understand, and respond to rapidly-changing conditions resulting from the COVID-19 pandemic by providing timely, credible, and actionable analytics to help anticipate and mitigate risks to force and mission.

Study Director/POC Signature: ***Phone:*** 703-806-5382

Study Director/POC: Mr. Kyle S Minor

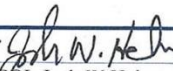
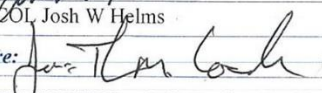
PART 2

Background/Statement of Problem:
In March 2020, the United States began to experience a global pandemic of the novel coronavirus-19 (COVID-19). In response, the Department of Defense - the Army in particular - leveraged resident technical expertise to model the spread of the disease and understand the impact to the force and the nation. CAA's study objective is to create informative models that give decision advantage to HQDA, Army components of Combatant Commands (COCOMs), other major commands (MACOMs), and other interested stakeholders. The models must support a decision makers ability to understand the impact of COVID-19 with respect to preservation of the force and implication to mission.

Scope:

Issues:

Milestones:

Signatures	***CAA Division Chief Signature:*** 	***Date*** 10 Jun 2020
	CAA Division Chief Name: COL Josh W Helms	
	Sponsor Concurrence Signature: 	***Date*** 11 Aug 2020
	Sponsor Name (COL/DA Div Chief/GO/SES): MG Bradley Gericke	

Print Date: 10-Aug-20

(THIS PAGE INTENTIONALLY LEFT BLANK)

APPENDIX C BIBLIOGRAPHY

Bar-On, Yinon, Ron Sender, Avi Flamholz, Rob Phillips, Ron Milo. "A quantitative compendium of COVID-19 epidemiology." 2020.

CDC Flu Incidence Data. Accessed December 2020. https://www.cdc.gov/flu/weekly/weeklyarchives2020-2021/data/NCHSData09.csv

"COVID-19 Vaccine Efficacy Summary." Institute of Health Metrics Evaluation, April 26, 2021, http://www.healthdata.org/covid/covid-19-vaccine-efficacy-summary

"COVID-19 Results Briefing, The United States of America." Institute of Health Metrics Evaluation, May 5 2021, http://www.healthdata.org/sites/default/files/files/Projects/COVID/2021/102_briefing_United_States_of_America_17.pdf

"COVID-19 Frequently asked questions." Centers for Disease Control and Prevention, Updated May 7, 2021. Coronavirus (COVID-19) frequently asked questions | CDC

"COVID-19 Pandemic Planning Scenarios." Centers for Disease Control and Prevention, Updated 19 March, 2021. COVID-19 Pandemic Planning Scenarios | CDC

Cox, Matthew. "82nd Airborne Disputes Claims of 'Mass Outbreak' of COVID-19 at Training Center." *Military.com,* February 5, 2021. https://www.military.com/daily-news/2021/02/05/82nd-airborne-disputes-claims-of-mass-outbreak-of-covid-19-training-center.html

Cui, Jingan, Yanan Zhang, and Zhilan Feng. "Influence of non-homogeneous mixing on final epidemic size in a meta-population model." *Journal of biological dynamics* 13 no. 1 (June 2018): 31–46. https://doi.org/10.1080/17513758.2018.1484186

Delamater, Paul, Erica Street, Timothy Leslie, Tony Yang, and Kathryn Jacobsen. "Complexity of the Basic Reproduction Number (R0)." *Emerging Infectious Diseases* 25, no. 1 (2019): 1-4. https://doi.org/10.3201/eid2501.171901

Ferretti, Luca, Chris Wymant, Cichelle Kendall, Lele Zhao, Anel Nurtay, Lucie Abeler-Dorner, Michael Parker, David Bonsall, and Christophe Fraser. "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing" *Science* 368. No 6491 (May 2020): https://doi.org/10.1126/science.abb6936

Freund, Yoav, and Robert E. Schapire. "A decision-theoretic generalization of on-line learning and an application to boosting." *Journal of computer and system sciences* 55 no. 1 (1997): 119-139.

Galloway, Summer, Prabasaj Paul, Duncan MacCannell, Michael Johansson, John Brooks, Adam MacNeil, Rachael MacCannell et al. "Emergence of SARS-CoV-2 B.1.1.7 Lineage — United States, December 29, 2020–January 12, 2021." *MMWR Morb Mortal Wkly Rep* 70. (2021): 95–99. http://dx.doi.org/10.15585/mmwr.mm7003e2

He, Xi, Eric Lau, Peng Wu, Xilong Deng, Jian Wang, Xinxin Hao, Yiu Lau et al. “Temporal dynamics in viral shedding and transmissibility of COVID-19.” *Nature Medicine* 26 (April 2020): 672–675. https://doi.org/10.1038/s41591-020-0869-5

Hethcote, Herbert, “The Mathematics of Infectious Diseases.” *SIAM Review* 42, no. 4 (December 2000): 599-653.

Hu, Hao, Karima Nigmatulina, and Philip Eckhoff. “The scaling of contact rates with population density for the infectious disease models.” *Mathematical biosciences* 244, no. 2 (2013): 125–134. https://doi.org/10.1016/j.mbs.2013.04.013

Jones, James. “Notes on R_0.” (Lecture Notes, University of Stanford, Stanford CA, May 1, 2007.)

Kasper, Matthew, Jesse Geibe, Christine Sears, Asha Riegodedios, Tina Luse, Annette Von Thun, and Michael McGinnis et al. “An Outbreak of COVID-19 on an Aircraft Carrier.” *New England Journal of Medicine* 383 (December 2020): 2417-2426

Kiskowski, Maria. “A Three-Scale Network Model for the Early Growth Dynamics of 2014 West Africa Ebola Epidemic.” *PLoS currents* 6 (November 2014). https://doi.org/10.1371/currents.outbreaks.c6efe8274dc55274f05cbcb62bbe6070

Kucharski, Adam, Timothy Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, and Rosalind Eggo. “Early dynamics of transmission and control of COVID-19: a mathematical modelling study.” *The Lancet Infectious Diseases* 20, no. 5 (May 2020): 553 – 558

Kucirka, Lauren, Stephen Lauer, Oliver Laeyendecker, Denali Boon, and Justin Lesser. “Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time since Exposure.” *Annals of Internal Medicine* 173 no. 4 (May 2020): 262-267 https://doi.org/10.7326/M20-1495

Lauer, Stephen, Kyra Grantz, Qifang Bi, Forrest Jones, Qulu Zheng, Hannah Meredith, Andrew Azman, Nicholas Reich, and Justin Lesser. “The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application.” *Annals of internal medicine* 172, no. 9 (2020): 577-582. https://doi.org/10.7326/M20-0504

Letizia, Andrew, Irene Ramos, Ajay Obla, Carl Goforth, Dawn Weir, Yongchao Ge, and Marcas Bamman et al. “SARS-CoV-2 Transmission among Marine Recruits during Quarantine.” *New England Journal of Medicine* 383 (December 2020): 2407-2416.

Li, Ruiyun, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman. “Substantial undocumented infection facilitates the rapid dissemination of the novel coronavirus (SARS-CoV-2).” *Science* 368 no. 6490 (May 2020). https://doi.org/10.1126/science.abb3221

Lipsitch, Marc, Christl Donnelly, Christophe Fraser, Isobel Blake, Anne Cori, Llaria Dorigatti, Neil Ferguson et al. “Potential Biases in Estimating Absolute and Relative Case-Fatality Risks during Outbreaks.” *PLoS Neglected Tropical Diseases* 9, no. 7 (July 2015). https://doi.org/10.1371/journal.pntd.0003846

Ma, Shujuan, Jiayue Zhang, Minyang Zeng, Qingping Yun, Wei Guo, Yixiang Zheng, Shi Zhao, Maggie Wang, and Zuyao Yang. “Epidemiological parameters of coronavirus disease 2019: a

pooled analysis of publicly reported individual data of 1155 cases from seven countries." *medRxiv.* Preprint (March 2020): https://doi.org/10.1101/2020.03.21.20040329

Massad, Eduardo, Marcelo Burattini, Luis Lopez, and Francisco Coutinho. "Forecasting versus projection models in epidemiology: the case of the SARS epidemics." *Medical hypotheses* 65, no.1 (2005): 17-22. https://doi.org/10.1016/j.mehy.2004.09.029

Nishiura, Hiroshi, and Gerardo Chowell. "The Effective Reproduction Number as a Prelude to Statistical Estimation of Time-Dependent Epidemic Trends." *Mathematical and Statistical Estimation Approaches in Epidemiology* (2009): 103–121. https://doi.org/10.1007/978-90-481-2313-1_5

Nishiura, Hiroshi, Natalie Linton, Andrei Akhmetzhanov. "Serial interval of novel coronavirus (COVID-19) infections." *International Journal of Infectious Diseases* 93 (March 2020): 284-286. https://doi.org/10.1016/j.ijid.2020.02.060

NOAA Integrated Surface Database Station History. Accessed November 2020. https://www1.ncdc.noaa.gov/pub/data/noaa/isd-history.txt

Scola, Bernard, Marion Bideau, Julien Andreani, Van Thuan Hoang, Clio Grimaldier, Philippe Colson, Philippe Gautret, and Didier Raoult. "Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards." *European Journal of Clinical Microbiology & Infectious Diseases*. 39 no. 6 (Apr 2020): 1059-1061. https://doi.org/10.1007/s10096-020-03913-9

Smit, Albertus, Jennifer Fitchett, Francois Engelbrecht, Robert Scholes, Godfrey Dzhivhuho and Neville Sweijd. "Winter Is Coming: A Southern Hemisphere Perspective of the Environmental Drivers of SARS-CoV-2 and the Potential Seasonality of COVID-19." *International journal of environmental research and public health* 17 no. 16 (August 2020): 5634. https://doi.org/10.3390/ijerph17165634

Tian, Huaiyu, Yonghong Liu Yidan Li, Chieh-His Wu, Bin Chen, Moritz Kraemer, and Bingying Li et al. "An Investigation of Transmission Control Measures During the First 50 days of the COVID-19 Epidemic in China." *Science* 368 no. 6491 (May 2020). https://doi.org/10.1126/science.abb6105

United States Census Data; Customized Table. Accessed April 2020. https://data.census.gov

Walsh, Kieran, Karen Jordan, Barbara Clyne, Daniel Rohde, Linda Drummond, Paula Byrne, Susan Ahern et al. "SARS-CoV-2 detection, viral load and infectivity over the course of an infection." *The Journal of infection* 81 no. 3 (September 2020): 357–371. https://doi.org/10.1016/j.jinf.2020.06.067

Wikramaratna Paul, Paton Robert, Ghafari Mahan, Lourenço José. "Estimating the false-negative test probability of SARS-CoV-2 by RT-PCR." *Euro Surveillance* 25 no. 50 (Dec 2020): https://doi.org/10.2807/1560-7917.ES.2020.25.50.2000568

(THIS PAGE INTENTIONALLY LEFT BLANK)