NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

VIRAL TESTING PROTOCOLS FOR U.S. NAVY SHIPS

by

Jamie C. Miller

June 2021

Thesis Advisor: Roberto Szechtman
Second Reader: Moshe Kress

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

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

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2021	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE
VIRAL TESTING PROTOCOLS FOR U.S. NAVY SHIPS

5. FUNDING NUMBERS

6. AUTHOR(S) Jamie C. Miller

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

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

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

The current best practices for minimizing the effects of COVID-19 are social distancing, the use of masks, the early detection of the virus among the population, and rapid quarantining of those who are infected. Early detection and rapid quarantining rely on testing, which is vital to the containment of the pandemic but also demands a large logistical effort in terms of testing sites, materials, labs, and staffing. This project develops tools to dynamically allocate these testing resources based on a model of the spread of COVID-19 on a Navy ship. Navy ships are highly compartmentalized environments and necessitate new models that relax the homogeneous-mixing assumption common to most epidemic models. We create a simulation model of the virus spread in the non-homogenous population, where people are spatially clustered into interconnected groups such as divisions or berthing on a Navy ship. The goal of this project is to develop testing protocols that minimize the impact of the virus to the ship's operational capabilities. To do so, we create a simulation-optimization model, which focuses on detection, not mitigation and response.

14. SUBJECT TERMS testing protocol, pandemic, viral testing	**15. NUMBER OF PAGES** 91
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

VIRAL TESTING PROTOCOLS FOR U.S. NAVY SHIPS

Jamie C. Miller
Ensign, United States Navy
BS, United States Naval Academy, 2020

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN APPLIED SCIENCE
(OPERATIONS RESEARCH)**

from the

**NAVAL POSTGRADUATE SCHOOL
June 2021**

Approved by: Roberto Szechtman
Advisor

Moshe Kress
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The current best practices for minimizing the effects of COVID-19 are social distancing, the use of masks, the early detection of the virus among the population, and rapid quarantining of those who are infected. Early detection and rapid quarantining rely on testing, which is vital to the containment of the pandemic but also demands a large logistical effort in terms of testing sites, materials, labs, and staffing. This project develops tools to dynamically allocate these testing resources based on a model of the spread of COVID-19 on a Navy ship. Navy ships are highly compartmentalized environments and necessitate new models that relax the homogeneous-mixing assumption common to most epidemic models. We create a simulation model of the virus spread in the non-homogenous population, where people are spatially clustered into interconnected groups such as divisions or berthing on a Navy ship. The goal of this project is to develop testing protocols that minimize the impact of the virus to the ship's operational capabilities. To do so, we create a simulation-optimization model, which focuses on detection, not mitigation and response.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

COVID, COVID-19	SARS-CoV-2
E	Exposed but not yet infectious (SEIR model)
I	Infectious (SEIR model)
IU	Infected Unaware (our model)
Q	Quarantines (our model)
R	Recovered, Removed
S	Susceptible
WSC	Winter Simulation Conference

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

This project researches the spread of the COVID-19 virus on U.S. Navy ships in an effort to find an optimal testing protocol to implement in order to maintain operational capabilities of ships and prevent outbreaks. Shipboard environments present a greater risk for viral spreading, especially for airborne viruses like COVID-19 because sailors cannot effectively social distance due to the close quarters and the ventilation system throughout the ship. The work environment is often stressful and this, combined with the lack of sleep, can weaken the immune system (Vera et al. 2014). We do not research mitigation techniques such as social distancing and wearing masks and instead focus on early detection through testing cycles to rapidly respond to outbreaks.

Epidemic modeling usually follows a standard SEIR flow model, which is a set of deterministic differential equations derived from empirical observations. In a SEIR model, entities move between susceptible (S), exposed but not yet infectious (E), infectious (I), and removed or recovered (R) using the inputted differential equations. Our model is a continuous time Markov model based on the SEIR flow where we also differentiate between infected unaware (IU) and quarantined (Q) sailors. We replace the (I) stage with two stages that define the difference between someone who is known to be sick or not. There is no (E) stage in our model, instead we assume sailors join the (I) stage immediately upon transmission and all infectious sailors are assigned a time until symptoms show. Infected unaware sailors are asymptomatic and continue to mix with the general population of the ship as normal, which is the main cause of the spread of the virus on board. Our Markov model is shown in Figure 1 with possible paths shown for false positive tests, false negative tests, and a standard infection route.

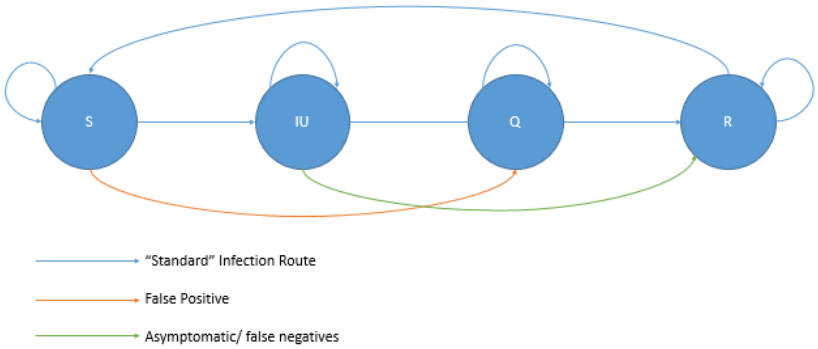


Figure 1. 4-Stage Markov Model

The Markov model we develop is implemented in the stochastic, agent-based computer simulation software SIMIO with the input parameters given in Table 1. These input parameters are developed from past studies on COVID-19 and are not ship-specific. For full robustness of the model we include pessimistic, optimistic, and nominal values to represent what a "best case" and "worst case" scenario might be. These parameters are put into the SIMIO model and determine how sailors move through the infection process.

Table 1. Model Input Parameters

Parameter	**"Best Case" Value**	**Nominal Value**	**"Worst Case" Value**	**Source(s)**
P(False Negative)	0.10			(Pray 2021) (Cashore et al. 2020)
P(False Positive)	0.001			(Pray 2021) (Cashore et al. 2020)
Time Until Symptomatic	~Poisson(3) Days			(Luo et al. 2020) (Cashore et al. 2020)
Time Until Recovered	~Poisson(12) Days			(Luo et al. 2020) (Cashore et al. 2020)
Time in Quarantine	10 Days	14 Days	14 Days	(Department of the Navy [DON] 2020) (Letizia et al. 2020)
P(Infectious at Initialization of Model)	0.01	0.01	0.25	(Letizia et al. 2020)
P(Symptoms Show)	0.821	0.65	0.55	(DON 2020) (Letizia et al. 2020)

P(infection transmission | contact between infectious and susceptible)	0.01	0.026	0.102	(Luo et al. 2020) (Cashore et al. 2020)
P(re-infection)	0.014			(Van Beusekom 2021)
Mixing Constant (κ_D)	15			
Mixing Constant (κ_{ND})	20			
Number of Drivers (For each ship)	15			

Naval ships vary greatly in sizes, which may affect how certain tests perform. To best grasp the importance of ship size, we run scenarios on three different ship sizes with an estimate of the number of sailors on board. The three ship scenarios are Destroyer (300 sailors), Cruiser (500 sailors), and Carrier (6,000 sailors).

Since ships are compartmentalized, we include the effects of non-homogeneous mixing between different job types, divisions, berthing, et cetera. To do so, we track a subpopulation of the ship "drivers" which have a different mixing constant, κ, than the non-drivers. This mixing constant represents the number of contacts a sailor will have in a given day and affects the time until a healthy sailor comes into close contact with an infectious sailor. The time until next infection is assumed to be exponential i.i.d and relies on the conditional probability of infection given a contact at time t (p_t), the mixing constant, the number of susceptible sailors (S), and the number of infected unaware (IU) sailors on board. The time until next infection is calculated in the simulation as:

$$\sim Exponential(mean = p_t * \kappa * S * IU)$$

Our model makes several assumptions based on the limitations of modeling in SIMIO. First, we assume that as soon as a sailor presents symptoms, they will self-report and get tested with a probability of 1. Second, we assume that there are no outside contacts/ no new infections brought on to the ship. If there are no infected unaware sailors on the

ship, then the time until next infection is infinity. Our model (and parameters) were adapted under the assumption that all social distancing and mask-wearing policies are followed to the best extent possible in the confined space, and the parameters are based on past studies that were not based on ship-specific data. Finally, we assume all sailors will eventually make a full recovery and do not model lasting effects.

We adjust the SIMIO model to represent three different testing protocols, which are compared to each other in terms of effectiveness. The three testing protocols are:

(1) Deterministic testing: every sailor on board is tested after a predefined, deterministic interval. For example, the entire ship is tested once a month.

(2) Subpopulation-based testing: a subpopulation of interest, such as the drivers, is tested more frequently than the rest of the ship's population. These intervals are also predefined and deterministic. An example is testing the driver subpopulation every week and the rest of the ship every month.

(3) Dynamic testing: as more sailors become infected and enter quarantine, the interval between testing events decreases and sailors are tested more often to try to reduce outbreaks. As fewer sailors become sick, tests become more infrequent.

With all three of these testing policies, we also have the possibility of a sailor to be tested immediately upon presenting symptoms instead of waiting for the next scheduled testing event. A true positive or a false positive test will result in a sailor being quarantined from the rest of the ship's population. A false negative will allow an infectious sailor to remain in the infected unaware group and continue to mix with other sailors and spread the virus.

We have three different models to represent the three different testing protocols. Each model is run with identical infection processes and simulates a 60 day-long period. We change input parameters for the time between tests under nominal parameters to determine the "best" testing interval for each ship and testing protocol. Then, using this testing interval, we vary the input parameters to represent the pessimistic, optimistic, and nominal parameter values. This results in 99 distinct simulations to be run. Each time a simulation is run, we conduct 50 replications. In the analysis, we use the following metrics:

- Average number of healthy drivers across all observations.
- Average number of healthy sailors across all observations.
- Average number of infected unaware sailors over all observations.
- The ranges (maximum and minimum recorded values by the simulation) of each of these metrics.
- The probability (given by a ratio of time) that the total number of healthy drivers or healthy sailors will dip below a dangerous level. There are four metrics within this: driver danger level, driver inoperable level, total ship danger level, and total ship inoperable level. These metrics are to ensure the ship is never at a dangerous level of personnel or inoperable due to lack of healthy sailors. For our purposes, the "danger" level is when there are less than 10 healthy drivers or less than 75% of the ship. The "inoperable" levels are when there are less than 6 drivers or less than 50% of the ship's sailors.
- The time-cost of testing: assuming a test will take 30 minutes out of a 16 hour work day, this is a calculation of what percentage of a sailor's day will be spent away from work to be tested. This is a conservative estimate of how long it might take a sailor to transit to the test taking site, receive the test, and wait for results.

First, we vary the time between testing events with the nominal input parameters. Each ship scenario is run through deterministic, subpopulation-based, and dynamic testing protocols with 5 experiments in each for a total of 45 experiments. The metrics are calculated and the major takeaway is that under nominal parameters, the time between tests does not have an effect on the operational readiness of the ship with our metrics. The thresholds for dangerous and inoperable levels of personnel are never met.

For the sake of our analysis, we choose the longest inter-testing period lengths and then compare the testing policies directly to each other using optimistic and pessimistic parameters. We also add in a fourth testing protocol to compare to the original three, called

reactive testing. Under reactive testing, the general population of the ship is never routinely tested and sailors are only ever individually tested when they self-report with symptoms.

The time between tests we choose to explore, which are independent of ship size, are as follows:

(1) A deterministic test every 4 weeks.

(2) A subpopulation based protocol which only tests drivers every 4 weeks and does not test the general population.

(3) A dynamic test every $\frac{2000}{Q+1}$ hours, where Q is the number of quarantined sailors.

(4) Under reactive testing, there is no routine testing event.

Using these testing intervals, we vary the input parameters and compare testing policies to each other within ship size scenarios. It is found that the nominal and optimistic input values returned similar, if not the same, results within the models. To help prevent the "worst case" scenario, the pessimistic scenario output values become our primary metrics of interest.

We conclude that the testing protocols do not have a significant impact on the operational capability of the ship. When the initial number of infected unaware sailors is low, barely testing the general ship population leads to a larger number of infected unaware sailors onboard but relatively few false positives. Testing at a high frequency results in a larger number of false positives and a lower number of infected unaware sailors. The total number of sailors with a positive test result is similar. Since both true and false positive sailors are quarantined, the operational impact on a ship is almost identical.

Therefore, independent of ship size, the testing protocol chosen is inconsequential as long as sailors self-report when symptoms show to be tested immediately and as long as the initial prevalence of infection on board is low. Simply testing the symptomatic sailors as they self-report is sufficient.

Because there is no significant difference between testing protocols, we conclude that a reactive testing protocol is an adequate choice for U.S. naval ships to implement.

This test is low on cost (monetary, time, and resource costs) and still has the same operational impact as a high-frequency, ship-wide, testing event.

This model is far from perfect, as it is impossible to create a simulation model that mirrors all aspects of ship-life accurately. Many social interactions were ignored in this project for simplicity. For example, only one subpopulation was created to represent ship drivers. In future research, more subpopulations should be created to better represent ship life and the interactions between groups could be better quantified. Sensitivity analysis should be conducted on the values of the mixing parameters, κ, as well as developing ship-specific input parameters using Navy data. A second, more sensitive testing procedure should be inputted into the model to minimize false positives. Finally, the model can be adjusted to include vaccinated sailors, outside interactions during port calls, and close contacts tracing.

At the conclusion of this project, we aspire to have provided a useful analysis of the three testing protocols. We are also hopeful that the methods we used can be applied for future situations with COVID-19 or be expanded to model future epidemics. It was found that the dynamic testing protocol results in the lowest number of infected unaware sailors, the highest danger driver threshold probability, and the highest time cost. The reactive testing policy results in the highest number of healthy sailors on board, no false positives, the lowest ratio of the driver danger threshold being met, the lowest time cost, and the highest number of infected unaware sailors on board. Depending on senior leader's priorities, either of these testing policies would be sufficient to implement on U.S. Navy ships.

References

Cashore M, Duan N, Janmohamed A, Wan J, Zhang Y, Henderson S, Shmoys D, Frazier P (2020) COVID-19 mathematical modeling for Cornell's fall semester.

Department of the Navy (2020) U.S. Navy COVID-19 standardized operational guidance version 3.0. NAVADMIN 266/20. https://www.navy.mil/Resources/NAVADMINs/Message/Article/2366564/us-navy-covid-19-standardized-operational-guidance-version-30/.

Letizia AG, Ramos I, Obla A, Goforth C, Weir DL, Ge Y, Bamman MM et al. (2020) SARS-CoV-2 Transmission among Marine recruits during quarantine. *N. Engl. J. Med* 383(25):2407–2416.

Luo L, Liu D, Liao X long, Wu X bo, Jing Q long, Zheng J zhen, Liu F hua et al. (2020) Modes of contact and risk of transmission in COVID-19 among close contacts. *medRxiv*. https://www.medrxiv.org/content/10.1101/2020.03.24.20042606v1.full.

Pray IW (2021) Performance of an antigen-based test for asymptomatic and symptomatic SARS-CoV-2 testing at two university campuses — Wisconsin, September–October 2020. *MMWR Morb. Mortal. Wkly. Rep.* 69.

Van Beusekom M (2021) Previous COVID-19 may cut risk of reinfection 84%. *CIDRAP*. Retrieved (June 7, 2021), https://www.cidrap.umn.edu/news-perspective/2021/04/previous-covid-19-may-cut-risk-reinfection-84.

Vera DM, Hora RA, Murillo A, Wong JF, Torre AJ, Wang D, Boulay D et al. (2014) Assessing the impact of public health interventions on the transmission of pandemic H1N1 influenza a virus aboard a Peruvian navy ship. *Influenza Other Respir. Viruses* 8(3):353–359.

ACKNOWLEDGMENTS

I would like to extend my thanks to my research advisors, Dr. Roberto Szechtman and Dr. Moshe Kress, for their support, encouragement, and grace over the past several months of working together. I would also like to acknowledge the entire Operations Research faculty here at the Naval Postgraduate School for their exceptional job overcoming the hardships of online learning during this difficult time.

Finally, I want to thank my family and friends who motivated me to complete this research and were always willing to help in any way possible.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

SARS-CoV-2 (COVID, or COVID-19) infection has been the defining feature of 2020 and of 2021 thus far. This global pandemic has permeated every aspect of our lives and day-to-day activities, reducing many to work from home virtually and social distance. Understandably, working from home is not an option for many of those in the armed forces. The United States Navy's mission areas explicitly require naval forces capable of winning wars, deterrence, and forward deploying (US Navy's Values and Missions). To do so, the United States Navy must remain operational and take every measure to do so.

Current measures to prevent COVID-19 are social distancing, the use of masks, and early detection of the virus leading into rapid quarantining of those who are infected. Sailors (and Marines) on U.S. naval ships are at a greater risk for viral infection due to the stressful work environment, inability to effectively social distance due to close quarters and the ventilation system (Vera et al. 2014). All of these factors contribute to higher susceptibility for airborne viruses such as COVID-19. As sailors are unable to effectively and totally socially distance on a ship, much of the countermeasures are dependent on early testing of the population, especially the asymptomatic population. It is estimated that anywhere between 35% and 45% of COVID cases will be asymptomatic, meaning an infectious person will continue to interact with a healthy population and thus the virus may spread easily (Department of the Navy [DON] 2020) (Pray 2021). Because of this, it is imperative a good testing protocol be implemented to prevent ship outbreaks.

We present a case study to describe how ships have been handling COVID prevention during deployment, which is based off a real ship's operations but has been generalized:

Starting about two weeks before deployment, ship personnel are not allowed to leave and no contractors or transfers are allowed onboard. Everyone is tested pierside before deployment, but this is the only mass testing event for the ship. Once underway, a two week quarantine is required of new sailors joining the ship. On port calls, no liberty is allowed and no contractors are allowed on board. If contractors must come aboard for

maintenance, keep “contaminated” areas of the ship secured to personnel until sanitized with bleach. Any sailors working with contractors must quarantine for two weeks. At any time if a sailor on board feels ill, they provide a list of every contact they had spent more than 15 minutes within 6 feet over the past 96 hours. This entire list is tested and quarantined until testing results return. Watch bills and berthings are adjusted to accommodate these close-contact quarantines. Throughout the deployment, face masks are required onboard, the gyms are closed or limited to one person at a time, and there are limits on the number of people on mess decks and other common areas. Six feet of spacing required throughout the ship.

To effectively build a testing protocol we rely on a Markov model, which we simulate in the program SIMIO. We do this by first creating a baseline model similar to historical viral models such as the SEIR model (explained in depth in chapter II) and specify it to U.S. Navy Ships to the best of our abilities. Due to the compartmentalized nature of ships between berthing, watch rotations, and divisions, we implement a way to account for non-homogeneous mixing between sailors. We also implement a way to track a subpopulation of interest; which can be expanded upon to track multiple subpopulations as desired in the future.

It is important to note here that a perfect model of any epidemic is impossible to create. We do our best to make this model as robust as possible while representing the real-world dynamics of shipboard life. Every model comes with two types of uncertainty: unknown parameters and random outcomes based on known parameters. This model contains both types of uncertainty, although the second type regarding random outcomes is represented within the stochasticity of the model. Minimizing the first type of uncertainty is difficult as many COVID-19 parameters are still unknown. We describe how we acquired the parameters we used in Chapter III, but it is important to note that the parameter uncertainty limits our abilities to perfectly model this system.

The background on epidemic modeling and a literature review is provided in Chapter II. We then present our baseline model and parameters for the homogenous case developed from the SEIR model in Chapter III. Also included in Chapter III is the expansion of our model to represent Navy ships with non-homogeneous mixing,

subpopulation tracking, and the implementation of three different testing procedures. In Chapter IV we provide the analysis of these alternative testing procedures and conduct analysis around input parameters. Finally, we draw conclusions and close the thesis in Chapter V.

At the end of this study we present an analysis of three different testing protocol alternatives. We are hopeful that these methods can be applied for future situations with COVID-19 or be expanded to model future epidemics.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND AND LITERATURE REVIEW

There are many factors to consider when choosing how to model an epidemic, as models range in complexity and scale. Tools can be as simple as a compartmentalized model and as complex as microsimulations, similar to what we are presenting in this paper. In his keynote speech at the Winter Simulation Conference (WSC) on December 15th, 2020, Stephen Chick discusses the lessons learned from past outbreaks and how the model chosen has drastic effects on the results that guide policy making (Chick 2020). The models he discusses start at the most basic, which assume homogeneity in the population, and range through individual patient-level simulation models.

Chick outlines several popular ways to model outbreaks and healthcare challenges. The simplest and most used tool in healthcare is a decision tree, where expected value decision making is used to assess the best choice to make. Usually decision trees assume homogeneity within the population—that is, each patient has the same risk of infection and recovery. One of the main flaws with decision trees is that there is no time component, which is remedied by using a Markov Model. This tool steps through time to model the flow of patients through each "compartment" of the model using predefined flow rates. The Markov model is incremented deterministically, whereas the continuous time dynamics model allows population changes to states or compartments according to continuous flow rates. An example of a continuous time system model is the SIR model, where a patient (or group of patients) moves between susceptible (S), infected (I), and removed (R) states. Tracking the state changes in population over time is the key advantage of this model. Better yet, the stochastic Markov chain model builds on the continuous time system model with flow rates for the system linked with the probability of an event. Finally, Chick introduces the patient level simulation, or micro-simulation, that is used to model specific events, such as an interaction between two people or a single patient's course through the disease process. Each of these tools increases in complexity and adds in new factors such as immunocompromised populations, heterogeneous population mixing, social interactions, and stochastic uncertainty. Many papers, a few of which are outlined

below, have been published on the subject of epidemic modeling, resource allocation, and the effects of social mixing constraints.

Starting with the simple compartmentalized model, Resource Allocation for Epidemic Control Over Short Time Horizons demonstrates a general compartmental epidemic model "to maximize the number of [new infections] averted over a fixed time horizon" (Zaric and Brandeau 2001). This paper focuses on resource allocation and uses heuristics to reach a solution. As mentioned, compartmentalized models are very general tools and there is often a need for models that can be incremented deterministically, which gives way to Markov Models, often SEIR. One advantage of the SEIR, SEIRS, and SIS models is that they track the movement of an entity through the cycle of the disease and also track the overall population's movement between categories using underlying differential equations. The solutions result in a deterministic model, which are not adequate for estimations with smaller population sizes and do not represent true variability of an epidemic (Sanchez and Sanchez 2015). Therefore, stochastic models have been developed, called Continuous Time Markov Models.

As mentioned by Dr. Chick in his keynote address, one of the most common Continuous Time models for an epidemic is the SEIR model, where entities move between susceptible (S), exposed but not yet infected (E), infected (I), and removed either through recovery or death (R) (Sanchez and Sanchez 2015). Each time increment, the rate of movement between compartments is calculated to represent a continuous-time model. The SEIRS, SIS, and SIR models are an extension of the SEIR, where the SEIRS allows recovered patients to become susceptible again. In the simpler SIS model, infected agents re-join the susceptible population immediately after recovering whereas the SIR does not model the time between exposure and infection.

Many recent studies using the SEIR model (or an extension), such as Spatial Resource Allocation for Emerging Epidemics: A Comparison of Greedy, Myopic, and Dynamic Policies, use epidemic modeling to allocate intervention resources in anticipation of the spreading disease (Long et al. 2018). The SIS model used in Optimal Control of Epidemics in Metapopulations demonstrates an example of resource allocation when an epidemic is spread throughout "different but interconnected regions" (Rowthorn et al.

2009). Their research uses a spatial structure of populations and has a goal of minimizing the number of new infections with respect to economic and epidemic constraints. The Influence of Non-Homogeneous Mixing on Final Epidemic Size in a Meta-Population Model expands on the SIR model to research the difference between homogenous and heterogeneous mixing (Cui et al. 2018). This paper challenges the assumption of homogeneous mixing and how contacts within the same sub-group or the variation of activities between subgroups affects the model's accuracy if heterogeneous mixing is not accounted for. The Effect of Social Mixing Controls on the Spread of Smallpox- A Two-Level Model presents one such way to account for social mixing. Although focused on Smallpox, the two-level social model created to account for an individual moving between a household and a secondary location (such as a school, a place of work, et. cetera) through a general meeting site like public transport or crowded streets (Kress 2005). This paper focuses on possible response strategies we have seen implemented in the past months, such as social distancing, closing schools, and reducing capacity on public transportation.

A recent COVID-19 model published by researchers at Cornell University is a hybrid between the stochastic SEIR and a Python simulation model. COVID-19 Mathematical Modeling for Cornell's Fall Semester uses both an excel model of initial testing estimates and a Python compartmental simulation built off of standard SEIR model to predict the outcomes for student and faculty return to campus before the Fall 2020 quarter (Cashore et al. 2020). This model accounts for outside mixing with the non-university population and assumes random testing every five days to create a prediction on infection rates, quarantine capacity, and overall effect of student return to campus. Of course, a campus is not nearly as close-quartered as a ship. In 2014, Vera et al. uses a daily time series model to evaluate the transmission of H1N1 on board a Peruvian naval ship with data from an outbreak in 2009. The authors use a stochastic epidemic model to prove heterogeneous mixing has an effect on viral transmission and that the military population in particular is prone to respiratory infections due to ventilation systems, close proximity in confined spaces over long periods of time, and high stress (Vera et al. 2014).

At the beginning of 2020, there was a COVID-19 outbreak onboard the Diamond Princess Cruise Ship off the coast of Japan. As a result of this outbreak and the daily test

results of the passengers, several papers were published to model the outbreak of passengers and crew and to estimate the transmission potential and asymptomatic rates of the (then novel) Coronavirus. Mizumoto and Chowell approach the issue using time-series data and a discrete-time integral equation and estimate model parameters with Monte Carlo Markov Chain (Mizumoto and Chowell 2020). Mizumoto, Kagaya, Zarebski, and Chowell used the daily time series testing results to create a confidence interval on the proportion of asymptomatic individuals aboard (Mizumoto et al. 2020). Rocklov, Sjodin, and Wilder-Smith use a SEIR model to estimate the transmission rates both with and without the ship's countermeasures in place. They used differential contact rates among groups and found the reproductive number onboard was four times higher on the ship than in the epicenter of a city, due to the population density of the ship (Rocklöv et al. 2020). The outbreak on board the cruise ship was helpful for initial estimates of transmission rates and asymptomatic proportions since it was one of the first closed-systems with high levels of data available at the beginning of the pandemic.

In 2015, Sanchez and Sanchez published a paper on the NPS Pandemic Model as a new model to infectious diseases. This model is a discrete-event, stochastic, agent-based simulation model as an extension of the SEIR model to preserve variability and incorporate different stochastic event impacts. Implemented in the computer program Ruby, this tool produces individual outcomes without the need for interconnectivity graphs between every agent. It also is flexible to incorporate different stochastic behavior such as social distancing and testing protocols. Presented at the WSC in December 2020, Testing-Based Interventions for COVID-19 Pandemic Policies uses the NPS pandemic model to assess testing-based interventions for COVID-19 management. The authors create two models, one to simulate the impact of COVID viral load on secondary infections and one to model the potential spread of COVID through a population with test-based interventions (Regnier et al. 2020). This ongoing study researches when and how often to test a population to minimize the effect of the virus, given the potential high viral load of asymptomatic infections.

A similar paper focusing on asymptomatic screening for COVID-19 mitigation, also presented at WSC 2020, works on determining the optimal testing cycle length to

minimize the final prevalence of infection at the end of a given time period (Lin et al. 2020). The authors pool multiple samples together and perform group tests in both linear and square arrays, which reduces the number of tests but also increases false negative rates. The two group testing policies are compared and the ongoing study hopes to identify the optimal time between testing cycles for minimal infection.

At the conclusion of his keynote address at the WSC 2020, Dr. Chick reiterates that every stochastic model, especially those concerning epidemics, have uncertainties. There is always stochastic uncertainty as well as systemic uncertainty due to unknown parameters (Chick 2020). The best mathematicians can do is to use the research preceding them to create the most accurate version of the model as possible, which is what we attempt to do in the next chapter.

THIS PAGE INTENTIONALLY LEFT BLANK

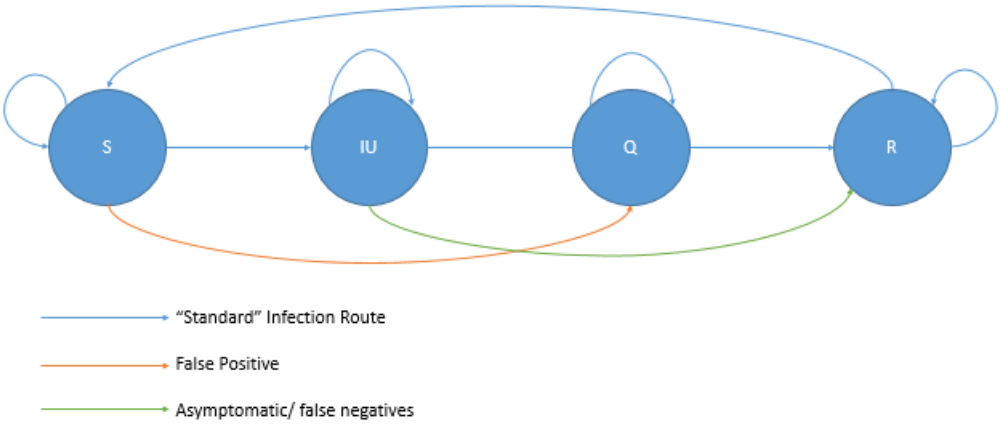
III. METHODS AND MODELS

A. BASIC MODEL

The first model we build and describe is the basic infection model that does not account for non-homogenous mixing and only implements a deterministic testing schedule. In Chapters III.C and III.D, we describe how we built upon this model to account for non-homogeneous mixing and three different testing protocols.

1. Model Dynamics

Our model follows a modified SEIR flow, where sailors move between "stages" as they progress through time; see Figure 1. The testing delay is the time between testing events, which is a user input and can range from 24 hours (testing every day) to 30 days (testing once a month).



A visual description of movement through the 4 stages of our model. Each stage and potential path through the model is described in the text.

Figure 1. 4-Stage Markov Model

Susceptible (S): Sailors are considered healthy; there is no change to routine or mixing with other people on the ship. It is assumed that each person is tested after the testing delay.

Infected Unaware (IU): Sailors have become infected with the disease, but do not know yet. They will either be administered a test according to the testing delay or a test if and when symptoms show. Since the sailors in this stage are either asymptomatic or unknowingly sick, they continue to mix with other people on the ship normally. If a sailor in this stage never tests positive, they will continue to stay in this stage throughout their entire illness and will eventually move to the recovered stage.

Quarantined (Q): A sailor is quarantined if they are confirmed to have the virus. They can be quarantined if they test positive during a routine test or after symptoms show after becoming infected. This includes false positive tests from the susceptible population as well as true positive tests from the infected unaware population. A quarantined sailor does not mix with the rest of the ship's population.

Recovered (R): Sailors are recovered after their disease has come to an end. In this model, we assume all sailors will make a full recovery. It is possible that a sailor can stay in infected unaware the entire course of the illness if symptoms never show or if false negatives occur. More likely, a sailor will move to the recovered stage after completing their time in quarantine. After a sailor is recovered, it is possible to return to the susceptible population for a secondary infection.

Note the difference in our model from the standard SEIR flow. We split the (I) stage into infected unaware and quarantined in order to differentiate between sailors who are infected and know it or not. There is no (E) stage in our model because we assume sailors will become infected unaware as soon as a transmission happens.

Each sailor is assigned a status label to indicate if they are sick, healthy, quarantined, or recovered. This label is a modeling characteristic that allows us to track the proportion of sailors that are healthy, infected unaware, or quarantined at any given time. As sailors move between the virus stages, their label is updated. For example, any sailor moving from susceptible to infected unaware will have a status change from "healthy" to "infected unaware." When they move to quarantined and later recovered, their labels will read "quarantined" and "recovered," respectively.

2. Parameters

The model parameters are assigned a nominal value as well as best and worst case scenario values. Each parameter is either based on previous studies, official DOD publications, or our own mathematical reasoning. We describe the parameters for the general model first in Table 1.

Table 1. Parameter Values and Sources

Parameter	"Best Case" Value	Nominal Value	"Worst Case" Value	Source(s)
P(False Negative)	0.10			(Pray 2021) (Cashore et al. 2020)
P(False Positive)	0.001			(Pray 2021) (Cashore et al. 2020)
Time Until Symptomatic	~Poisson(3) Days			(Luo et al. 2020) (Cashore et al. 2020)
Time Until Recovered	~Poisson(12) Days			(Luo et al. 2020) (Cashore et al. 2020)
Time in Quarantine	10 Days	14 Days	14 Days	(DON 2020) (Letizia et al. 2020)
P(Infectious at Initialization of Model)	0.01	0.01	0.25	(Letizia et al. 2020)
P(Symptoms Show)	0.821	0.65	0.55	(DON 2020) (Letizia et al. 2020)
P(infection transmission | contact between infectious and susceptible)	0.01	0.026	0.102	(Luo et al. 2020) (Cashore et al. 2020)
P(re-infection)	0.014			(Van Beusekom 2021)
Mixing Constant (κ)	20			

We wish to make our model parameters as robust as possible, to best represent the scenario. However, data surrounding the infection rates on U.S. naval ships are not available to the public. In order to keep this research both as accurate as possible and

unclassified, we use parameters developed from other COVID-19 studies and reports. Many of these aforementioned parameters come from the same process as the mathematicians modeling the spread of COVID on Cornell's campus (Cashore et al. 2020). The sources for each parameter is listed in Table 1 in the corresponding row and each will be elaborated on.

First, we list the false positive and false negative test results for our model. The PCR sensitivity is reported to be about 90%, meaning the false negative rate is 10%. Similarly, the false positive rates are reported to be about 0.1% (Cashore et al. 2020) and (Pray 2021).

Luo et al.'s work on close-contact transmission and infection rates in Guangzhou estimates the proportion of asymptomatic cases, the time until symptoms show, the time until fully recovered, and the probability of infection given a close contact (Luo et al. 2020).

Letizia et al. conducts a similar study on the spread and transmission rates among young adults, in particular in Marine recruits. Letizia et al. gives insight to military-aged men and women's susceptibility to the virus. In particular, we use the time in quarantine, the initial infection probability, and the probability that symptoms show that was reported in this research (Letizia et al. 2020). After a sailor reaches the recovered stage, they may re-enter the susceptible bucket for a re-infection with probability of 1.4% (Van Beusekom 2021).

As we explain in chapter III.A.1, the sailors move between susceptible and infected unaware via a pure birth-death process, so we need to model the time until the next infection in our model. To do this, we use the minimum between the time until the next testing event and the time until next infection. The time until next infection is exponentially distributed and depends on a mixing constant, κ, which may vary based on what "group" or division a sailor belongs to. In this basic model, it will not vary as we assume homogeneous mixing at first. Varied mixing constants are implemented in the non-homogeneous model in chapter III.C.

We assume random interactions between two groups of people, the susceptible group (S) and the infectious group (IU), so that the possible number of contacts between

these groups is S*IU. Under the assumption that the times between contacts are i.i.d. exponentially distributed, the time until the first contact is

$$\sim Exponential(mean = \kappa * S * IU)$$

where the constant κ is the contact rate between any pair of sailors. (Note that κ also accounts for multiple interactions between the same pair of people.) We assume a baseline value of κ=20, meaning that each sailor contacts 20 people per day, on average. In reality, however, the value of κ varies based on the job a sailor performs on the ship, their division, or some other subgrouping method. This issue is addressed in Chapter III.C.

Once this interaction happens according to the exponential distribution, there is a conditional probability that the interaction results in transmission of the virus. We present three different values for this conditional probability as a pessimistic, nominal, and optimistic parameter based on the work from Cornell (Cashore et al. 2020) and the research regarding contact risk of transmission among close contacts (Luo et al. 2020). We denote this probability of transmission at time *t* as p_t . The time until next infection becomes

$$\sim Exponential(mean = p_t * \kappa * S * IU)$$

3. Assumptions

Our model makes several fundamental assumptions based on the limitations of SIMIO and the unknown parameters. First, we assume that as soon as a sailor becomes symptomatic, they will self-report to get tested with a probability of 1.

We assume no outside contacts in this model, so in the case there are no infected unaware sailors in the system, the time until next infection is set to infinity.

Our model (and parameters) are adapted under the assumption that all social distancing and mask-wearing policies are followed to the best extent possible in the confined space.

Finally, we assume all sailors eventually make a full recovery and do not model lasting effects.

4. Infection and Testing Processes

There are several underlying "processes" that are implemented in this model to ensure it mirrors the real-life infection dynamics as closely as possible.

Routine Testing process. The routine testing process simulates the testing event that occurs every user-inputted "testing delay" period. Initializing the model begins this process and after every testing delay (for example, 48 hours) each sailor in the (S) and (IU) stages is tested. A true negative test of a healthy sailor in (S) will keep them in (S). A false positive test of a healthy sailor in (S) will mistakenly move them to (Q). A true positive test of a sick sailor in (IU) will move them to (Q). A false negative test of a sick sailor in (IU) will keep them in (IU), thus allowing them to keep mixing with the rest of the healthy population.

Susceptible to Infected Unaware parameter process. The susceptible bucket is modeled as a pure Birth-Death process in which only the next sailor to get sick is modeled one at a time. Each sailor has their own parameters that determine if they re-join the (S) queue or if they become infected. This process sets the parameters for sailors on an individual basis that rely upon the rest of the model's states. We calculate the time until next infection using the exponential distribution in III.A.2 and then calculate the time until the next routine testing event for the model. If the time until the next infection is greater than the time until the next testing event, the sailor rejoins the (S) stage of the model. If the time until the next testing event is greater than the time until the next infection, then the susceptible sailor in question becomes infected and is moved to the (IU) stage of the model.

Infected Unaware Input process. Each time a sailor becomes infected, they do not know they are sick until either they get a positive test or symptoms show. This process is initialized each time a sailor enters the (IU) stage and assigns them a time until symptoms show, taking into consideration the probability that symptoms will show at all. If the sailor is tested before their symptoms show, and the test is a true positive, they will be moved to quarantine. If the sailor's symptoms show before they receive a true positive test from a routine test, they are tested immediately and will be moved into quarantine. A false negative test result in either an individual test or a routine test will keep the sailor in the

(IU) stage until a true positive is received. If a sailor never tests positive, then they will eventually move to the recovered stage without ever knowing they were ill.

5. Tracking Population and Subpopulations

The model tracks the number of healthy (susceptible and recovered), infected unaware, and quarantined sailors using the aforementioned status labels. At any given time in the simulation, we know the proportion of the total population that is sick and healthy.

We are most interested in tracking the status of specific subpopulations, since a ship is only operational when it meets a minimum threshold of healthy qualified sailors. For example, we track the subpopulation called "drivers" that represent the sailors able to drive the ship. We track the number of healthy, infected unaware, quarantined, and recovered drivers much in the same way we track the total population. By monitoring these numbers, we can create several measures of performance relating to the operational capability of the ship. Of course, ships have many subpopulations beyond drivers, but we choose to model only drivers versus non-drivers. Using similar techniques, this can be expanded to include every subpopulation of interest.

6. Model Scenarios

We wish this model to be as adaptable for use of the Navy. There are many different types of ships with different sizes and sub-populations to monitor, so we run this model on four different scenarios:

Table 2. Model Scenario Inputs

Type of Ship	Total Crew Size (approx.)	Number of "Drivers" (approx.)
Aircraft Carrier	6000	15
Destroyer	300	15
Cruiser	500	15

In each of these scenarios, we maintain the number of "drivers" as a constant 15. There may not be exactly 15 drivers on each of these ships, but it is a rough estimate of

how large a sub-population of interest might be on a ship, such as number of people in a division, in a particular berthing, et cetera. Keeping the number of "drivers" to track constant throughout each scenario enables us to compare how the size of a ship might affect the spread of COVID-19 and how important the inclusion of non-homogeneous mixing is to the model.

B. IMPLEMENTATION OF BASIC MODEL IN SIMIO

We implement the model in the computer program SIMIO. We create servers for each "stage" and named them susceptible_bucket, infected_unaware_bucket, quarantine_bucket, and recovered_bucket. Sailors can be in various stages of the illness upon initialization of the model, meaning at time "zero" of the simulation, we can have sailors in susceptible, quarantine, and infected unaware.

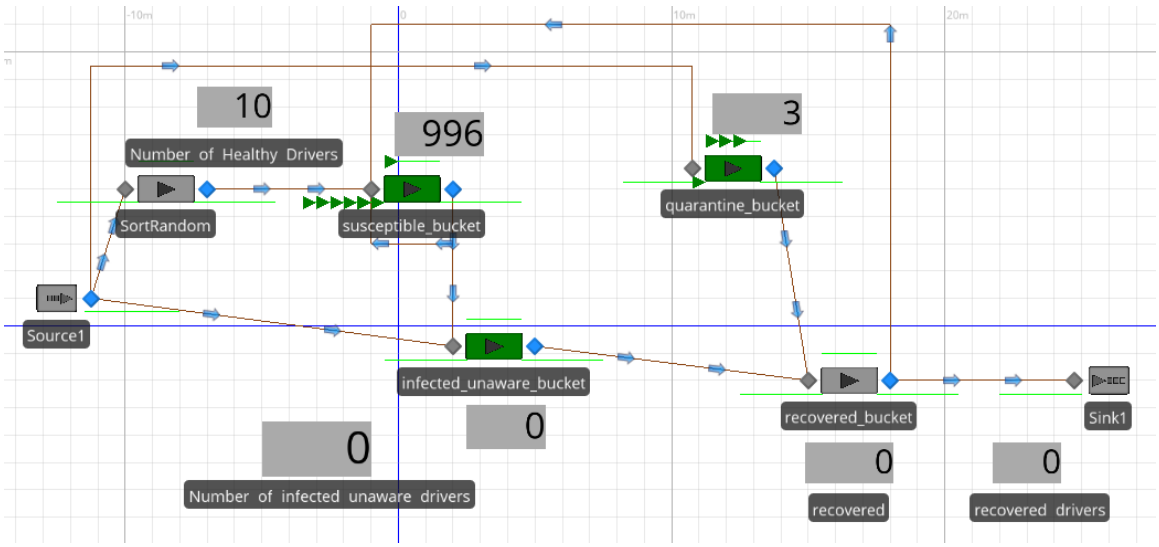


Figure 2. Screenshot of Basic SIMIO Model, Run With 1,000 Entities

Since susceptible is a pure birth-death process with a capacity of one, the entities must be processed in a randomized order (according to a random Uniform (0,1) priority assigned upon entry to the simulation). This randomized order is implemented in the server SortRandom. Status labels track the overall status of the system, where we can see the total

number of susceptible, infected unaware, or recovered sailors as well as the number of "drivers" in those groups as well.

C. NON-HOMOGENEOUS MIXING

As we discuss in Chapter III.A, a Navy ship population mixes differently than the general public. In this chapter we consider a non-homogeneous mixing setting, where the mixing rate depends on the job type. We pose the simplest non-trivial setup, where susceptible ship drivers interact with IU randomly with a mixing rate κ_d and the susceptible non-drivers mix with the IUs under a different mixing rate, κ_{nd}. To make the model realistic, we assume $\kappa_d \leq 20 \leq \kappa_{nd}$, corresponding to a scenario where the ship drivers, being a key subpopulation, is kept more isolated than the rest of the population.

D. TESTING PROTOCOLS

There are three different testing policies we decide to compare, which are built off the model described. Each testing policy is compared in Chapter IV.

1. Deterministic Testing Schedule

The first testing protocol is the one defined in the basic model description. We assume a deterministic testing schedule, where every sailor is be tested either individually or with a group testing protocol as defined in the Lin et al. paper presented at the WSC (Lin et al. 2020). These deterministic times between tests can range from every other day, to every week, every two weeks, or once a month. If an infected sailor presents symptoms between the routine testing events, they will be tested individually and sent to quarantine as necessary.

2. Subpopulation-Based Testing Schedule

The second testing protocol is a similar deterministic protocol, but each subpopulation is tested at different intervals based on importance. For example, since we are really interested in the subpopulation of "drivers" staying healthy to maintain operational readiness of the ship, then we test that subpopulation at shorter intervals than

the rest of the ship. This can be done for any subpopulation; in our model we will test the drivers and non-drivers at different intervals.

This testing policy can be implemented by expanding the deterministic testing model to include a second real state variable, time between tests for drivers. Then, we replace the original time between tests variable with a Boolean expression that evaluates either the driver or non-driver state value depending on a specific entity's job label. To show how this could be implemented, we set the driver's testing schedule to be once every three days and the non-drivers get tested once a week.

3. Dynamic Testing Schedule

The third testing protocol is dynamic. The testing period length varies based on prevalence of the virus on the ship. As more sailors test positive, the intervals between testing of the general population decreases. The theory is that more testing will catch more infected unaware individuals and therefore the overall infection rate will decrease as a response to closer testing periods. Once the overall infection rate decreases again on the ship, the testing schedule can be relaxed and have longer intervals between testing. Logically, this makes sense as on a long deployment, if the number of infectious sailors aboard is low (or zero, in an ideal situation) then the whole ship will not need to be tested every week as there will likely not be any more infections.

To implement the dynamic testing protocol in our model, we reassign the value of time between tests every time a sailor tests positive. As the number of quarantines sailors increases, the time between tests decreases, so we set the time between tests to be 500/(# of sailors in quarantine + 1) hours.

4. Close Contacts and Contact Tracing

A fourth testing protocol that we do not implement in this study is the testing of close contacts of a newly quarantined sailor. Every sailor that tests positive would trigger the test of every close contact over the previous several days. With our simulation, this testing protocol is not possible to implement, but is recommended to be researched in future studies.

E. FINAL MODELS

We end up with three models to run in SIMIO and compare the testing protocols. Each of these models has the same skeleton that accounts for the SEIR model dynamics described as the basic model in chapter III.A.1. Each accounts for non-homogenous mixing of the subpopulations using the mixing constant methods described in III.C. The models are run with several experiments to vary parameters and ship size scenarios. Chapter IV discusses both the experiments as well as the results.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. ANALYSIS AND RESULTS

We run each of the models created in Chapter III with the parameters described. We also run experiments on each model with a sensitivity analysis on the parameters as well as different scenarios for each ship size. The scenario is run for a time period of 2 months (60 days), and each instance is run with 50 replications. The metrics we analyze are given in IV.A.

In IV.B, only the mean values of observations across these replications is reported for each metric to two decimal places. In IV.C, we present histograms as well as the ranges (minimum and maximum values recorded) for the number of healthy drivers, the number of healthy sailors, and the number of infected unaware sailors.

Many of the results in Chapter IV look very similar to each other or even result in the same averages across observations between different testing protocols. We conclude that the reason this happens is due to the input parameters and assumptions of the model. Since we operate under the assumptions of social distancing, mask wearing, self-reporting symptoms with a probability of 1, and no outside interaction/ infection possibility, the model itself is very similar across all runs. With such low initial prevalence rates and probability of transmission given contact each time the model is run there are very few sailors who end up infected unaware. This means there are also very few new infections from interactions between infected unaware and susceptible sailors and since there is no outside infection possibility included in this model, the time until next infection is set to infinity once everyone on the ship is "healthy" or quarantined.

The metrics we describe in IV.A seem insensitive due to the low prevalence of infection on ships within the scenarios with nominal input parameters in IV.B. In contrast, in IV.C there is a significant difference between the averages across runs between the pessimistic and nominal scenarios. We conclude the metrics are not insensitive and the model run with nominal parameters results in such low numbers of infected unaware sailors that the metric outputs look the same. This conclusion is supported by the averages reported

for a scenario where no routine test is administered in IV.B, which are very similar to the averages of the testing scenarios.

A. METRICS OF ANALYSIS

To compare each testing alternative, we need a standard set of metrics. We develop these metrics off of what the Navy would deem important to operational ability.

Number of healthy drivers over time: this metric tracks the temporal distribution of the number of healthy drivers and delivers the probability that the ship is below some critical threshold of drivers needed to be fully operational. This critical threshold is developed by assuming a ship needs two watch standers (port and starboard) for each of the three watch rotations. Any less than 6 healthy drivers would render the ship inoperable. A dangerous level of healthy drivers is below 10; meaning only 5 of the 15 can be sick at any given time. We track the proportion of time that the ship is below both this danger level and the inoperability level.

Number of total sailors healthy over time: Similar to tracking the healthy drivers, we will track the total number of healthy sailors on the ship. A ship will be at dangerous levels when reduced to 75% of total starting strength and will be inoperable once below 50%. This metric tracks the number of susceptible sailors in the model and reflects potential false positives, meaning if a susceptible sailor receives a false positive and is sent to quarantine, the total number of healthy sailors will decrease, even though that sailor is technically "healthy."

Number of total sailors infected unaware over time: We will also track the total number of infected unaware over time. The goal of each testing protocol is to keep this number as close to zero as possible.

Time-Cost of Testing: The non-monetary cost of testing sailors. If we assume each test takes about 30 minutes to conduct, that is 30 minutes of the sailor's day that they are unable to be working, resting, et cetera. We report this "time cost" as a percentage of the working day (16 hours) that is spent on COVID testing. A test every two days would be calculated as

$$\frac{30}{2*16*60}*100 = 1.5625\%$$

We refer to this percentage as the time-cost of an individual for a particular testing protocol. We recognize this time-cost as the only cost metric we evaluate because we do not consider budgetary constraints when designing these testing protocols. As the time-cost represents time not working during the working day, we wish to select a protocol with a lower time-cost.

Threshold probabilities: We also report the ratio of time that the ship's drivers and total populations are below the danger and inoperable thresholds. This ratio can be interpreted as the probability that the ship will be at dangerous or inoperable levels.

B. EXPERIMENTS ON TIME BETWEEN TESTS

Each one of the three testing alternatives is compared using the metrics in IV.A. We are first interested in the "best" length of testing interval within each testing protocol for each ship size scenario. This section explores the impact of different testing schedules and see if varying the testing length will affect the model greatly. If the model is not affected significantly, then the longest testing period (that is not "never") is chosen since we wish to reduce operational and monetary strain of COVID testing.

Recall that under our model assumptions that any sailor who is infected unaware and presents symptoms will self-report and be tested individually, even under the "never" testing policy. An inter-testing period length of "never," or "no test" means that there is no routine testing policy for the whole ship and instead, only sailors who present symptoms and self-report will be tested. This testing protocol will have no false positive test results, as only true infected unaware sailors will be tested when they present symptoms. This is evident in the tables, where the "no test" policy will report a higher average number of healthy sailors since there are no false positives. We explore the tradeoff between no false positive results and catching the infected unaware sailors sooner through routine testing to find the "best" interval.

Once we find the testing interval for each ship size and each testing protocol, we compare the different protocols to each other as well as the best and worst case scenarios

in Chapter IV.C. This section of the chapter focuses solely on varying the time between tests.

We run a total of 45 sub-experiments using the three models. Each testing protocol has four experiments that correspond to ship size of Destroyer, Cruiser, and Carrier. Each of these experiments changes the number of sailors in the system and has multiple sub-experiments to vary input parameters. The scenarios are run with the "nominal" parameters where only the time between tests changes within each ship size. The results are reported in table format, with one table for each of the 9 experiments. Only averages are reported in these tables, to two decimal places, and further detail on the metrics such as value ranges and histograms are presented in Chapter IV.C.

1. Deterministic Testing

The deterministic testing model has five experiments per ship scenario with a time between tests of 3 days, 1 week, 3 weeks, and 4 weeks. Each is reported in contrast to never doing a routine test as well. A total of 15 experiments of deterministic testing are run.

Table 3. Destroyer: Deterministic Testing

Inter-Testing Period	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of IU Sailors	Time-Cost of Test
3 days	14.69	293.33	0.77	1.04%
1 week	14.77	293.68	1.47	0.44%
3 weeks	14.78	293.92	1.97	0.14%
4 weeks	14.76	293.90	2.30	0.11%
No test	14.76	293.86	2.46	0%

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over the others, meaning that the longest inter-testing period of 4 weeks is chosen for analysis in chapter IV.C.1.

Table 4. Cruiser: Deterministic Testing

Inter-Testing Period	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of IU Sailors	Time-Cost of Test
3 days	14.65	487.78	1.07	1.04%
1 week	14.75	489.60	2.16	0.44%
3 weeks	14.79	490.15	3.04	0.14%
4 weeks	14.78	490.42	3.12	0.11%
No test	14.78	490.55	3.30	0 %

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over the others, meaning that the longest inter-testing period of 4 weeks is chosen for analysis in Chapter IV.C.2.

Table 5. Carrier: Deterministic Testing

Inter-Testing Period	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of IU Sailors	Time-Cost of Test
3 days	14.67	5826.87	5.45	1.04%
1 week	14.75	5860.21	10.92	0.44%
3 weeks	14.78	5876.23	30.80	0.14%
4 weeks	14.78	5878.97	30.65	0.11%
No test	14.78	5881.49	24.18	0%

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over the others, meaning that the longest inter-testing period of 4 weeks is chosen for analysis in Chapter IV.C.3.

2. Subpopulation-Based Testing

The subpopulation-based testing model has five experiments per ship scenario varying the time between tests for non-drivers and driver subpopulations. Listed in pairs of (driver interval, non-driver interval), the experiments are (2 days, 3 days), (3 days, 1 week),

(1 week, 3 weeks), (3 weeks, 4 weeks), and (4 weeks, never test). There are a total of 15 experiments run on the subpopulation-based testing model. The time-cost of each experiment is calculated with respect to the proportion of drivers and their time lost added to the proportion of non-driver's time loss.

Table 6. Destroyer: Subpopulation-Based Testing

Inter-Testing Period	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of IU Sailors	Time-Cost of Test
(2d, 3d)	14.71	293.36	0.73	1.06%
(3d, 1w)	14.73	293.65	1.48	0.47%
(1w, 3w)	14.78	293.94	1.98	0.16%
(3w, 4w)	14.76	293.94	2.31	0.11%
(4w, never)	14.76	293.81	2.45	0.01%

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over the others, meaning that the longest inter-testing period of (4 weeks, never) is chosen for analysis in Chapter IV.C.1.

Table 7. Cruiser: Subpopulation-Based Testing

Inter-Testing Period	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of IU Sailors	Time-Cost of Test
(2d, 3d)	14.65	487.78	1.07	1.05%
(3d, 1w)	14.71	489.73	2.25	0.46%
(1w, 3w)	14.74	490.15	3.01	0.16%
(3w, 4w)	14.75	490.30	3.08	0.11%
(4w, never)	14.78	490.52	3.31	0.00%

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over

the others, meaning that the longest inter-testing period of (4 weeks, never) is chosen for analysis in Chapter IV.C.2.

Table 8. Carrier: Subpopulation-Based Testing

Inter-Testing Period	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of IU Sailors	Time-Cost of Test
(2d, 3d)	14.61	5839.67	5.44	1.04%
(3d, 1w)	14.68	5870.94	10.91	0.44%
(1w, 3w)	14.76	5879.56	30.46	0.15%
(3w, 4w)	14.80	5884.10	29.82	0.11%
(4w, never)	14.80	5890.78	24.18	0.00%

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over the others, meaning that the longest inter-testing period of (4 weeks, never) is chosen for analysis in Chapter IV.C.3.

3. Dynamic Testing

The dynamic testing model has five experiments per ship scenario and varies the numerator of the time between tests calculation. Recall the time between tests in the dynamic model is $\frac{500}{Q+1}$ hours, where Q is the number of sailors in quarantine. We vary the numerator to be 100, 200, 500, 1000, and 2000. A total of 15 experiments are run for the dynamic testing model. Since the testing policy is dynamic, the time-cost of this protocol is not reported.

Table 9. Destroyer: Dynamic Testing

Inter-Testing Period	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of IU Sailors
$\frac{100}{Q+1}$	14.68	292.38	0.74
$\frac{200}{Q+1}$	14.74	293.47	0.90
$\frac{500}{Q+1}$	14.75	293.60	1.12
$\frac{1000}{Q+1}$	14.75	293.57	1.11
$\frac{2000}{Q+1}$	14.75	293.58	1.12

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over the others, meaning that the longest inter-testing period of $\frac{2000}{Q+1}$ hours is chosen for analysis in Chapter IV.C.1.

Table 10. Cruiser: Dynamic Testing

Inter-Testing Period	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of IU Sailors
$\frac{100}{Q+1}$	14.37	479.96	0.78
$\frac{200}{Q+1}$	14.62	487.95	1.19
$\frac{500}{Q+1}$	14.74	489.44	1.72
$\frac{1000}{Q+1}$	14.75	489.62	1.80
$\frac{2000}{Q+1}$	14.75	489.65	1.82

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over the others, meaning that the longest inter-testing period of $\frac{2000}{Q+1}$ hours is chosen for analysis in Chapter IV.C.2.

Table 11. Carrier: Dynamic Testing

Inter-Testing Period	**Avg # of Healthy Drivers**	**Avg # of Healthy Sailors**	**Avg # of IU Sailors**
$\frac{100}{Q+1}$	7.24	2926.95	11.22
$\frac{200}{Q+1}$	7.59	3048.23	13.04
$\frac{500}{Q+1}$	14.71	5843.79	26.63
$\frac{1000}{Q+1}$	14.78	5877.39	22.34
$\frac{2000}{Q+1}$	14.78	5879.84	21.83

The ratios for driver danger levels, driver inoperability levels, total danger levels, and inoperable total levels are all zero. There is no advantage of one testing period over the others, meaning that the longest inter-testing period of $\frac{2000}{Q+1}$ hours is chosen for analysis in Chapter IV.C.3.

4. Addressing the Non-monotonicity

In each of Table 3 through Table 11 above, there is an inconsistency with the average number of susceptible sailors reported by the model. As fewer tests are conducted, the average number of healthy sailors actually increases. This does not make sense from an epidemiological base, so we investigate further. The difference is exceptionally large in the Carrier scenario, where there are 6,000 sailors on board.

We attribute the differences to false positives. As more frequent tests are conducted, there are more false positives. Even if there are no infected unaware sailors on board, a testing event will falsely label a healthy sailor as sick. With no tests, a sailor will only be tested once they show symptoms. This allows for more interaction between susceptible and infected unaware sailors, resulting in more infections, but there are no false positives so the number of average susceptible sailors reported by the model is higher.

This issue presents itself in all three testing protocols, but it is most evident for the dynamic testing procedure. Recall the dynamic testing protocol depends on the number of sailors in quarantine. When a false positive occurs, the time until next test is shortened. This next test may result in more false positives which in turn will shorten the time between tests again. The cycle continues until the number of quarantined sailors is so high from false positives that the calculated time between tests is operationally infeasible. The most evident example of this happening in the model is in Table 11, where the average number of healthy sailors ranges from 2926.95 when tests are conducted every $\frac{100}{Q+1}$ hours to 5879.84, when tests are conducted every $\frac{2000}{Q+1}$ hours.

To further explore this inconsistency, we explore the relationship between the false positives and this non-monotonicity. We increase the probability of false positives to three times the original, 0.003, and also decrease it to zero. The results of these runs are shown for all three testing protocols only for the Carrier scenario.

Table 12. Deterministic Testing: Average Number of Healthy Sailors

Inter-Testing Period	**0.001 (Standard)**	**0.00 (Optimistic)**	**0.003 (Pessimistic)**
1 week	5860.21	5885.58	5815.99
3 weeks	5876.23	5884.54	5852.96
4 weeks	5878.97	5884.62	5871.08
No test	5881.49	5884.56	5881.63

Table 13. Subpopulation-Based Testing: Average Number of Healthy Sailors

Inter-Testing Period	0.001 (Standard)	0.00 (Optimistic)	0.003 (Pessimistic)
(3d, 1w)	5870.94	5881.75	5815.75
(1w, 3w)	5879.56	5881.54	5852.17
(3w, 4w)	5884.10	5881.70	5871.07
(4w, never)	5890.78	5881.65	5881.61

Table 14. Dynamic Testing: Average Number of Healthy Sailors

Inter-Testing Period	0.001 (Standard)	0.00 (Optimistic)	0.003 (Pessimistic)
$\frac{500}{Q+1}$	5843.79	5881.61	5794.03
$\frac{1000}{Q+1}$	5877.39	5881.56	5818.46
$\frac{2000}{Q+1}$	5879.84	5881.52	5858.78

It is clear that the false positives has a large effect on the non-monotonicity of the model. With a false positive rate of zero, the average number of healthy sailors decreases or remains the same as the tests become less frequent. With a higher rate of false positives, the number of healthy sailors increases as the tests become less frequent. Therefore, we can attribute the non-monotonicity of the model to the false positive rate of 0.001.

This modeling issue can be remedied by adding a second, more sensitive, test to anyone who has tested positive with a PCR test. We do not implement this improvements in our model and leave it to be explored in future work.

5. Takeaways

The major takeaways from the analysis of the time between tests over the 9 different scenarios is that under nominal parameters and the assumptions listed in III.A.3, the time between tests does not have a large effect on the operational readiness of the ship under

our defined metrics. The thresholds for dangerous levels of infection are never met, even if ships do not employ a regular testing event.

Specifically looking at the metric of the average number of infected unaware sailors; it makes sense that the number of infected unaware sailors increases as tests become less frequent. Conversely, high-frequency testing minimizes the number of infected unaware sailors. This pattern is protocol-independent.

The metrics for each scenario do not vary significantly across inter-testing time periods. We can conclude that the time between tests is inconsequential. The operational rationale for these results is that when the initial number of infected unaware sailors is low, barely testing the susceptible population and testing those with symptoms leads to a relatively larger number of infected people and a relatively small number of false positives. Conversely, when the testing is done aggressively, we end up with a relatively small number of infected people and a large number of false positives. In either scenario, the total number of sailors with a positive test result is similar. Since both true and false positive sailors are quarantined, the operational impact on ship operations is almost identical. The same conclusion can be drawn for intermediate testing protocols. Hence, an important take-away from our simulation model is that time between testing doesn't matter to the extent that: (i) The initial number of infected unaware sailors is low; and (ii), that people are tested soon after showing symptoms.

We create a fourth testing protocol in the next Section, called reactive testing. Under reactive testing, there is no scheduled ship-wide testing event and tests are conducted only on those who self-report due to symptoms. For the sake of our analysis, we choose the longest length testing period as the inter-test length for the three original testing protocols and compare each to each other and the reactive testing policy in the following section of this chapter.

C. ANALYSIS OF TESTING POLICIES

This chapter compares the four testing policies with the time between tests determined in IV.B. The testing policies are compared directly to each other and broken

down into different ship scenarios. Each testing policy is applied to nominal parameters, optimistic parameters, and pessimistic parameters for a fully robust analysis.

Within each ship scenario we present histograms and the ranges (minimum and maximum) for the number of healthy drivers, the number of healthy sailors, and the number of infected unaware sailors. Since the input parameters are now different, we will see differences in the metric averages across the observations that we did not see in IV.B.

1. Destroyer

The Destroyer scenario is run with a deterministic test every 4 weeks, a subpopulation based protocol that only tests drivers every 4 weeks (and does not test the general population), a dynamic test every $\frac{2000}{Q+1}$ hours, and a reactive test.

a. Average Number of Healthy Drivers

Shown in Figure 3, the average healthy sailors has a mean of 11.13 for the pessimistic case, 14.76 for the nominal case, and 14.76 for the optimistic case. The pessimistic case saw as low as 8 average healthy drivers and has a driver danger threshold ratio of 0.098. All other thresholds remained zero. The optimistic and nominal cases both recorded between 12 and 15 healthy drivers.

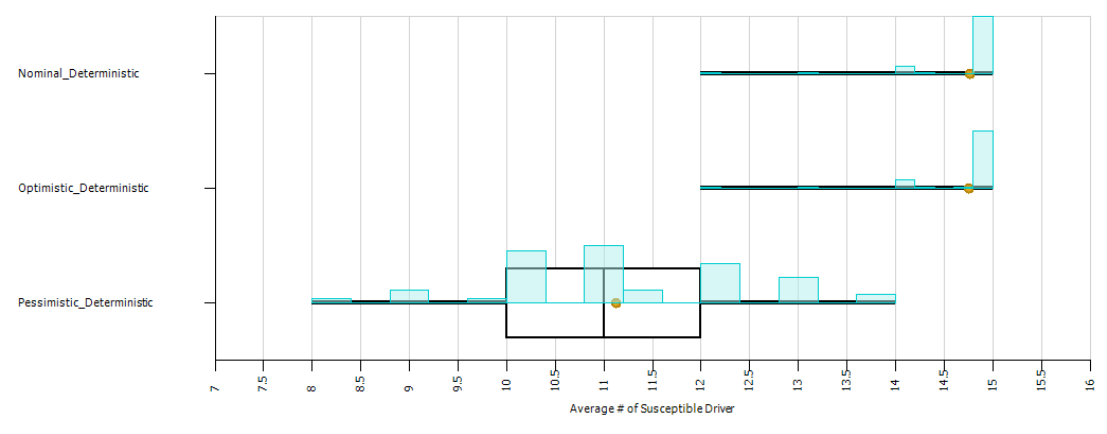


Figure 3. Destroyer Healthy Drivers, Deterministic Testing

When a subpopulation-based testing protocol is used, the mean value of all observations on healthy drivers is 11.14 for the pessimistic case, 14.75 for the nominal case, and 14.76 for the optimistic case. The nominal and optimistic averages both range between 12 and 15 and the pessimistic averages range between 8 and 14. The driver danger threshold ratio is 0 for all cases.

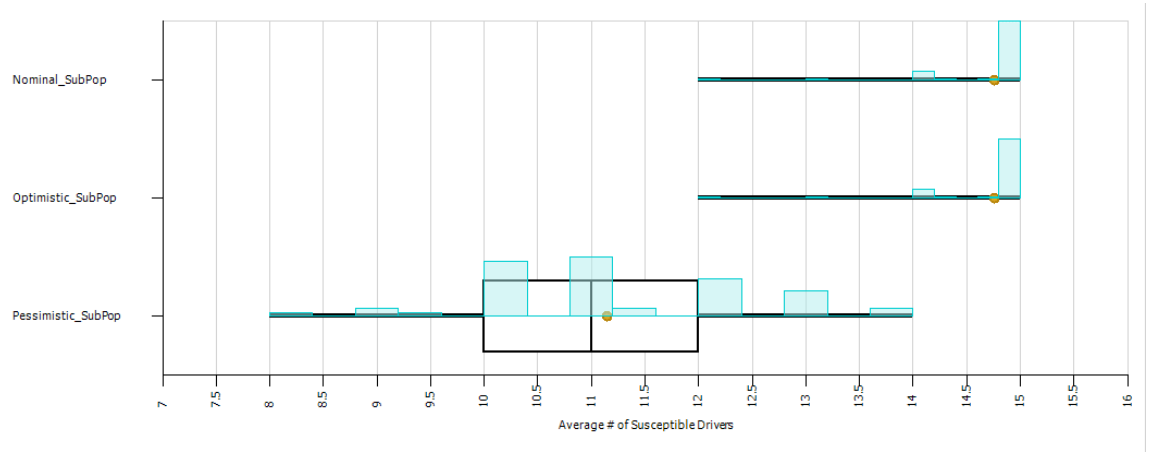


Figure 4. Destroyer Healthy Drivers, Subpopulation-Based Testing

The dynamic testing protocol results are shown in Figure 5. The average number of healthy drivers is 11.04 for the pessimistic case, with values ranging between 8 and 14. The optimistic case has a mean of 14.66 and 14.75 for the nominal case. Both have values between 12 and 15. The driver danger threshold ratio is 0.267 for the pessimistic case and zero for all other cases. Recall the threshold is so high due to the non-monotonicity discussed in IV.B.4.

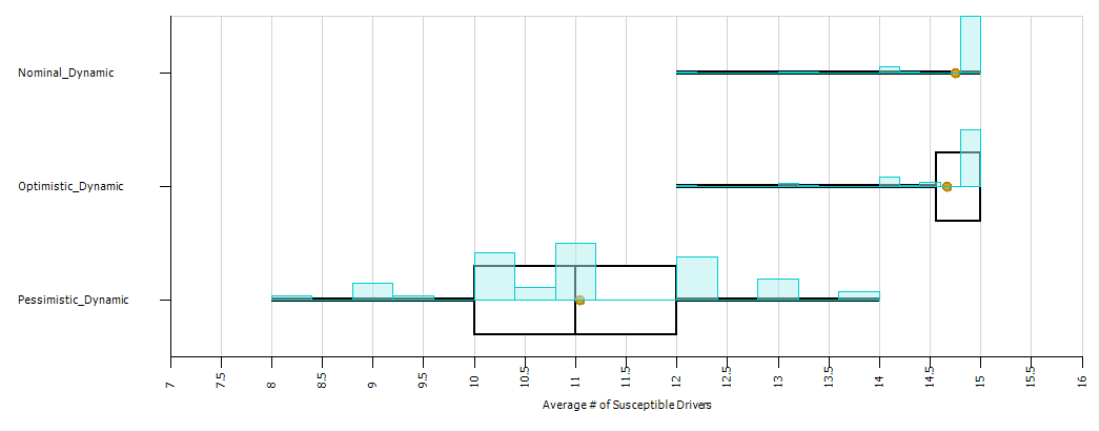


Figure 5. Destroyer Healthy Drivers, Dynamic Testing

The reactive testing protocol in shown in Figure 6 and returns an average number of healthy drivers of 14.76 in both the nominal and optimistic cases, with values ranging between 12 and 15. The pessimistic case returns an average of 11.14 with values between 8 and 14. All the danger and inoperability thresholds are zero for all three input parameter cases.

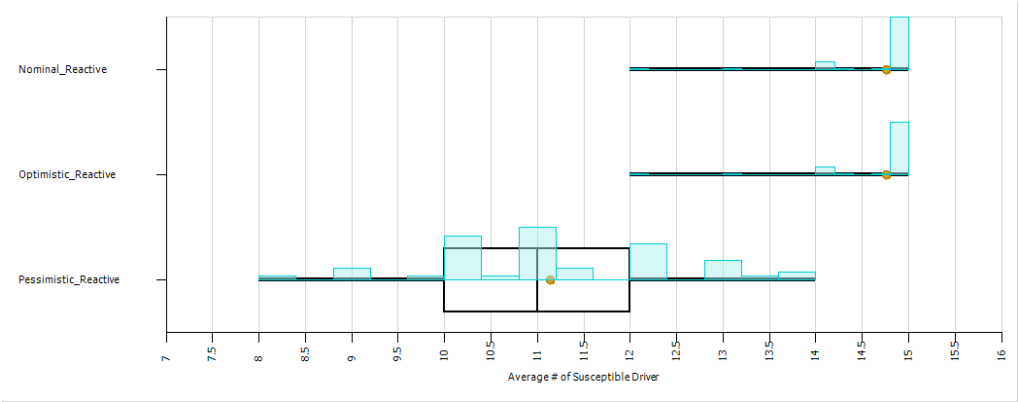


Figure 6. Destroyer Healthy Drivers, Reactive Testing

b. Average Number of Healthy Sailors

With a deterministic testing schedule, the average healthy sailors has a mean of 222.57 for the pessimistic case, 293.90 for the nominal case, and 293.84 for the optimistic case. The pessimistic observations ranged from as low as 199.84 healthy sailors to as high

as 236 sailors. The nominal and optimistic cases both ranged between 286 and 300 sailors. These results are presented in Figure 7.

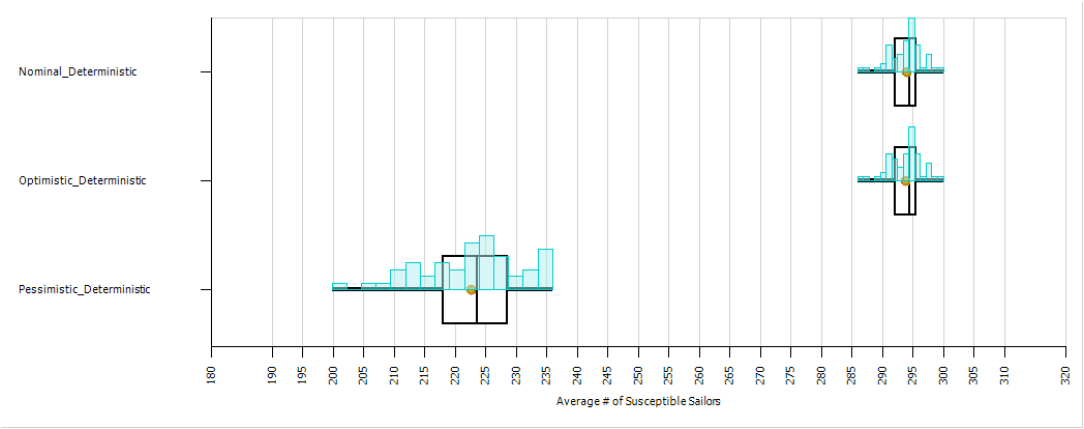


Figure 7. Destroyer Healthy Sailors, Deterministic Testing

When a Destroyer uses a subpopulation-based testing, the average across all observations is 223.16 for the pessimistic case with values ranging between 201.60 and 236.00. The optimistic case has an average of 293.78 and the nominal case has an average of 293.81. Both range between 286 and 300.

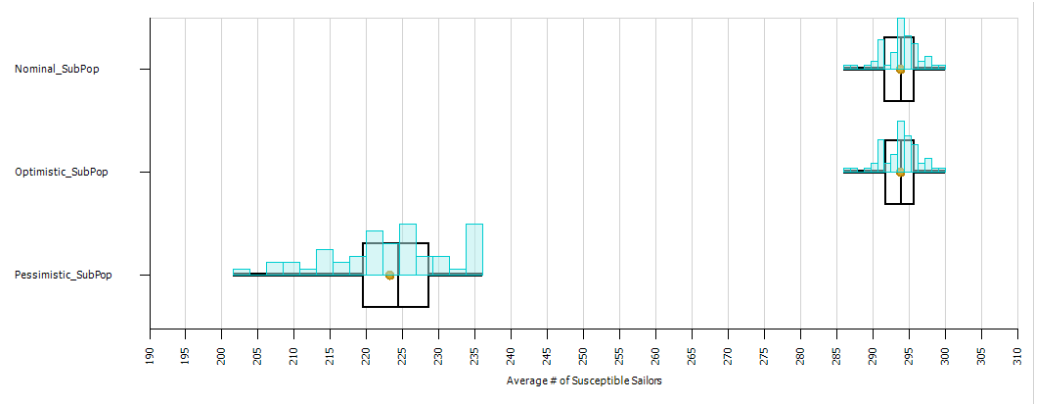


Figure 8. Destroyer Healthy Sailors, Subpopulation-Based Testing

Presented in Figure 9, the average number of healthy sailors using dynamic testing is 221.74 in the pessimistic case, 292.52 in the optimistic case, and 293.58 in the nominal

case. The pessimistic case has values ranging between 199.61 and 235.60, the optimistic case ranges between 285.51 and 297.66 healthy sailors, and the nominal case ranges between 286 and 298.

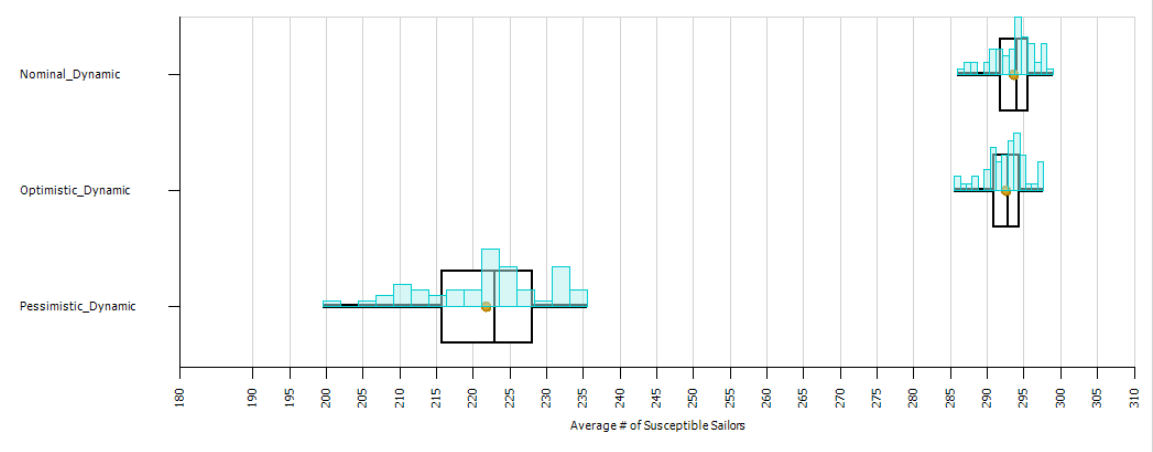


Figure 9. Destroyer Healthy Sailors, Dynamic Testing

The reactive testing protocol returns an average number of healthy sailors of 293.86 for the nominal case, 293.86 for the optimistic case, and 223.15 for the pessimistic case. The optimistic and nominal case averages both range between 286 and 300 and the pessimistic case averages range between 201.42 and 236.

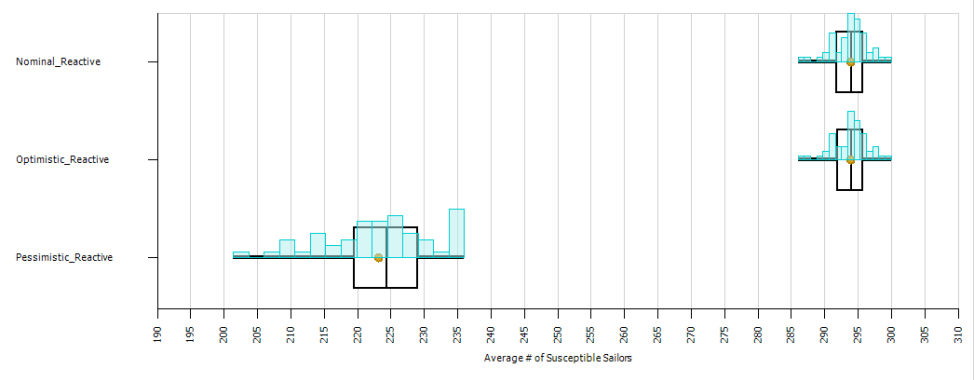


Figure 10. Destroyer Healthy Sailors, Reactive Testing

c. Average Number of Infected Unaware Sailors

The average number of infected unaware sailors in the destroyer scenario using deterministic testing is 20.75 for the pessimistic case, 2.21 for the optimistic case, and 2.30 for the nominal case. The pessimistic values range between 14.47 and 28.86, the optimistic values and the nominal values both range between 0 and 6.59 sailors. These results are in Figure 11.

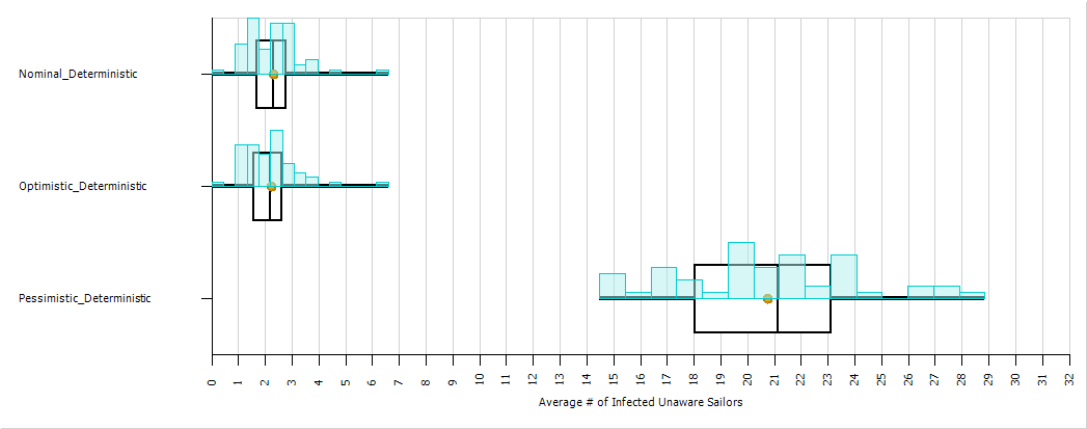


Figure 11. Destroyer IU Sailors, Deterministic Testing

Using a subpopulation-based technique, the average number of infected unaware sailors was 17.86, 2.45, and 2.45 for the pessimistic, optimistic, and nominal cases, respectively. The pessimistic case recoded values between 13.78 and 24.06 and the nominal and optimistic cases both have values between 0 and 6.59 sailors, as shown in Figure 12.

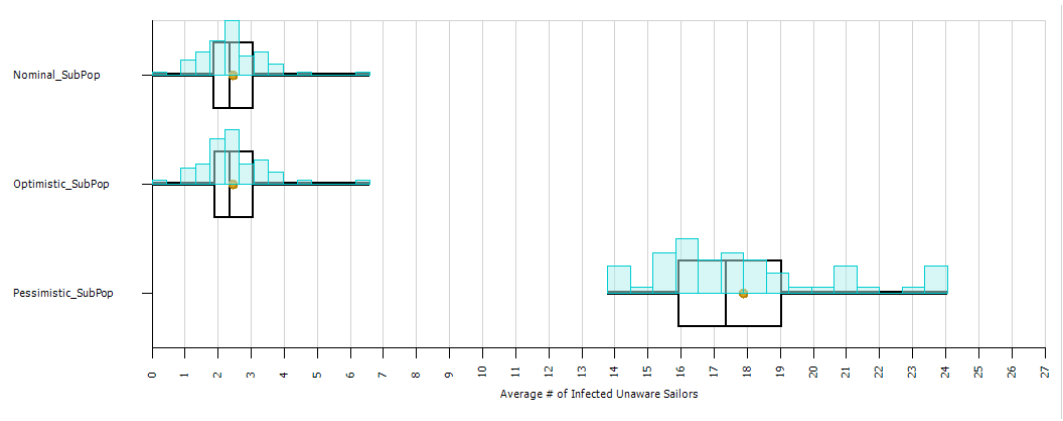


Figure 12. Destroyer IU Sailors, Subpopulation-Based Testing

The average number of infected unaware sailors using a dynamic testing protocol is 12.54 for the pessimistic case, 1.02 for the optimistic case, and 1.12 for the nominal case. The pessimistic case reports values between 7.43 and 18.30, the optimistic case reports between 0 and 3.17, and the nominal case reports between 0 and 2.914 infected unaware sailors.

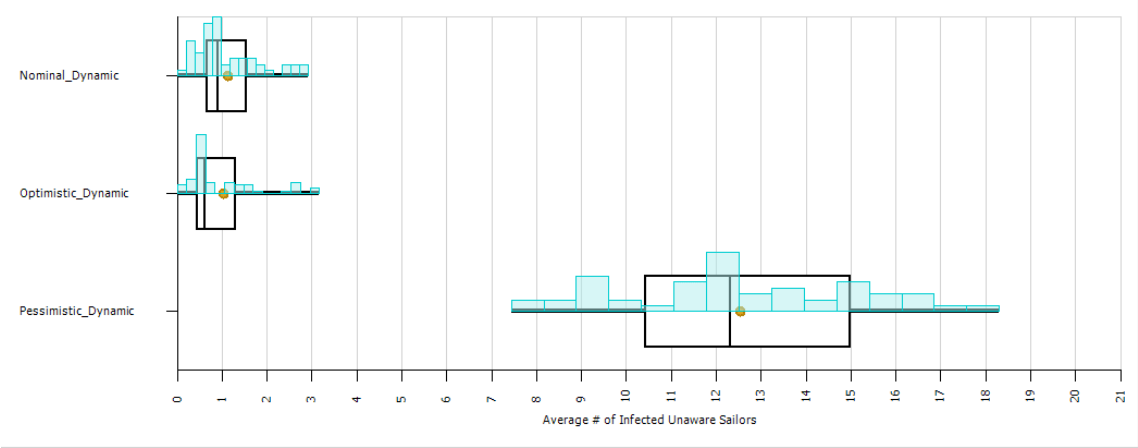


Figure 13. Destroyer IU Sailors, Dynamic Testing

Figure 14 details the results on the reactive test, which returns an average number of infected unaware sailors of 2.46, ranging between zero and 6.59, for both the nominal and optimistic cases. The pessimistic case average is 17.83 and ranges between 13.78 and 24.06.

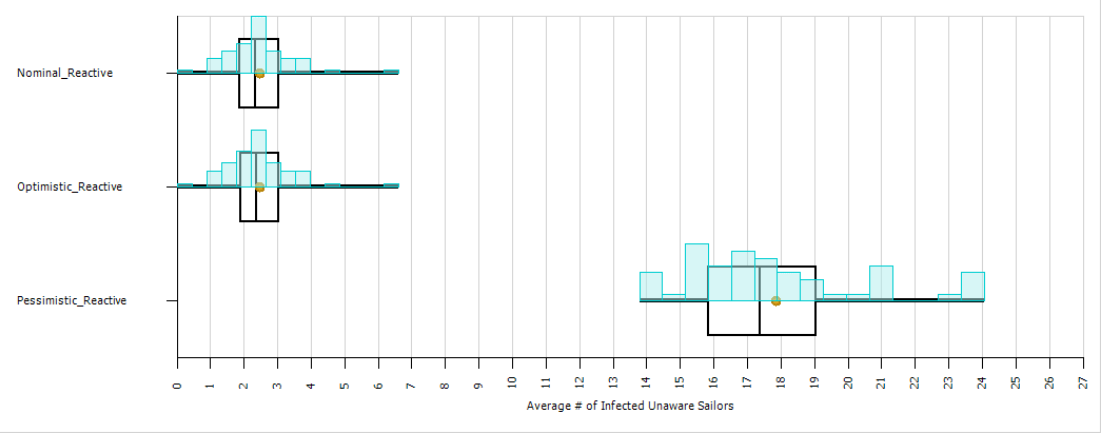


Figure 14. Destroyer IU Sailors, Reactive Testing

2. Cruiser

The Cruiser scenario is run with a deterministic test every 4 weeks, a subpopulation based protocol that only tests drivers every 4 weeks (and does not test the general population), a dynamic test every $\frac{2000}{Q+1}$ hours, and a reactive test.

a. Average Number of Healthy Drivers

The average across all observations on a Cruiser using deterministic testing is 11.08 for the pessimistic case and 14.78 for both the nominal and optimistic case. The pessimistic average healthy drivers range between 8 and 14 whereas the nominal and optimistic cases range between 12 and 15. The pessimistic case results in a driver danger threshold ratio of 0.118 and zero for all other ratios. These results are presented in Figure 15.

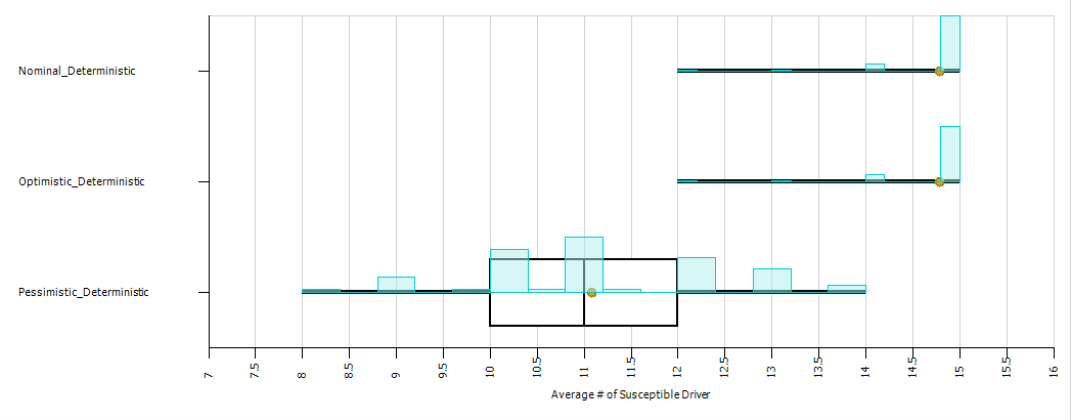


Figure 15. Cruiser Healthy Drivers, Deterministic Testing

The average across all observations on a Cruiser using subpopulation-based testing is 11.08 for the pessimistic case and 14.78 for both the nominal and optimistic case. The pessimistic average healthy drivers range between 8 and 14 whereas the nominal and optimistic cases range between 12 and 15. All of the danger ratios and inoperable ratios are zero using this testing method.

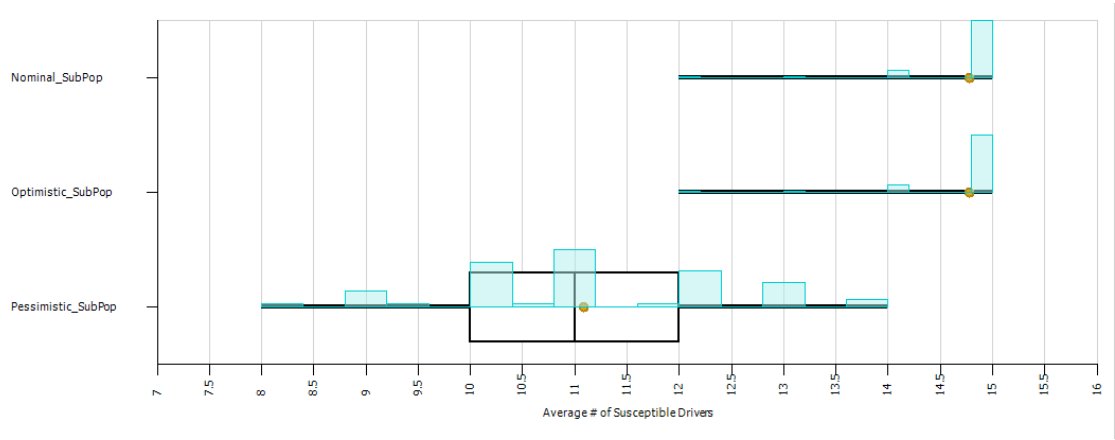


Figure 16. Cruiser Healthy Drivers, Subpopulation-Based Testing

The average across all observations on a Cruiser using dynamic testing is 11.11, 14.73, and 14.76 for the pessimistic, optimistic, and nominal cases. Given in Figure 17, the pessimistic average healthy drivers range between 8 and 14. The optimistic averages range

between 12 and 15 and the nominal cases range between 12.44 and 15. The pessimistic case results in a driver danger threshold ratio of 0.158 and zero for all other ratios.

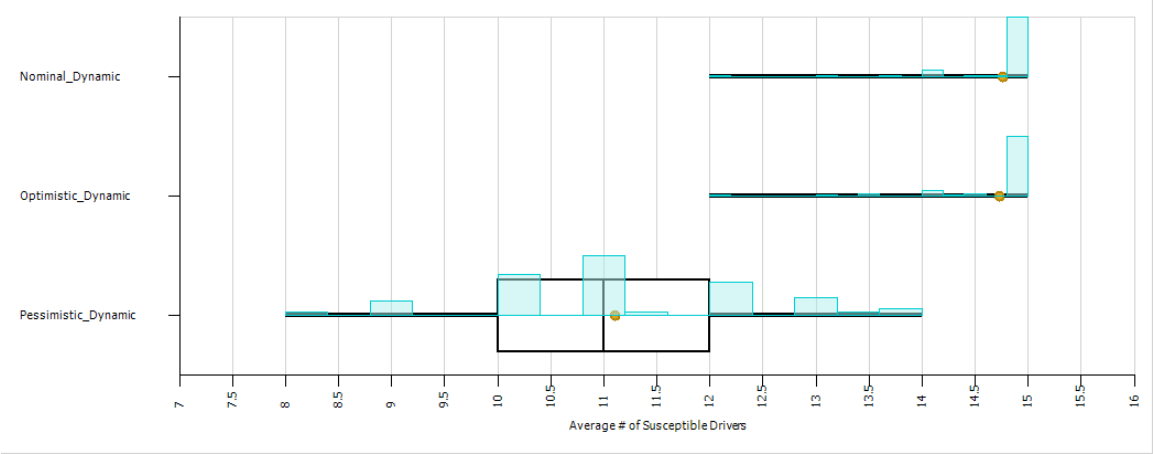


Figure 17. Cruiser Healthy Drivers, Dynamic Testing

The reactive testing protocol returns an average number of healthy drivers of 14.78, ranging between 12 and 15, for both the nominal and optimistic cases. The pessimistic case returns an average of 11.14 and has values between 8 and 14. All danger and inoperability threshold ratios are zero.

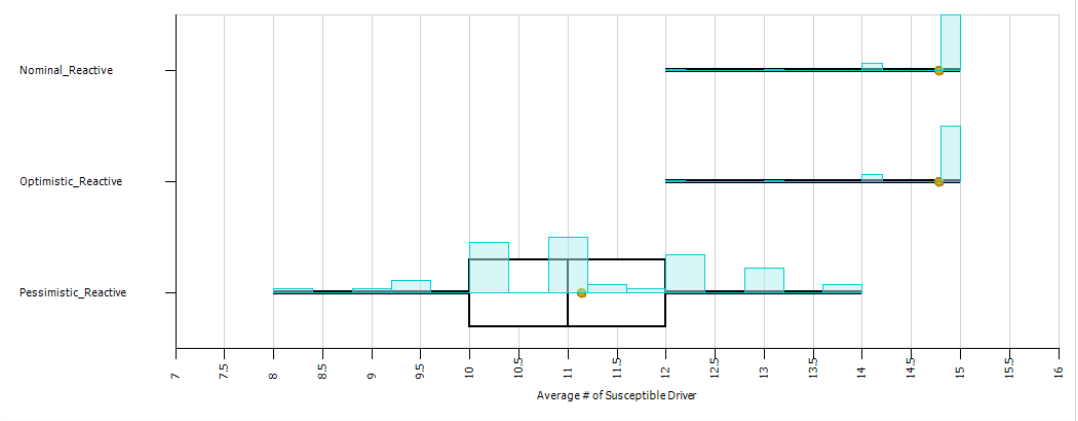


Figure 18. Cruiser Healthy Drivers, Reactive Testing

b. Average Number of Healthy Sailors

The average across all observations on a Cruiser using deterministic testing is 370.31 with values ranging between 346.14 and 390.24 for the pessimistic case. The optimistic case results in an average of 490.41 across all observations ranging between 484.36 and 498.45. Finally, the nominal case has an average of 490.42 healthy sailors and ranges between the values of 484.36 and 498.45.

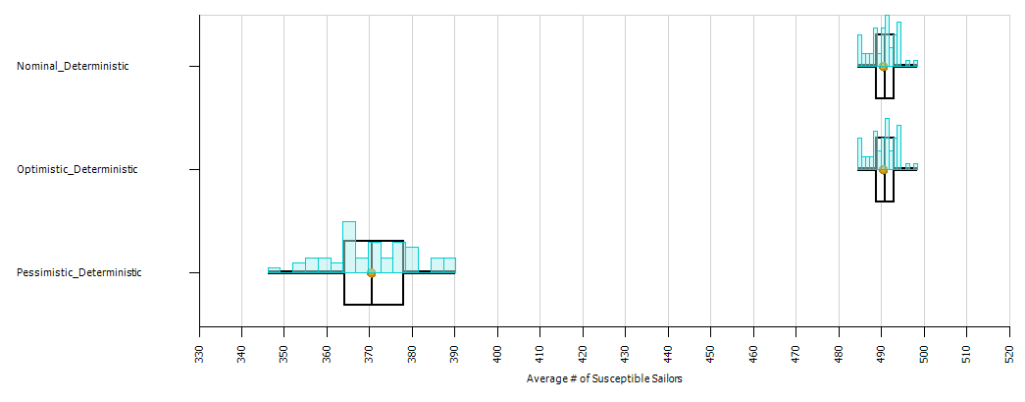


Figure 19. Cruiser Healthy Sailors, Deterministic Testing

The average across all observations on a Cruiser using subpopulation-based testing is 371.12 for the pessimistic case and 490.53 for both the nominal and optimistic case. The pessimistic average healthy sailors range between 347.81 and 391.36 whereas the nominal and optimistic cases range between 484.57 and 498.45. These averages and ranges are presented in Figure 20.

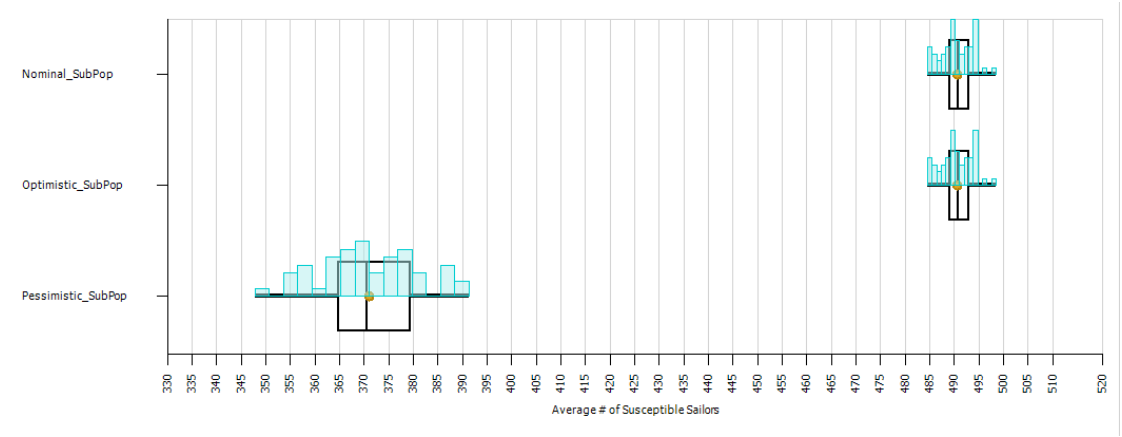


Figure 20. Cruiser Healthy Sailors, Subpopulation-Based Testing

Shown in Figure 21, the average across all observations on a Cruiser using dynamic testing is 369.37, 489.35, and 489.65 for pessimistic, optimistic, and nominal inputs, respectively. The values range between 345.82 and 387.96 for the pessimistic case. The optimistic case ranges between 481.93 and 495.25 and the nominal case ranges between 482.08 and 496.00.

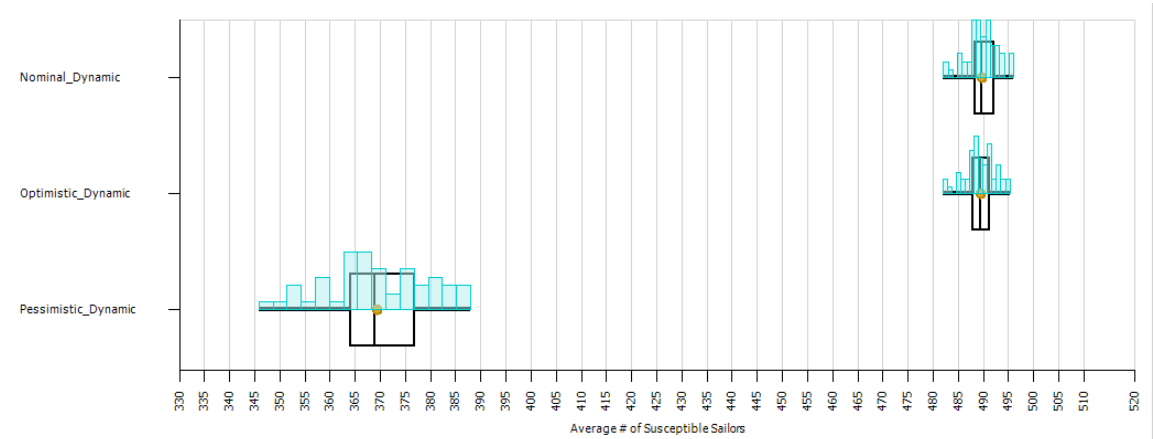


Figure 21. Cruiser Healthy Sailors, Dynamic Testing

The reactive testing protocol returns an average number of healthy sailors of 490.55 for the nominal case, 490.54 for the optimistic case, and 371.14 for the pessimistic case.

The nominal and optimistic cases both return values between 484.94 and 498.45 and the pessimistic case returns values between 347.72 and 391.22.

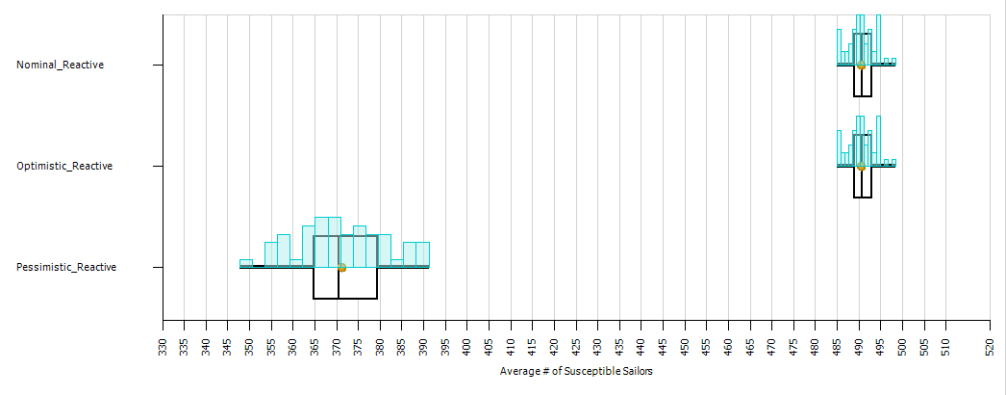


Figure 22. Cruiser Healthy Sailors, Reactive Testing

c. Average Number of Infected Unaware Sailors

The average number of infected unaware sailors across all observations on a Cruiser with deterministic testing is 33.91, 2.96, and 3.12 for the pessimistic, optimistic, and nominal cases, respectively. The pessimistic averages range between 23.49 and 45.05, the optimistic averages range between 1.14 and 6.07, and the nominal averages range between 1.14 and 6.71. These results are given in Figure 23.

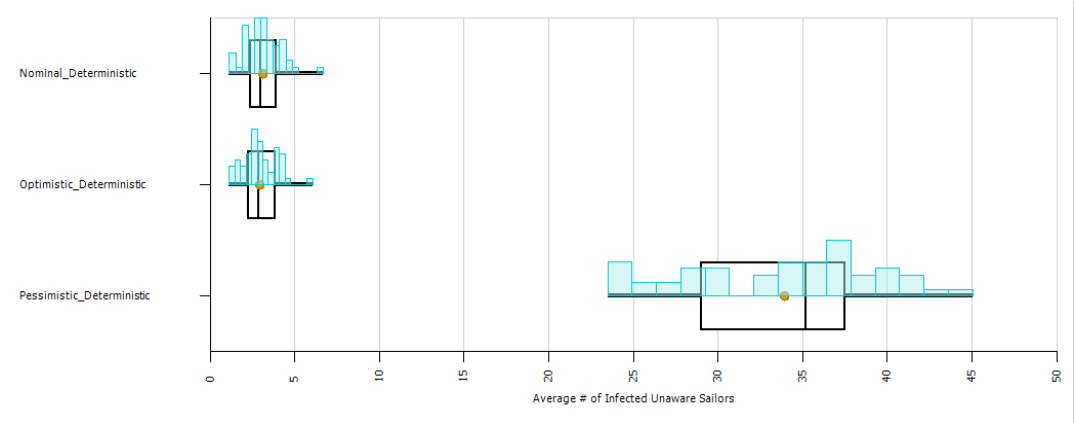


Figure 23. Cruiser IU Sailors, Deterministic Testing

The average number of infected unaware sailors across all observations on a Cruiser with subpopulation-based testing is 27.47, 3.29, and 3.31 for the pessimistic, optimistic, and nominal cases, respectively. The pessimistic averages range between 19.72 and 35.64, the optimistic and nominal averages both range between 1.14 and 7.27.

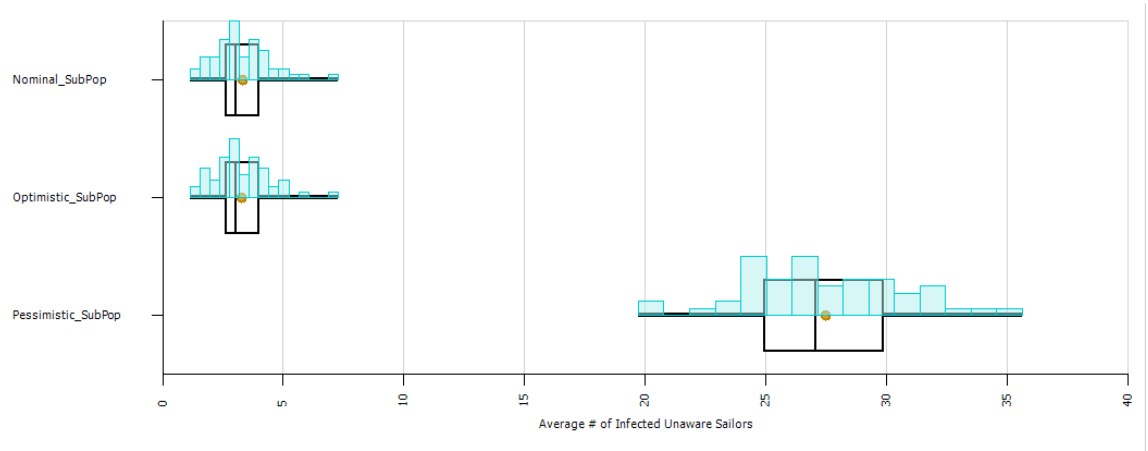


Figure 24. Cruiser IU Sailors, Subpopulation-Based Testing

Figure 25 shows the average number of infected unaware sailors across all observations on a Cruiser with dynamic testing is 18.62, 1.61, and 1.82 for the pessimistic, optimistic, and nominal cases, respectively. The pessimistic averages range between 8.38 and 25.55, the optimistic averages range between 0.09 and 5.92, and the nominal averages range between 0.14 and 5.92.

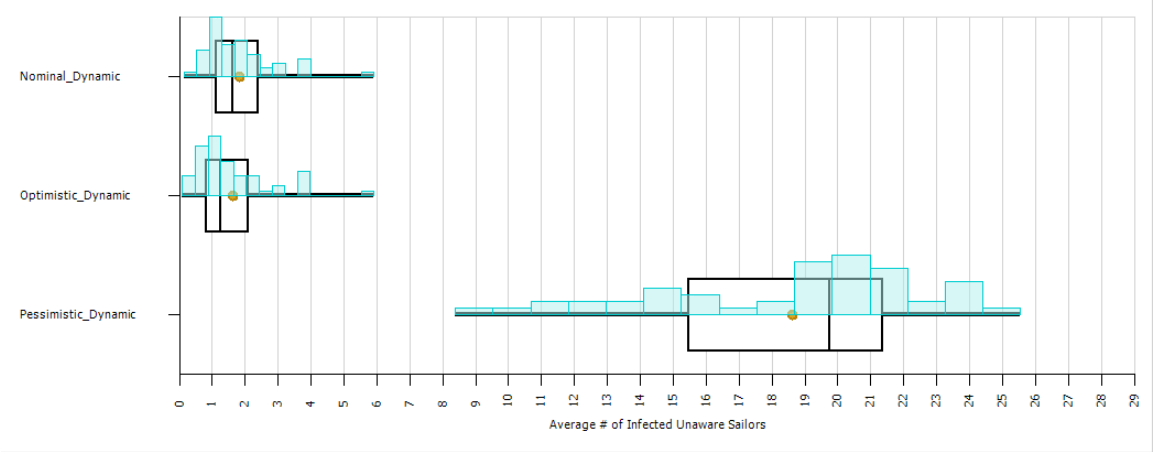


Figure 25. Cruiser IU Sailors, Dynamic Testing

The reactive testing protocol returns an average number of infected unaware sailors of 3.30 for the nominal case, 3.29 for the optimistic case, and 27.62 for the pessimistic case. The values of the averages range between 1.14 and 7.27 for both the nominal and optimistic cases and the pessimistic case ranges between 19.72 and 35.64.

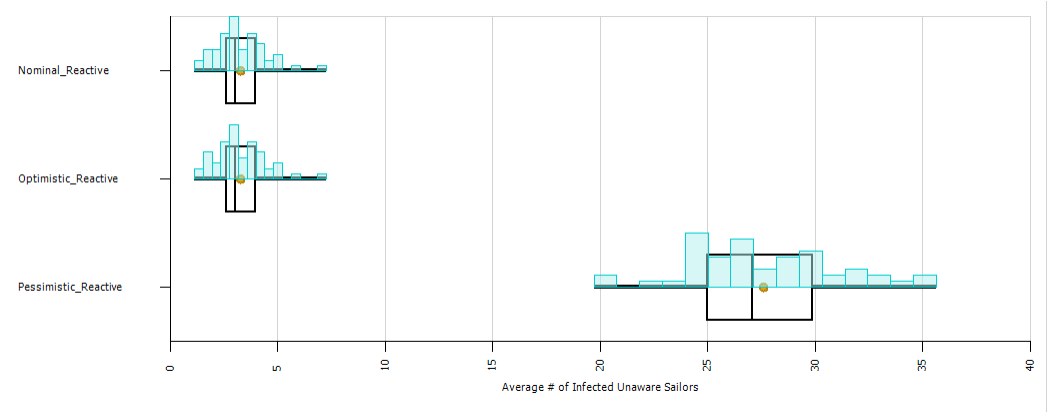


Figure 26. Cruiser IU Sailors, Reactive Testing

3. Carrier

The Carrier scenario is run with a deterministic test every 4 weeks, a subpopulation based protocol that only tests drivers every 4 weeks (and does not test the general population), a dynamic test every $\frac{2000}{Q+1}$ hours, and a reactive test.

a. Average Number of Healthy Drivers

The average number of healthy drivers across all observations on a Carrier with deterministic testing is 11.14 for the pessimistic case and 14.78 for both the optimistic and nominal case. The pessimistic averages range between 8 and 14 and the optimistic and nominal averages range between 12 and 15. The pessimistic case has a driver danger ratio of 0.169 and the remaining ratios are zero. These results are presented in Figure 27.

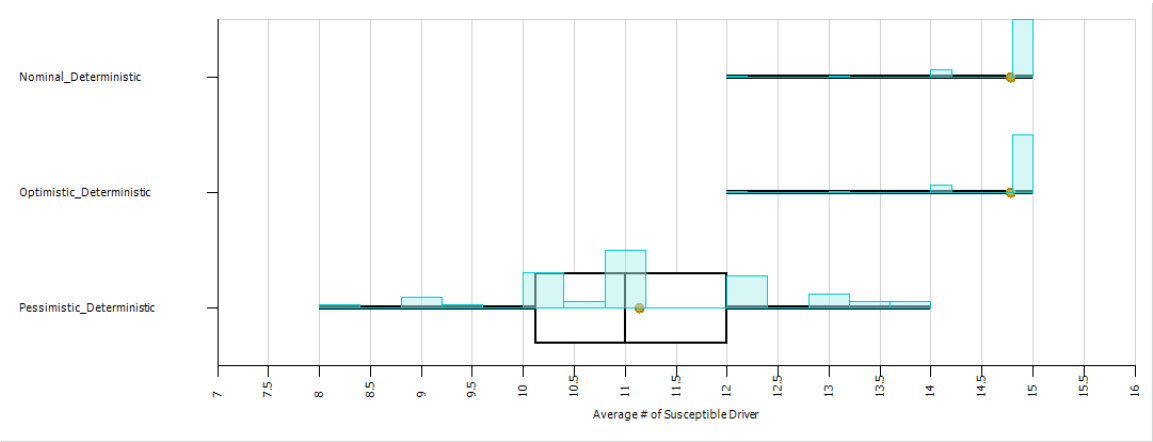


Figure 27. Carrier Healthy Drivers, Deterministic Testing

The average number of healthy drivers across all observations on a Carrier with subpopulation-based testing is 11.12 for the pessimistic case and 14.78 for both the optimistic and nominal case. The pessimistic averages range between 8.16 and 14, and the optimistic and nominal averages range between 12 and 15. All danger threshold ratios are zero.

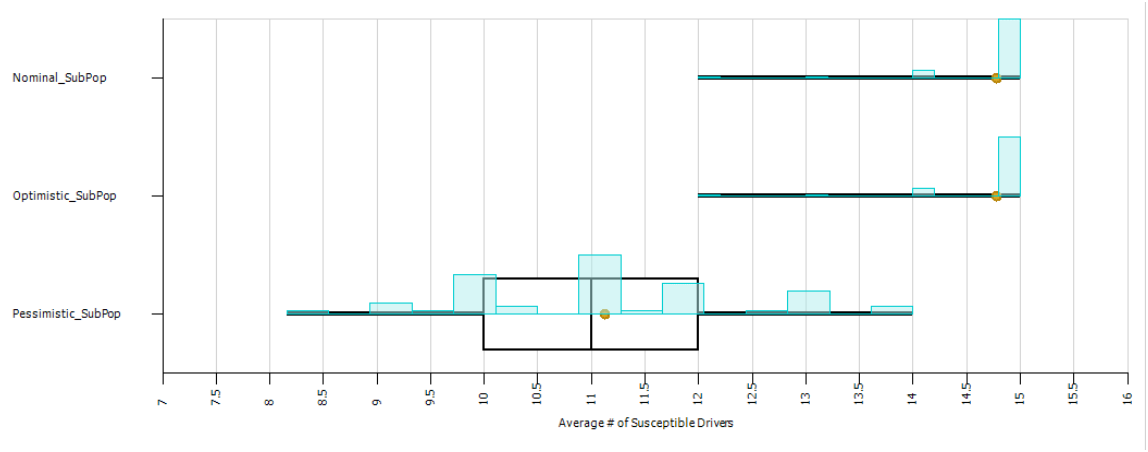


Figure 28. Carrier Healthy Drivers, Subpopulation-Based Testing

The average number of healthy drivers across all observations on a Carrier with dynamic testing is 11.16 for the pessimistic case and 14.78 for both the optimistic and nominal case. As shown in Figure 29, the pessimistic averages range between 8 and 14 and the optimistic and nominal averages range between 12 and 15. The pessimistic case has a driver danger ratio of 0.018 and the remaining ratios are zero.

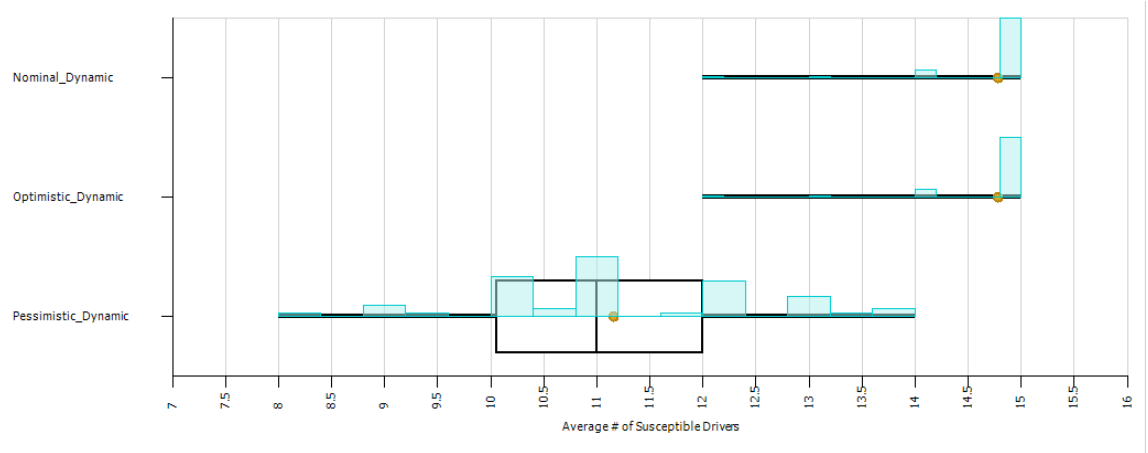


Figure 29. Carrier Healthy Drivers, Dynamic Testing

The reactive testing protocol returns an average number of healthy drivers of 14.78, ranging between 12 and 15, for both the nominal and optimistic cases. The pessimistic case

ranges between 8 and 14.31 with an average of 11.13. All danger and inoperability thresholds are zero.

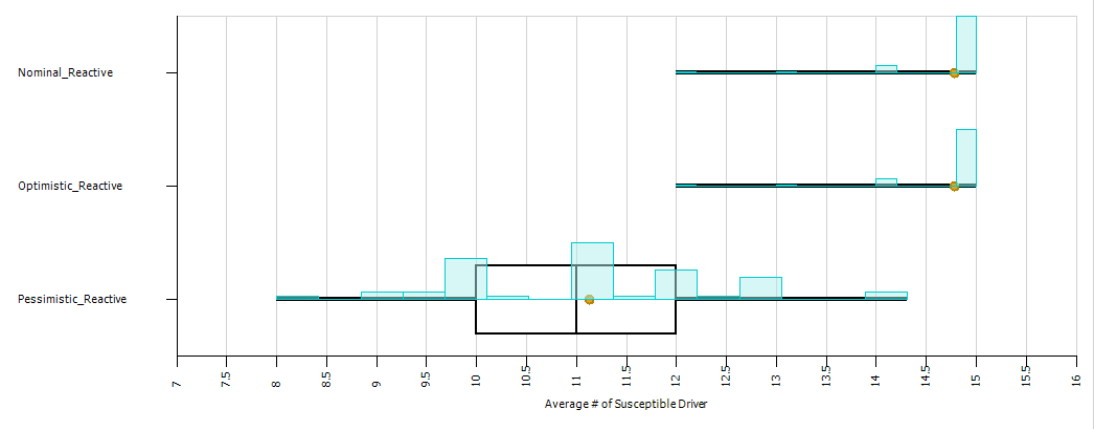


Figure 30. Carrier Healthy Drivers, Reactive Testing

b. Average Number of Healthy Sailors

The average number of healthy sailors across all observations on a Carrier with deterministic testing is 4461.80 for the pessimistic case, 5878.80 for the optimistic case, and 5878.97 for the nominal case. The pessimistic averages range between 4378.89 and 4539.14, the optimistic values range between 5855.27 and 5909.67, and the nominal averages range between 5855.31 and 5910.16. These results are presented in Figure 31.

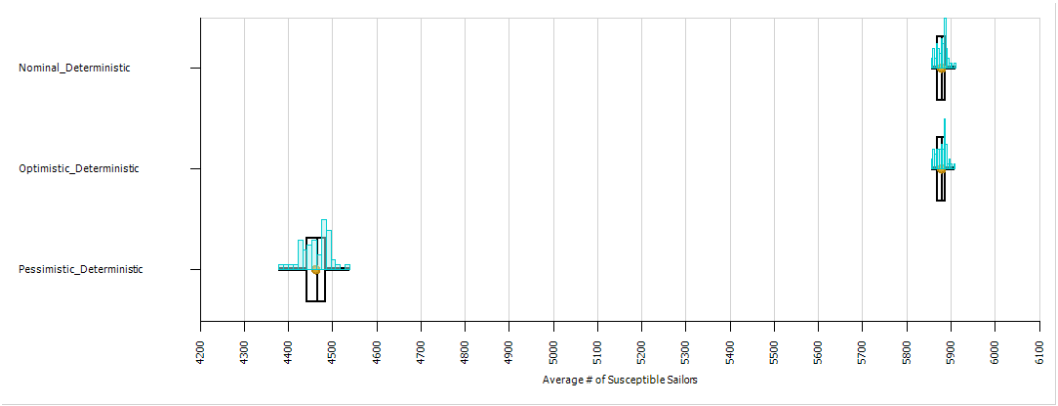


Figure 31. Carrier Healthy Sailors, Deterministic Testing

Given in Figure 32, the average number of healthy sailors across all observations on a Carrier with subpopulation-based testing is 4466.53, 5881.50, and 5881.52 for the pessimistic, optimistic, and nominal cases, respectively. The pessimistic values range between 4390.07 and 4546.36. The minimum value recorded for the optimistic case is 5857.41 and the maximum is 5912.61. The values for the nominal case fall within 5857.41 and 5912.61.

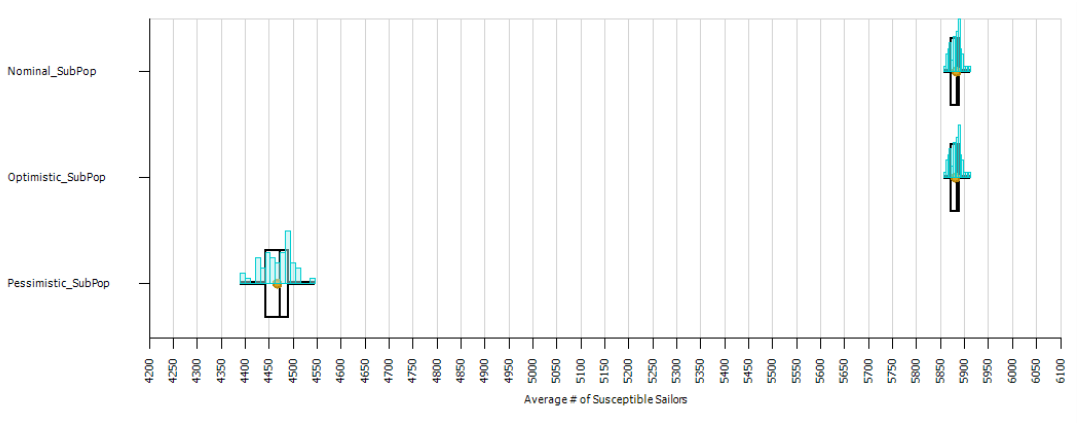


Figure 32. Carrier Healthy Sailors, Subpopulation-Based Testing

The average number of healthy sailors across all observations on a Carrier with dynamic testing is 4465.43 for the pessimistic case, 5879.48 for the optimistic case, and 5879.84 for the nominal case. The pessimistic averages range between 4388.85 and 4547.44, the optimistic averages range between 5815.76 and 5913.09, and the nominal averages range between 5831.71 and 5913.09.

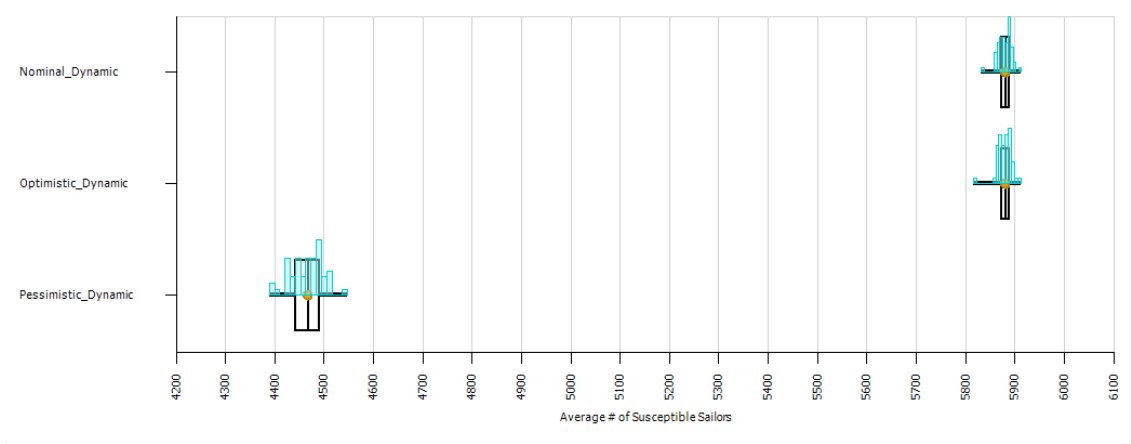


Figure 33. Carrier Healthy Sailors, Dynamic Testing

The reactive testing protocol returns an average number of healthy sailors of 5881.49 for the nominal case, 5881.48 for the optimal case, and 4466.56 for the pessimistic case. Shown in Figure 34, the nominal values range between 5857.36 and 5913.02, the optimistic values range between 5857.36 and 5912.99, and the pessimistic values range between 4389.95 and 4546.27.

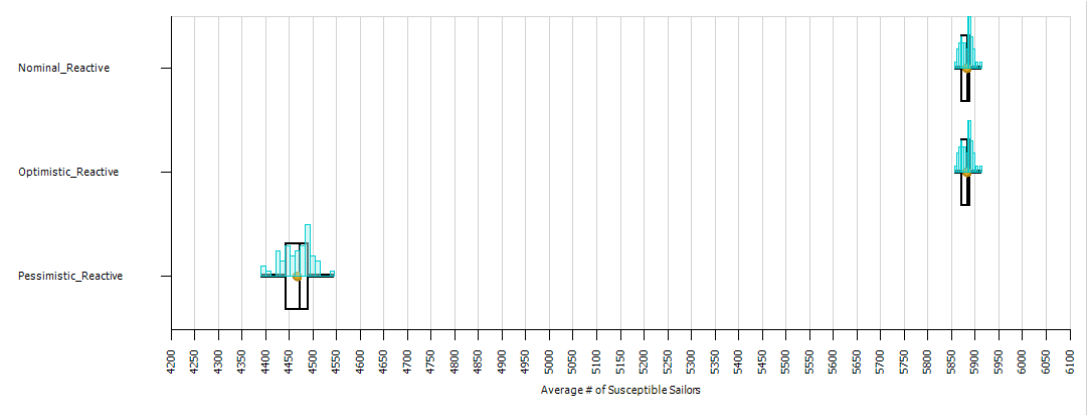


Figure 34. Carrier Healthy Sailors, Reactive Testing

c. Average Number of Infected Unaware Sailors

The average number of infected unaware sailors across all observations on a Carrier with deterministic testing is 294.87 for the pessimistic case, 28.37 for the optimistic case,

and 30.64 for the nominal case. The pessimistic averages range between 273.32 and 417.04, the optimistic averages range between 21.95 and 37.08, and the nominal averages range between 19.74 and 40.98.

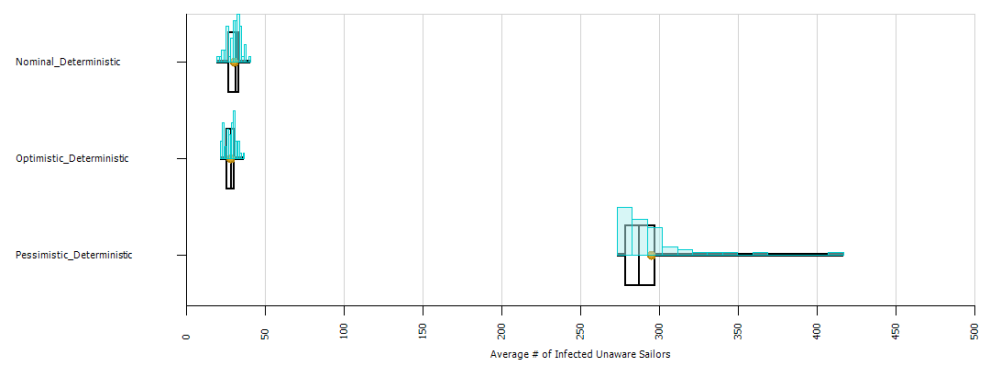


Figure 35. Carrier IU Sailors, Deterministic Testing

In Figure 36, we present the average number of infected unaware sailors across all observations on a Carrier with subpopulation-based testing. The average is 307.76 for the pessimistic case, 24.18 for the optimistic case, and 24.18 for the nominal case. The pessimistic averages range between 283.38 and 332.34, the optimistic averages range between 17.71 and 32.56, and the nominal averages range between 17.71 and 32.56 as well.

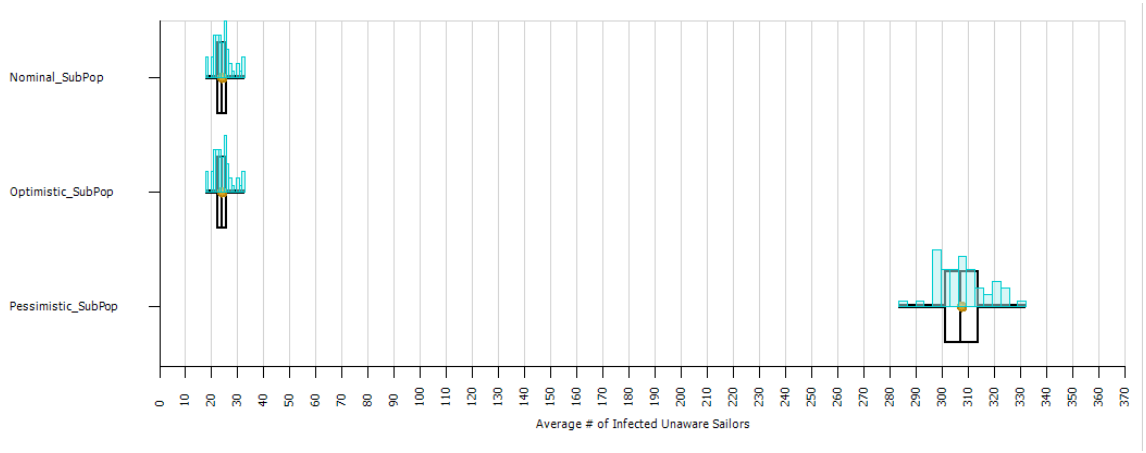


Figure 36. Carrier IU Sailors, Subpopulation-Based Testing

The average number of infected unaware sailors across all observations on a Carrier with dynamic testing is 272.21 for the pessimistic case, 21.80 for the optimistic case, and 21.82 for the nominal case. Given in Figure 37, the pessimistic averages range between 121.42 and 320.09, the optimistic averages range between 6.39 and 30.54, and the nominal averages range between 7.57 and 30.54.

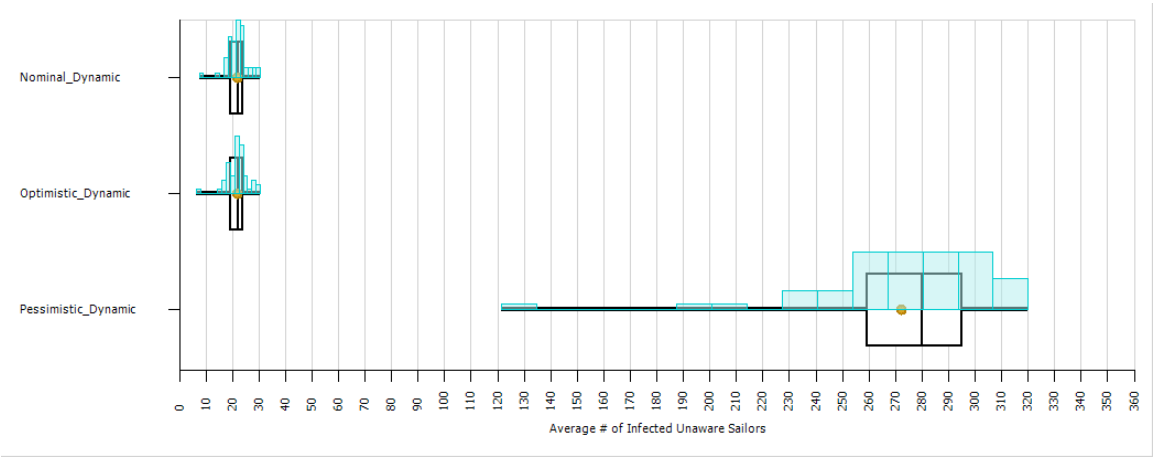


Figure 37. Carrier IU Sailors, Dynamic Testing

Under the reactive testing protocol, the average number of infected unaware sailors is 24.18 for both the nominal and optimistic cases, with values ranging between 17.71 and 32.56. The pessimistic case returns an average of 307.81 and ranges between 283.38 and 332.35.

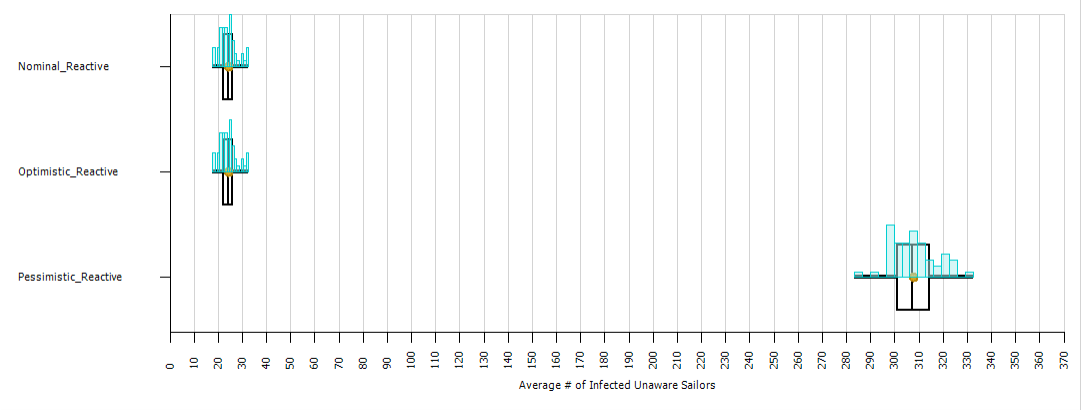


Figure 38. Carrier IU Sailors, Reactive Testing

4. Results

The results given suggest that the dynamic testing policy or the reactive testing policy may be the best approach for U.S. naval ships, depending on which metric is deemed more important by decision makers.

For all three ship types, there were very small differences between the results in the optimal case and in the nominal case. The variance of those parameters did not seem to make a difference in the metrics we reported. The pessimistic input parameters made a significant difference in every scenario analyzed. Since the Navy is interested in protecting itself against worst case scenarios (i.e., an outbreak like some of the minimum values recorded for the pessimistic cases) we focus primarily on the worst case metrics for analysis and take secondary consideration to the nominal case metrics. We present a summary of the results for Destroyers, Cruisers, and Carriers in Tables 15, 16, and 17.

The time-cost of test is calculated using the average number of sailors in the quarantine stage, which is outputted by the model report. The equation for this calculation shown for $\frac{2000}{Q+1}$ is:

$$\frac{30}{(\frac{2000}{Q+1})hr*(\frac{1day}{24hr})*16*60}*100=\frac{30*100*24}{\frac{2000}{Q+1}*16*60}$$

Table 15. Destroyer Summary of Pessimistic / Nominal Results

Test Type	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of Infected Unaware
Deterministic	11.13 / 14.76	222.57 / 293.90	20.75 / 2.30
Subpopulation-based	11.14 / 14.75	223.16 / 293.78	17.86 / 2.45
Dynamic	11.04 / 14.66	221.74 / 292.52	12.54 / 1.02
Reactive	11.14 / 14.76	223.15 / 293.86	17.83 / 2.46

The Destroyer average number of healthy drivers was similar across all four testing types, with a pessimistic value around 11 and a nominal value around 14.75. The average number of total healthy sailors was the lowest using a dynamic testing policy and reaches a maximum using the reactive testing protocol. The average number of infected unaware sailors was also lowest using dynamic testing and the ratio of danger drivers is the lowest at 0 for the deterministic testing protocol. The time-cost for the deterministic test is 0.11%, the cost of the subpopulation test is 0.001%, the time-cost of the pessimistic dynamic test is 1.03%, and the cost of the nominal dynamic test is 0.09%. The time cost for the reactive test is 0.00%. Both the pessimistic and nominal cases have testing time-cost values because as there are more infected sailors in the system, there will be more tests so the time-cost increases. Because we wish to maximize the number of healthy drivers, maximize the number of healthy sailors, and minimize the number of infected unaware sailors, the reactive and dynamic tests are best suited for Destroyers.

Table 16. Cruiser Summary of Pessimistic / Nominal Results

Test Type	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of Infected Unaware
Deterministic	11.08 / 14.78	370.31 / 490.41	33.91 / 3.12
Subpopulation-based	11.08 / 14.78	371.12 / 490.53	27.47 / 3.31
Dynamic	11.11 / 14.76	369.37 / 489.65	18.62 / 1.82
Reactive	11.14 / 14.78	371.14 / 490.55	27.62 / 3.29

The Cruiser average number of healthy drivers was similar across all four testing protocols for both the pessimistic and nominal cases. The number of healthy sailors is maximized with the reactive testing protocol and the minimum average number of infected unaware sailors is reached with the dynamic testing protocol. However, the ratio of driver danger threshold is actually highest for the pessimistic case of the dynamic testing at 0.158 and is the lowest for the reactive testing. The time-cost for the deterministic test is 0.11%, the cost of the subpopulation test is 0.001%, the time-cost of the pessimistic dynamic test is 1.28%, the cost of the nominal dynamic test is 0.13%, and the cost of reactive testing is 0.00%.

Table 17. Carrier Summary of Pessimistic / Nominal Results

Test Type	Avg # of Healthy Drivers	Avg # of Healthy Sailors	Avg # of Infected Unaware
Deterministic	11.14 / 14.78	4461.80 / 5878.97	294.87 / 30.64
Subpopulation-based	11.12 / 14.78	4466.53 / 5881.52	307.76 / 24.18
Dynamic	11.16 / 14.78	4465.43 / 5879.84	272.21 / 21.82
Reactive	11.13 / 14.78	4466.56 / 5881.49	307.81 / 24.18

The Carrier average number of healthy drivers is similar across all four testing policies, with no notable minimum for either the pessimistic or the nominal case. The average number of healthy sailors is maximized using the subpopulation-based policy or the reactive policy. The average number of infected unaware sailors is minimized with the dynamic testing policy with a value of 272.21 for pessimistic and 21.82 for nominal cases. The minimum ratio of driver danger threshold is reached with the reactive testing protocol. The time-cost for the deterministic test is 0.11%, the cost of the subpopulation test is 0.001%, the time-cost of the pessimistic dynamic test is 1.30%, and the cost of the nominal dynamic test is 0.15%. The time cost of the reactive testing protocol is 0.00%.

Based on what we discuss in this section as well as section B.5, the testing protocol is inconsequential and the Navy should focus on reactive testing of sailors when symptoms present themselves. Overtesting sailors results in the least number of infected unaware sailors, but also is a waste of time and resources and results in more false positives. Overtesting still ultimately makes no significant difference in terms of operational capabilities of the ship.

When sailors are not tested at all on a regular basis, testing the symptomatic sailors still catches enough that a reactive testing policy is just as good, if not better, than the other protocols. Saving time and testing resources, as well as the lack of false positives, makes reactive testing the best policy to implement on U.S. Navy ships. This result is the same regardless of ship size, as long as initial prevalence of the virus is low and sailors continue to self-report when they show symptoms.

V. CONCLUSION

Over the past several months we have created an adaptation on the SEIR model that accounts for non-homogeneous mixing, tracks subpopulations of interest over time, and is able to run three different testing procedures to compare. These three different testing procedures are: (1) a deterministic test of all sailors on board after a predetermined, fixed interval, (2) a subpopulation-based testing interval where a subpopulation, like "drivers" is tested more frequently than the rest of the ship's population, and (3) dynamic testing, where more tests are administered as more sailors become sick and enter quarantine and less tests are administered as fewer sailors are sick.

An analysis is conducted on these three testing protocols based on three different ship scenarios for Destroyers, Cruisers, and Carriers. Each was run first by varying the time between testing events (or how the time between was calculated, in the dynamic testing case) and a testing delay was selected for each ship size and protocol type. From here, we add a fourth testing protocol to represent never testing the general population and to only test when a sailor present symptoms. With these four testing protocols, another analysis was conducted using varying input parameters using nominal, pessimistic, and optimistic values.

It is found that the nominal and optimistic input values returned similar, if not the same, results within the models. To help prevent the "worst case" scenario, the pessimistic scenario output values became our primary metric of interest. The metrics used are the average number of healthy drivers over all observations, the average number of healthy sailors over all observations, the average number of infected unaware sailors over all observations, the ranges (maximum and minimum recorded values by the simulation) of each of these metrics, the probability (given by a ratio of time) that the number of healthy drivers will dip below a dangerous level, and the time-cost of testing.

At the end of Chapter IV it is reported that the metrics for each scenario do not vary significantly. Testing aggressively results in a lower number of infected unaware sailors but results in a higher number of false positives, which makes it seem like there are more

healthy sailors if there are less frequent tests. As long as the initial number of infected unaware sailors is low, the reactive testing protocol results in a larger number of infected unaware sailors but a small number of false positives. The total number of sailors with positive test results is similar, leading to similar operational effects on the ship. Our conclusion from the simulation model is that testing doesn't matter and each testing protocol is inconsequential. Because it does not test the general population, reactive testing only tests the symptomatic sailors. It results in less cost of both time and resources and therefore is the optimal testing protocol.

This model is far from perfect, as it is impossible to create a simulation model that mirrors all aspects of ship-life accurately. Many social tendencies were ignored in this project for simplicity. For example, only one subpopulation was created to represent ship drivers. In future research, more subpopulations should be created to better represent ship life and the interactions between groups could be better quantified. Sensitivity analysis should be conducted on the values of the mixing parameters, κ, as well. A second, more sensitive testing procedure should be inputted into the model to minimize false positives. Finally, the input parameters could be modified to include vaccinated sailors, outside interactions during port calls, et cetera.

At the conclusion of this project, we aspire to have provided a useful analysis of the four testing protocols and have proven the insignificance of which is chosen for all ship sizes in the Navy as long as initial prevalence is low on board. We are hopeful that the methods we used can be applied for future situations with COVID-19 or be expanded to model future epidemics.

LIST OF REFERENCES

Anon U.S. Navy's Values and Missions. military.com. Retrieved May 20, 2021, https://www.military.com/join-armed-forces/navy-missions-values.html.

Cashore M, Duan N, Janmohamed A, Wan J, Zhang Y, Henderson S, Shmoys D, Frazier P (2020) COVID-19 mathematical modeling for Cornell's fall semester.

Chick S (2020) Health resource allocation: Lessons for today from past outbreaks. Presentation, Winter Simulation Conference, 15 December 2020. https://informs-sim.org/wsc20papers/by_area.html.

Cui J, Zhang Y, Feng Z (2018) Influence of non-homogeneous mixing on final epidemic size in a meta-population model. *J. Biol. Dyn* 13:31–46.

Department of the Navy (2020) U.S. Navy COVID-19 standardized operational guidance version 3.0. NAVADMIN 266/20. https://www.navy.mil/Resources/NAVADMINs/Message/Article/2366564/us-navy-covid-19-standardized-operational-guidance-version-30/.

Kress M (2005) The effect of social mixing controls on the spread of smallpox—A two-level model. *Health Care Manag. Sci.* 8(4):277–289.

Letizia AG, Ramos I, Obla A, Goforth C, Weir DL, Ge Y, Bamman MM et al. (2020) SARS-CoV-2 Transmission among Marine recruits during quarantine. *N. Engl. J. Med* 383(25):2407–2416.

Lin Y, Ren Y, Wan Jingyuan, Cashore M, Wan Jiayue, Zhang Y, Frazier P, Zhou E (2020) Group testing enables asymptomatic screening for COVID-19 mitigation: Feasibility and optimal pool size selection with dilution effects. *ArXiv200806642 Math Q-Bio Stat*.

Long EF, Nohdurft E, Spinler S (2018) Spatial resource allocation for emerging epidemics: A comparison of greedy, myopic, and dynamic Policies. *Manuf. Serv. Oper. Manag.* 20(2):181–198.

Luo L, Liu D, Liao X long, Wu X bo, Jing Q long, Zheng J zhen, Liu F hua et al. (2020) Modes of contact and risk of transmission in COVID-19 among close contacts. *medRxiv*. https://www.medrxiv.org/content/10.1101/2020.03.24.20042606v1.full.

Mizumoto K, Chowell G (2020) Transmission potential of the novel coronavirus (COVID-19) onboard the diamond Princess Cruises Ship, 2020. *Infect. Dis. Model* 5:264–270.

Mizumoto K, Kagaya K, Zarebski A, Chowell G (2020) Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance* 25(10):2000180.

Pray IW (2021) Performance of an antigen-based test for asymptomatic and symptomatic SARS-CoV-2 testing at two university campuses — Wisconsin, September–October 2020. *MMWR Morb. Mortal. Wkly. Rep.* 69.

Regnier E, Sanchez SM, Sanchez PJ (2020) Testing-based interventions for COVID pandemic policies. *Proc. 2020 Winter Simul. Conf.* 10.1109/WSC.2015.7408160.

Rocklöv J, Sjödin H, Wilder-Smith A (2020) COVID-19 outbreak on the Diamond Princess cruise ship: estimating the epidemic potential and effectiveness of public health countermeasures. *J. Travel Med* 27(3).

Rowthorn RE, Laxminarayan R, Gilligan CA (2009) Optimal control of epidemics in metapopulations. *J. R. Soc. Interface* 6(41):1135–1144.

Sanchez PJ, Sanchez SM (2015) A scalable discrete event stochastic agent-based model of infectious disease propagation. *2015 Winter Simul. Conf. WSC*. 151–158.

Van Beusekom M (2021) Previous COVID-19 may cut risk of reinfection 84%. *CIDRAP*. Retrieved June 7, 2021, https://www.cidrap.umn.edu/news-perspective/2021/04/previous-covid-19-may-cut-risk-reinfection-84.

Vera DM, Hora RA, Murillo A, Wong JF, Torre AJ, Wang D, Boulay D et al. (2014) Assessing the impact of public health interventions on the transmission of pandemic H1N1 influenza a virus aboard a Peruvian navy ship. *Influenza Other Respir. Viruses* 8(3):353–359.

Zaric GS, Brandeau ML (2001) Resource allocation for epidemic control over short time horizons. *Math. Biosci* 171(1):33–58.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California