Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

MULTI-CRITERIA SIMULATION OPTIMIZATION FOR COVID-19 TESTING IN SCHOOLS

Yiwei Zhang

Operations Research
North Carolina State University
915 Partners Way
Raleigh, NC 27606, USA

Maria E. Mayorga
Julie S. Ivy
Julie L. Swann

Edward P. Fitts Department of Industrial and Systems Engineering
North Carolina State University
915 Partners Way
Raleigh, NC 27606, USA

ABSTRACT

Evidence has shown that random screening tests are effective in reducing COVID-19 infections in schools. However, test administration may be hindered due to a limited budget or low participation caused by pandemic fatigue. Thus, we seek to balance the number of tests administered with end-of-semester infections. To do this we use an SEIR model to simulate SARS-CoV-2 transmissions within K-12 schools, design a multi-objective simulation optimization problem, and tune an NSGA-II algorithm to find the best testing schedules. We find the Pareto front of optimal schedules of screening tests, which can be used by stakeholders to inform test administration strategies. We discuss insights about the characteristics of optimal strategies, for example, when there are limited number of tests available or a desire to use few tests, the optimal plan is to perform the tests earlier in the semester and at higher intensity.

1 INTRODUCTION

COVID-19 has dramatically disrupted the global healthcare system. In the US alone, the virus has been responsible for over 5.7 million hospital admissions (CDC 2023a) and over 1 million deaths (DCD 2023). For people below the age of 18, the disease has caused more than 177 thousand hospital admissions (CDC 2023a) and 1.6 thousand deaths (CDC 2023b). As the main places for children to study and socialize, schools can play a very important role in protecting children from infection. The measures taken by schools directly impact disease transmission and the health of students in schools.

Through disease modeling, evidence shows that testing policies have been useful tools for keeping COVID-19 incidence low in schools. In 2021, Asgary et al. (2021) simulated preventive testing policies in schools using an agent-based model to verify the effectiveness of testing policies within schools. They analyzed and compared a variety of testing scenarios, such as test frequency, test result days, and test expiration days. They concluded that frequent testing, expedited test results, and self-isolation together are effective in controlling in-school infections. In our previous work, we also demonstrated the effectiveness of random screening tests implemented in K-12 schools (Zhang et al. 2022), in particular when used as one of several layered interventions.

There also exist many studies that apply optimization methods while analyzing COVID-19 screening testing strategies. Van Pelt et al. (2021) utilized a decision tree model to optimize the detection of true positives and negatives as well as the number of tests needed among college students by designing and assessing five different strategies associated with RT-PCR tests. Aragón-Caqueo et al. (2020) built a mathematical model to optimize the group size in pool testing using RT-PCR tests, which as a result,

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

was able to boost the testing capacity. Abdin et al. (2023) proposed a non-linear programming model to minimize the spread of COVID-19 by optimizing the allocation of limited testing resources in France.

However, we noticed that K-12 school students and teachers were rarely considered as the target population in the studies of screening tests and optimization algorithms were not commonly used in the analysis of COVID-19 testing schedules (Jordan et al. 2021). In addition, almost all the studies we reviewed analyzed PCR tests instead of rapid antigen tests. Thus, we proposed this study to focus on screening tests using rapid antigen tests in K-12 schools and apply an optimization algorithm to solve the problem.

As the highly contagious subvariant of Omicron spread in late 2022 and early 2023, some school districts were supported by the government to utilize tests as a measure to protect children in school. For instance, the North Carolina Department of Health and Human Services (NCDHHS) encouraged schools to participate in the state-funded testing program for the 2022-23 school year, in which schools were responsible for deciding on their own testing plans and providing guidance for quarantine and isolation (NCDHHS 2022). Our goal is to provide guidance to decision makers about the most effective and efficient ways to administer random screening tests in schools. Therefore, in this study, we use an SEIR simulation model to capture the disease transmission within K-12 schools, then, we design an optimization problem based on the simulation model. Because there are trade-offs between testing and infection reduction, we modeled a bi-criteria optimization problem with the number of end-of-semester infections and the total number of screening tests administered as the objectives. The decision variables are the timing and intensity of tests over a semester. By solving the multi-criteria simulation optimization problem, this study aims to help K-12 schools optimally plan test administration over the entire semester using simulation optimization. In addition, due to the large decision space and the multi-objective nature of the problem, the NSGA-II algorithm is tuned and used to solve the optimization problem.

2 METHODS

2.1 Simulation Model Overview

The simulation model we used in the study was adapted from our previous work (Zhang et al. 2022). It is a multi-grouped SEIR model specific to K-12 schools, where individuals transition between 10 different disease states with a one-day time step, as shown in Figure 1. To capture the different social dynamics between students from different classrooms and teachers, we split the school population into subgroups. We parameterized three distinct contact matrices to accommodate the different social mixing patterns within elementary (K-5), middle, and high schools. We simulated the start of the Spring 2022 semester. As such, the parameters were chosen according to the Omicron variant and the best evidence at the time.

In our previous work, we mainly focused on quantifying the effect of different non-pharmaceutical interventions (NPIs) implemented within K-12 schools during the pandemic. Then, we estimated the impact of different and potentially layered NPIs, including baseline interventions of masks and random screening, as well as three strategies of contact reduction, school closures, and test-to-stay on COVID-19 transmission within K-12 schools. We focused what combinations of interventions were most effective in mitigating disease transmission within school settings under different assumptions about the school population (such as incoming protection). Under the strategy of contact reduction, our simulation model achieved face validity because the level of reduction in infections was consistent with Boutzoukas et al. (2022)'s analysis when the universal masking policy was implemented.

In this paper, we extended the analysis by adding the design of a multi-objective optimization problem to the simulation model. We also made some crucial changes to the model structure. These changes include updates to capture changes in the new subvariants of Omicron, refinement to allow for day-to-day decision-making, and updates based on changes in human behavior.

In late 2022 and early 2023, BA.4, BA.5, BQ.1, BQ.1.1, and XBB.1.5 took turns becoming the dominant subvariants of Omicron (Zimmer, Carl 2023). They are known to be able to dodge antibodies produced by vaccines or previous infections (Uraki et al. 2022), which means people who recovered from the disease

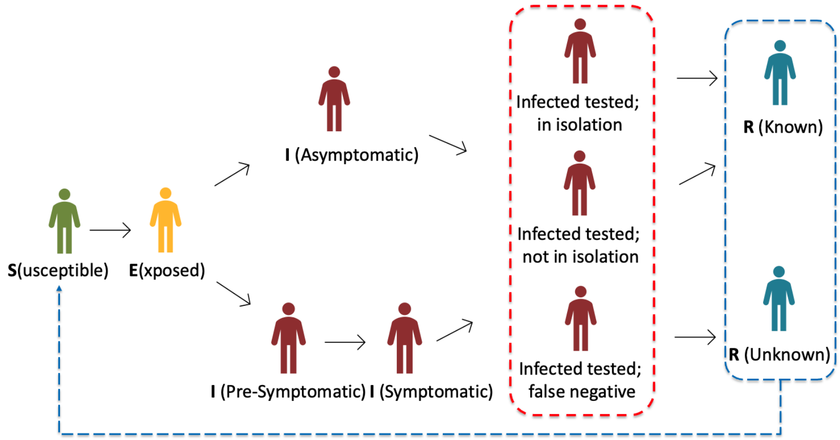

Figure 1: Flowchart of the COVID-19 disease model, including the transition from recovered to susceptible shown in blue added as part of this work.

can become infected again. Thus, we added the transition of reinfection to the previous model; that is, people who are recovered from the disease, can transition back to the "susceptible" state at a later date, as shown by the blue line in Figure 1. We assumed that the reinfection rate of the subvariants of Omicron was 13%, based on (Özüdoğru et al. 2022). We assume that on average, it takes about 100 days for a recovered individual to be susceptible to infection again (Nordström et al. 2022). Furthermore, for the optimal solution to return a precise screening schedule, we updated the model to treat weekdays and weekends differently, with the transmission rate within schools being zero on weekends. Note that we allow for outside infections that students bring to school from outside of school, as later discussed. One semester's total number of days is assumed to be 107 days, with a starting day of Monday based on the North Carolina school calendar.

Additionally, we updated the model to account for changes in recommendations and human behavior in early 2023 compared to early 2022. For instance, the CDC now suggests an isolation time of 5 days, which is shorter compared to when the original strain of Omicron was dominant. We assumed the social mixing among students and teachers is "well-mixed" with 0% contact reduction, representing that social contact patterns have returned to pre-pandemic levels with social distancing policies rescinded. The "well-mixed" contact pattern implies that each individual has an equal probability of making contact with other individuals from the same or different population groups. It is possible for symptomatic individuals to self-quarantine. However, because COVID-19 symptoms are often mild in children and can be confused with a common cold, we assume a low, 20%, level of self-quarantine. Moreover, we increased the level of incoming protection to represent a higher percentage of previous infection or vaccination among school-aged students and teachers. Lastly, we considered the isolation due to testing compliance rate to be 100%. This assumes the screened testing results are strictly enforced in schools. Thus the results represent the highest benefits that can be obtained from random screening. The full list of model parameters is provided in Appendix A.

We assume a school population of 500 individuals. As mentioned earlier, students can bring infections to the school from outside. To simulate a dynamic change of the disease prevalence in the community, we created discrete uniform random number generators $Y_i \overset{iid}{\sim} DU[3,4]$ and $Y_i \overset{iid}{\sim} DU[1,2]$, representing the newly exposed individuals coming from exposures occurring outside of school each week during the first month of the school semester, and a lower level of newly exposed each week later on, respectively. This was designed to simulate the surge of COVID-19 cases at the beginning of the semester after summer vacation or winter break. In addition, we assumed one student was already exposed to the disease at the beginning of the semester, which represents the virus invasion into the school environment.

In this study, we focused on rapid antigen tests, not PCR tests, as rapid antigen tests do not require lab work to retrieve results, which makes them desirable for screening purposes. One disadvantage of rapid antigen tests is the low accuracy of the testing results for asymptomatic people. However, the model contains a post-testing state of "Infected tested; false negative", thus it is able to accommodate the impact of undetected infections. The sensitivity rate and specificity rate shown in Table 1 (Dinnes et al. 2022).

Table 1: Sensitivity rate and specificity rate of symptomatic and asymptomatic people.

	Symptomatic	Asymptomatic
Sensitivity	73%	54.7%
Specificity	99.1%	99.7%

Based on the values of the disease parameters, the SEIR model was used to simulate the disease progression within the school over the entire semester. While the SEIR model is a deterministic compartmental model, because we generate the "newly exposed" using discrete uniform random number generators, different seeding results in distinct projections. When performing the parameter tuning for NSGA-II, we repeat the optimization process with different seeding and compare the results.

2.2 Problem Definition

The optimization problem designed in this study is aims to help K-12 schools find the optimal schedule for random screening tests. Specifically, when proposing a schedule of random screening for the semester, the schedule is broken down week by week. During each week, the school determines whether to schedule random screening tests for that week or not. If tests are scheduled, it is assumed that the possible test ratio is 20%, 40%, 60%, or 80%, which means that among all students and teachers who are present at school on that day, 20%, 40%, 60%, or 80% of them will be randomly selected to receive a rapid antigen test, respectively. Because we assume the isolation compliance rate is 100% despite symptoms, all the individuals who receive positive results will be isolated. If they do not schedule a set of random screening tests that week, the test ratio is set to be 0%. We call this planned testing.

By distinguishing weekdays and weekends in our simulation, we were able to specify the testing day for each week. The frequencies of planned testing we considered are 1, 2, or 5 days per week. If one planned test is scheduled during the week, it is set to performed on Wednesday. Two planned testing days are scheduled on Wednesday and Friday. Five planned testing days means random screening tests are performed every weekday. Note that the test ratio remains constant within the week regardless of the testing frequency, meaning that if 20% is chosen and we are testing two days then we test 20% of individuals each day. The model runs separately for different testing frequencies, and the testing frequency remains unchanged during the entire semester within each run. In other words, we optimize the weeks selected for testing and intensity of testing separately for one day, two days, and five days of testing per week. In future work we discuss the possibility of combining these decisions.

Thus, we defined a set of decision variables $\mathbf{X} = \{X_1, ..., X_j, ..., X_{16}\}$, where X_j represents the ratio of planned testing during week j. The possible values for X_j are 0%, 20%, 40%, 60%, and 80%. 0% means that there will be no planned testing for that week. The two cost functions together comprise our objective: C_1 is the end-of-semester infections within the school; C_2 is the number of random screening tests performed. These are obtained as a result of the simulation. Note that C_2 equals the sum of the test ratio multiplied by the total number of individuals in school on test days. Thus, the values of the cost functions change as the testing schedule changes. By minimizing the two objectives together, we aimed to determine the optimal weekly schedule and intensity of tests for the entire semester. The formulation of the multi-objective optimization problem is shown in Equation (1):

$$\begin{aligned} \min_{\mathbf{X}} \quad & F = (C_1(\mathbf{X}), C_2(\mathbf{X})) \\ \text{s.t.} \quad & X_j \in \{0\%, 20\%, 40\%, 60\%, 80\%\} \quad j = 1, 2, ..., 16 \end{aligned} \tag{1}$$

X_j is the planned testing rate of week j,
C_1 is the end-of-semester infection,
C_2 is the number of screening tests performed.

2.3 Parameter Tuning for NSGA-II

The Non-dominated Sorting Genetic Algorithm II (NSGA-II) is a powerful evolutionary algorithm that specializes in solving optimization problems with multiple objectives and large solution spaces. Within each iteration, the algorithm sorts the parent and offspring population of solutions and keeps the results that belong to the best non-dominated set for the next iteration until iterations are exhausted. To ensure that the solutions are diverse, solutions corresponding to the largest crowding distance will also be selected for the next iteration. Thus, in a single simulation run, the algorithm is able to find multiple diverse Pareto front solutions (Deb et al. 2002; Anagnostopoulos and Mamanis 2010; Adyatama, Arga 2020).

There are four main parameters in the NSGA-II algorithm: number of generations, population size, probability of crossover (P_c), and probability of mutation (P_m). Distinctive choices of the combination of these parameters may affect the solutions found by the algorithm. Thus, we performed parameter tuning for the NSGA-II's P_c and P_m parameters to check whether there is a better combination of these two parameters before taking a deeper look at the simulation optimization results.

3 RESULTS

The simulation model we used in the study is capable of simulating elementary schools (K-5), middle schools, and high schools separately. As previously stated, K-5, middle schools, and high schools have slightly different assumptions of structures and age-specific parameters. However, in this paper we limited our analysis to a K-5 school size of 500 individuals with a "well-mixed" contact pattern.

3.1 Statistical Analysis for Parameter Tuning

This multi-objective optimization problem (MOOP) was solved using the function "nsga2R" in the R package called "nsga2R". Here, we fixed the number of generations and the population size within each generation to be 100 and 50, respectively. These two values were chosen based on previous parameter tuning literature (Samsuri et al. 2019), computation time, and initial testing which show little improvement in solutions as these values were increased further. Among the 50 initial parents, two of them were assumed to be: 0% testing ratio every week and 80% testing ratio every week during each run. This helped to reduce the run time by seeding the two extreme solutions.

As seen from other studies, the conventional choice for P_c is 0.9, and $1/N$ for P_m, where N corresponds to the number of variables in the optimization problem, which in our case was 16 (Samsuri et al. 2019; Wang et al. 2019; Feng et al. 2016). Then, we expanded the possible values for P_c and P_m and performed statistical analysis on the two parameters of the NSGA-II. More specifically, we treated P_c and P_m as two independent variables. The possible values for P_c were 0.88, 0.9, and 0.92. For P_m, the 8 possible values we considered range from $1/N$ (0.0625) to $1/2N$ (0.03125) with 0.005 increment.

For each of the parameter combinations, we ran the simulation optimization process and obtained a set of Pareto front solutions, that is a set of non-dominated solutions. Thus, it is not possible to directly compare all solutions in the Pareto front between parameter combinations, as each may produce a different number and set of solutions. Thus, we needed a way to compare the set of solutions.

Since this is a minimization problem, when plotting the Pareto front solutions with the x-axis and y-axis representing the two cost functions (shown in Figure 3), the closer the optimal solutions to the axes the better. The Manhattan distance of a point is usually used to examine how far a point is from the axes and is calculated by summing up the distances between the x and y coordinates in the two-dimensional coordinate system. However, from Figure 3 we can see that the x-axis and y-axis have very different scales. Thus, we standardized the distances by dividing the distances by the maximum value of the corresponding axes. For each simulation run, we averaged the Manhattan distances of all the points in the Pareto front. In summary, we calculated the average standardized Manhattan distance of the optimal solutions in the Pareto front set and treated the results of distance as the response variable in the parameter tuning analysis.

One simulation run (for a parameter combination) of 100 generations took approximately 3.75 hours. Overall, the experiment was replicated 10 times with different initial seeding. Table 2 summarizes the mean and standard deviation of the Manhattan distance over the 10 replications under each combination of P_c and P_m.

Table 2: The mean and standard deviation of the Manhattan distance under each combination of P_c and P_m.

	P_m							
	0.03125 (n=10)	**0.0325** (n=10)	**0.0375** (n=10)	**0.0425** (n=10)	**0.0475** (n=10)	**0.0525** (n=10)	**0.0575** (n=10)	**0.0625** (n=10)
P_c **= 0.88**								
Mean (SD)	1.3423 (0.0092)	1.3350 (0.0145)	1.3381 (0.0164)	1.3432 (0.0121)	1.3397 (0.0124)	1.3310 (0.0101)	1.3425 (0.0116)	1.3361 (0.0172)
P_c **= 0.9**								
Mean (SD)	1.3326 (0.0099)	1.3355 (0.0143)	1.3431 (0.0173)	1.3417 (0.0114)	1.3416 (0.0118)	1.3421 (0.0145)	1.3397 (0.0148)	1.3395 (0.0168)
P_c **= 0.92**								
Mean (SD)	1.3381 (0.0102)	1.3429 (0.0118)	1.3407 (0.0138)	1.3374 (0.0096)	1.3497 (0.0103)	1.3333 (0.0097)	1.3339 (0.0119)	1.3346 (0.0158)

Figure 2 illustrates the relationship between the average values of the response variable. The nonparallel lines in the interaction plot indicate that the relationship between P_c and the Manhattan distance depends on the value of P_m. Both Table 2 and Figure 2 suggested that the best two pairs of P_c and P_m were 0.88 and 0.0525, 0.9 and 0.03125, since they corresponded to the two smallest distance values.

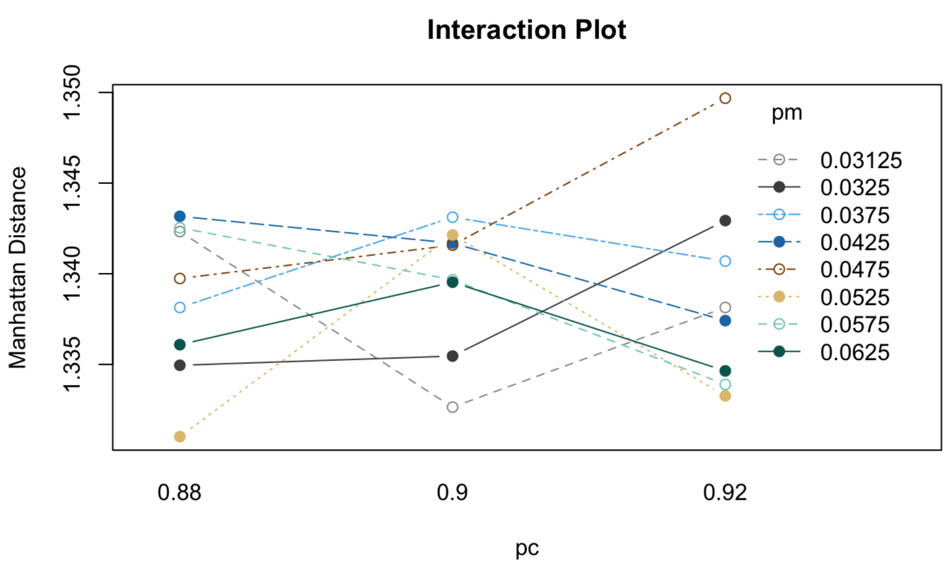

Figure 2: Interaction plot of the mean Manhattan distance under different combinations of P_c and P_m values.

Then, we performed a two-way ANOVA test on the results we obtained through simulation. The main effects as well as their interaction effect were considered. P-values in Table 3 suggested that there were no significant differences between the sample means.

Table 3: The result of two-way ANOVA with interaction effect.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
pc	2	0.00	0.00	0.12	0.8896
pm	7	0.00	0.00	1.21	0.2972
pc:pm	14	0.00	0.00	1.21	0.2672
Residuals	216	0.04	0.00		

The parameter tuning process suggested that the different values we assigned to P_c and P_m did not make a large impact on the Pareto front obtained through the algorithm. From Table 2 we saw that although the pair of P_c and P_m being 0.9 and 0.03125, was associated with the second-lowest mean Manhattan distance, it corresponded to a slightly smaller standard deviation than the pair associated with the lowest mean Manhattan distance, which indicates the stability of this group. Thus, we chose this pair and proceeded with more analysis.

3.2 Results with One Planned Testing Day Per Week

After determining the optimal choices for P_c and P_m, more random screening scenarios were considered in the study. By minimizing the end-of-semester infections and the number of random screening tests needed at the same time, we successfully obtain the optimal solution set (shown in Figure 3). Each of the optimal solutions within the set can be interpreted as the lowest number of tests needed for the end-of-semester infections to reach a certain level. For instance, the leftmost point in Figure 3 represents that the number of end-of-semester infections is 301 when no random screening tests were performed during the entire semester. Compared to the rightmost point, we can see that 31 infections were reduced when almost 6,000 tests were used. In a school of 500, this corresponds to about 12 tests per person over the semester. Thus, a lot of tests must be conducted to reduce infections by only a small about. This is due to several factors, such as limiting tests to once per week and lack of other interventions. We explore this further later.

Before we explore other results, we look at the solutions more closely. Each point in the Pareto Front corresponds to a specific week-by-week testing schedule. Figure 4 demonstrates the testing ratios in each week under different schedules. The schedules were listed in descending order of the end-of-semester infections, which were labeled on top of each corresponding schedule. For example, Schedule 1 (never test) results in 301 infections. Comparing Schedule 2 to 5, we can see that if testing were scheduled during the same week, the higher the testing ratio, the lower the end-of-semester infections. In addition, the first half of the heatmap also suggested that if the number of weeks for testing is limited, it is always optimal to arrange them earlier during the semester. Furthermore, we see that when testing does occur it is usually done at a high level (as seen by the yellow color in Figure 4), and that most policies are monotonic (they start at stop once). While not shown here, we tested the results by changing parameters (such as the number of initial infections) and found that the structure of the optimal policy was robust to changes in parameters.

3.3 Different Testing Frequencies

The problem was solved separately under the different testing frequency assumptions. In general, increasing the frequency allows more tests to be performed. The first row of Figure 5 demonstrates the Pareto front under different testing frequencies without the usage of masks. Performing 5 planned tests per week, 30,000 tests resulted in reducing 86 infections compared with no tests. Additionally, higher testing frequency makes screening tests more efficient at reducing infections. For example, if the random screening tests were scheduled 2 times per week, performing 10,000 random screening tests reduced the end-of-semester

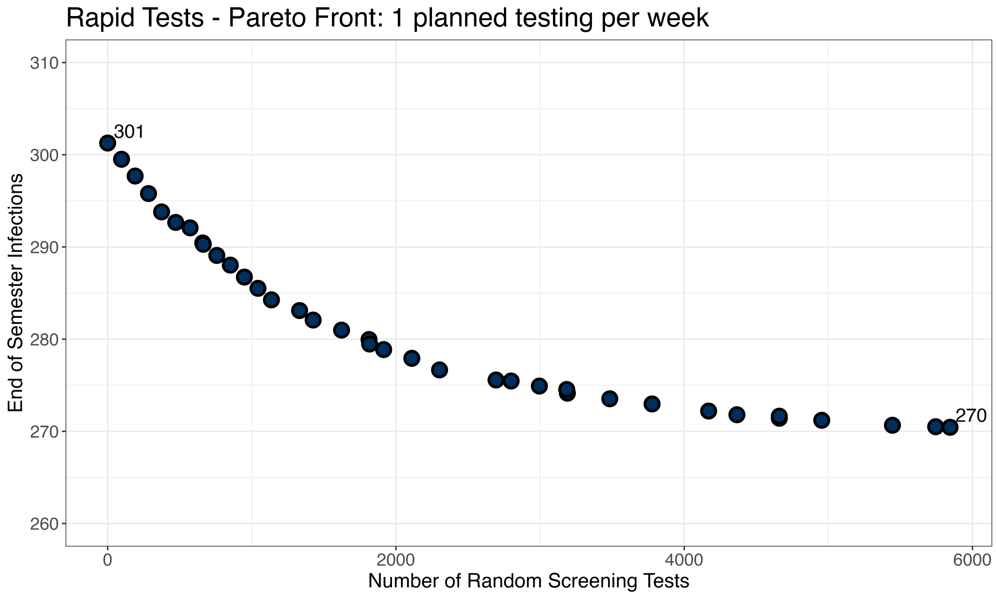

Figure 3: Pareto front of one planned testing per week.

infection to 257. By comparison, if increasing the testing frequency to 5 times per week, the same number of tests was able to reduce the end-of-semester infection to 250. This implies that there might be a benefit to more frequent tests over shorter periods of time.

3.4 Layered Intervention of Masks

Almost all K-12 schools lifted mask mandates in early 2022 (The Center for Disability, Equity, and Intersectionality 2022), as such the previous results assume no masks are in place. Unfortunately, evidence shows the removal of masks increased infections among students and staff in the Massachusetts school district (Cowger et al. 2022). Raifman and Green (2022) also stated that universal masks are effective as this measure reduces the overall density of the virus in shared air, which decreases the chances of being exposed within the environment. Thus, as the highly contagious subvariants of Omicron spread in late 2022 and early 2023, some school districts are brought back mask recommendations after winter break (Kekatos, Mary 2022).

In the study, an additional scenario of universal masking policy was considered. More precisely, masks were assumed to be utilized during the first 4 weeks of the semester. When masks are recommended, we assumed the overall effectiveness of masks is 25%, which is calculated by multiplying the mask efficacy of 50% by the mask adherence level of 50%. The second row of Figure 5 demonstrates the Pareto front results of testing when utilizing universal masking policy as a layered intervention under different testing frequencies. By comparing the results between the first and second rows of Figure 5, we can see that the universal masking policy itself could reduce the end-of-semester infections from 301 to 292, even if no tests are administered. Furthermore, a universal masking policy and random screening tests together provided additional benefits in reducing infections. For example, infections were reduced from 215 to 200 when 30,000 tests were administered (5 day testing frequency) when masks were utilized the first 4 weeks of the semester.

4 CONCLUSIONS AND FUTURE WORK

In the study, we designed and solved a simulation optimization problem to help K-12 schools plan their random screening test schedules using NSGA-II. The results provided multiple optimal schedules of random

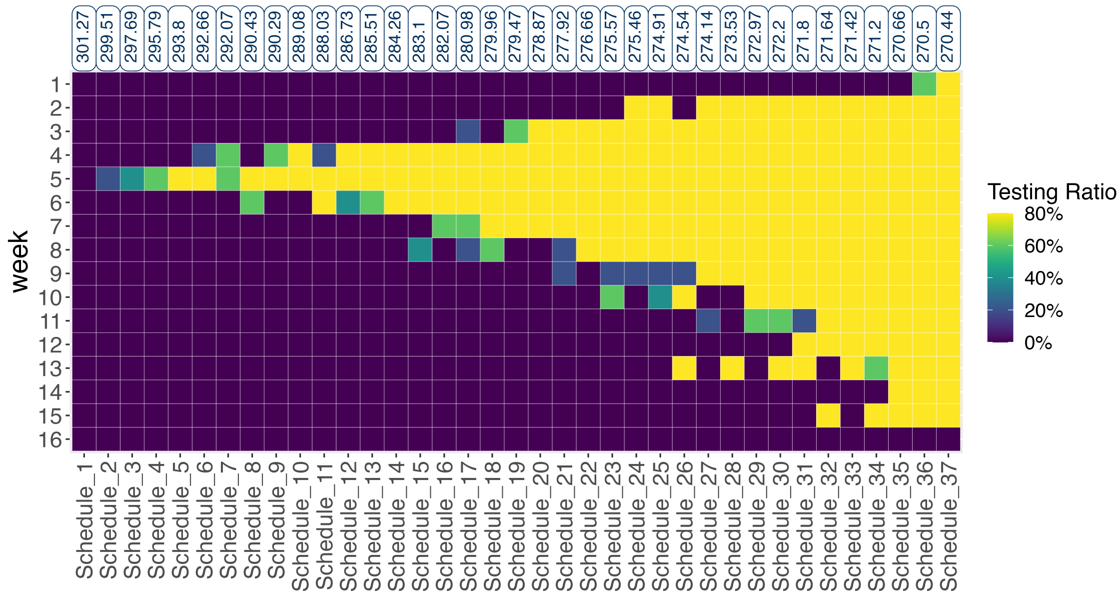

Figure 4: All the testing schedules in the Pareto front.

screening tests within K-12 schools, each with different testing requirements. By calculating descriptive statistics and performing a two-way ANOVA test, we conclude that there are no statistically significant relationships between P_c and P_m and the corresponding performance of NSGA-II, as the Pareto front of the optimal solutions generated behave very similarly. The statistical analysis suggests that NSGA-II is robust in general for solving the problem. However, the pair where P_c and P_m are equal to 0.9 and 0.03125 is slightly preferred due to its overall performance and stability. This suggests that the conventional choice for $P_c = 0.9$ works in our model. In addition, if the number of variables is small, $1/2N$ can be a good candidate for P_m.

Applying the tuned NSGA-II algorithm to our problem, we found that random screening tests are effective in reducing the number of cumulative infections within K-12 schools. However, a lot of tests must be conducted to reduce infections by modest amounts. Layered interventions, such as mask usage alongside screening can further reduce infections, even when mask recommendations are implemented for a short period of time. Higher frequencies of random screening in schools are also beneficial, as it improves the efficiency of tests. The same number of tests can prevent more infections when screening tests are scheduled more frequently in schools. In addition, the week-by-week schedule of testing plays an important part in the efficiency of tests as well. Our analysis shows that it is always optimal to schedule tests more intensively during the early weeks of the semester, such as screening 80% every week for three weeks during weeks 4 through 6, than scattered over the entire semester with low testing ratio. This is particularly important if testing resources are limited.

There are some limitations to the study. Assumptions such as the number newly exposed coming from outside of school each week may not be accurate. We assumed the number of newly exposed was higher during the first month of the semester, and, lower for the rest of the semester. However, there might be seasonal effects associated with outside infections. For instance, more newly exposed individual may return to school after holidays. The impact of these and other parameters (such as self-quarantine) adherence could be tested using sensitivity analysis. For example, when we increased the number of infected individuals at the beginning of the semester, testing started in earlier weeks (results not shown). The method we developed can be easily executed for changes to any parameters and can be tailored to any particular school.

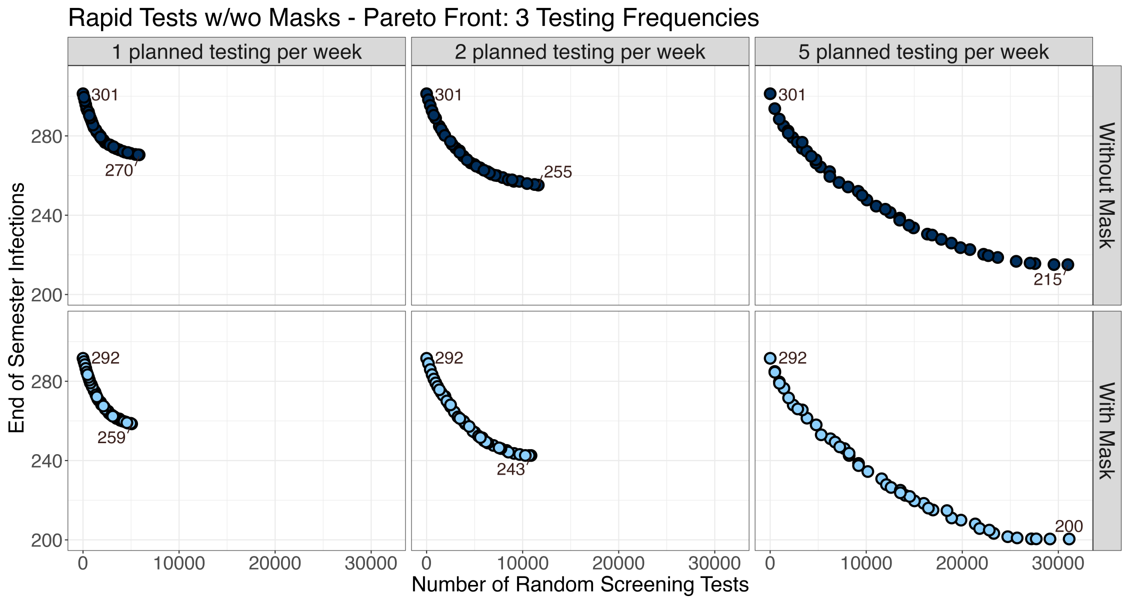

Figure 5: Pareto front under three different planned testing frequencies per week with and without masking implemented.

In addition, this framework can be useful in future pandemics to test intervention strategies as new data and information become available.

This study has demonstrated the effectiveness of masks in reducing infections and saving on the number of screening tests. We believe future work can be done by treating masks as another set of variables, that is adding a decision about which weeks to recommend masking. Then the redesigned optimization problem will be able to provide an optimal schedule for random screening tests and masks at the same time. In addition, we have obtained the results under different testing frequencies, we could let the testing frequency (within a week) be a decision variable, such that a schedule would produce not only which weeks to test but also how many days to test each week and at what intensity. Furthermore, we have only considered rapid antigen tests when simulating the effects of random screening tests in K-12 schools. However, if PCR tests are chosen instead of rapid antigen tests, the results may differ significantly due to the lower false negative rate and longer time to obtain results for PCR tests. If we consider both rapid antigen and PCR tests, the objective may need to change from total number of tests to total cost, as PCR tests are more expensive. Lastly, our study successfully demonstrated the application of NSGA-II, we believe future studies can be conducted by applying different algorithms to the optimization problem and comparing the optimal results and computational efficiency.

ACKNOWLEDGMENTS

This research was supported by Cooperative Agreement NU38OT000297 from the CSTE and the CDC. The research also received partial support from NC State University.

A APPENDIX

Table 4: Simulation parameters. Assumed, unless otherwise indicated.

PARAMETER	ESTIMATES and REFERENCES
School size	500
Day-1 exposed	1
Semester length	107 days
Incoming protection level	30% for K-5 students, 50% for teachers, U.S. Department of Health & Human Services (2022) and calibration
Proportion of asymptomatic	0.4, Ma et al. (2021)
Discount rate of asymptomatic transmission	0.55 Li et al. (2020)
Pre-symptomatic period	2 days, CDC (2022a)
Latent period	1 day, Grant et al. (2022) and Brandal et al. (2021)
Isolation/Recover time	5 days, CDC (2022b)
Isolation compliance rate	100%
Reinfection rate of Omicron	13% , Özüdoğru et al. (2022)
Reinfection period	100 days, Nordström et al. (2022) and calibration
Contact reduction level	0%
Self-quarantine	20%
New “exposed” from community	$\sim DU[3,4]$/week (week = [1,4]) $\sim DU[1,2]$/week (week = [5,16])

REFERENCES

Abdin, A. F., Y. P. Fang, A. Caunhye, D. Alem et al. 2023, 1. “An Optimization Model for Planning Testing and Control Strategies to Limit the Spread of a Pandemic – The Case of COVID-19”. *European Journal of Operational Research* 304(1):308–324.

Adyatama, Arga 2020. “RPubs - Multi-Objective Optimization with NSGA-II https://rpubs.com/Argaadya/moea”.

Anagnostopoulos, K. P., and G. Mamanis. 2010, 7. “A Portfolio Optimization Model with Three Objectives and Discrete Variables”. *Computers & Operations Research* 37(7):1285–1297.

Aragón-Caqueo, D., J. Fernández-Salinas, and D. Laroze. 2020, 10. “Optimization of Group Size in Pool Testing Strategy for SARS-CoV-2: A Simple Mathematical Model”. *Journal of Medical Virology* 92(10):1988–1994.

Asgary, A., M. G. Cojocaru, M. M. Najafabadi, and J. Wu. 2021, 12. “Simulating Preventative Testing of SARS-CoV-2 in Schools: Policy Implications”. *BMC Public Health* 21(1):1–18.

Boutzoukas, A. E., K. O. Zimmerman, M. Inkelas, M. A. Brookhart et al. 2022, 6. “School Masking Policies and Secondary SARS-CoV-2 Transmission”. *Pediatrics* 149(6).

Brandal, L. T., E. MacDonald, L. Veneti, T. Ravlo et al. 2021, 12. “Outbreak Caused by the SARS-CoV-2 Omicron Variant in Norway, November to December 2021”. *Eurosurveillance* 26(50):2101147.

CDC 2022a. “COVID-19 Contact Tracing | CDC”.

CDC 2022b. “Ending Isolation and Precautions for People with COVID-19: Interim Guidance”.

CDC 2023a. “CDC COVID Data Tracker: Hospital Admissions”.

CDC 2023b. “Provisional COVID-19 Deaths: Focus on Ages 0-18 Years | Data | Centers for Disease Control and Prevention”.

Cowger, T. L., E. J. Murray, J. Clarke, M. T. Bassett et al. 2022, 11. “Lifting Universal Masking in Schools — Covid-19 Incidence among Students and Staff”. *New England Journal of Medicine*.

DCD 2023. “NVSS - Provisional Death Counts for COVID-19 - Executive Summary”.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002, 4. “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II”. *IEEE Transactions on Evolutionary Computation* 6(2):182–197.

Dinnes, J., P. Sharma, S. Berhane, S. S. van Wyk et al. 2022, 7. “Rapid, Point-of-care Antigen Tests for Diagnosis of SARS-CoV-2 Infection”. *Cochrane Database of Systematic Reviews* 2022(7).

Feng, J., W. Z. Shen, and C. Xu. 2016, 10. “Multi-Objective Random Search Algorithm for Simultaneously Optimizing Wind Farm Layout and Number of Turbines”. *Journal of Physics: Conference Series* 753(3).

Grant, R., T. Charmet, L. Schaeffer, S. Galmiche et al. 2022, 2. "Impact of SARS-CoV-2 Delta Variant on Incubation, Transmission Settings and Vaccine Effectiveness: Results from a Nationwide Case-control Study in France". *The Lancet Regional Health - Europe* 13.

Jordan, E., D. E. Shin, S. Leekha, and S. Azarm. 2021. "Optimization in the Context of COVID-19 Prediction and Control: A Literature Review". *IEEE Access* 9:130072.

Kekatos, Mary 2022, 12. "Some School Districts Bring Back Mask Mandates to Stem Spread of COVID, Flu and RSV - ABC News https://abcnews.go.com/Health/school-districts-bring-back-mask-mandates-stem-spread/story?id=95975555".

Li, R., S. Pei, B. Chen, Y. Song et al. 2020, 5. "Substantial Undocumented Infection Facilitates the Rapid Dissemination of Novel Coronavirus (SARS-CoV-2)". *Science (New York, N.Y.)* 368(6490):489–493.

Ma, Q., J. Liu, Q. Liu, L. Kang et al. 2021, 12. "Global Percentage of Asymptomatic SARS-CoV-2 Infections Among the Tested Population and Individuals With Confirmed COVID-19 Diagnosis: A Systematic Review and Meta-analysis". *JAMA Network Open* 4(12):e2137257–e2137257.

NCDHHS 2022. "StrongSchoolsNC K-12 COVID-19 Testing Program | NC COVID-19".

Nordström, P., M. Ballin, and A. Nordström. 2022, 2. "Risk of Infection, Hospitalisation, and Death up to 9 Months After A Second Dose of COVID-19 Vaccine: A Retrospective, Total Population Cohort Study in Sweden". *The Lancet* 399(10327):814–823.

Özüdoğru, O., Y. G. Bahçe, and Ö. Acer. 2022. "SARS CoV-2 Reinfection Rate is Higher in the Omicron Variant Than in the Alpha and Delta Variants". *Irish Journal of Medical Science*:1.

Raifman, J., and T. Green. 2022, 11. "Universal Masking Policies in Schools and Mitigating the Inequitable Costs of Covid-19". *New England Journal of Medicine*.

Samsuri, S. F. M., R. Ahmad, M. Z. Zakaria, and M. Z. M. Zain. 2019, 8. "Parameter Tuning for Comparing Multi-Objective Evolutionary Algorithms Applied to System Identification Problems". *2019 IEEE 6th International Conference on Smart Instrumentation, Measurement and Application, ICSIMA 2019*.

The Center for Disability, Equity, and Intersectionality 2022. "State-by-State School Mask Mandates - The National Center for Disability, Equity, and Intersectionality".

Uraki, R., M. Ito, Y. Furusawa, S. Yamayoshi et al. 2022, 12. "Humoral Immune Evasion of the Omicron Subvariants BQ.1.1 and XBB.". *The Lancet. Infectious diseases* 23(1):30–32.

U.S. Department of Health & Human Services 2022. "COVID-19 State Profile Report - North Carolina | HealthData.gov".

Van Pelt, A., H. A. Glick, W. Yang, D. Rubin et al. 2021, 1. "Evaluation of COVID-19 Testing Strategies for Repopulating College and University Campuses: A Decision Tree Analysis". *Journal of Adolescent Health* 68(1):28–34.

Wang, Q., L. Wang, W. Huang, Z. Wang et al. 2019, 5. "Parameterization of NSGA-II for the Optimal Design of Water Distribution Systems". *Water 2019, Vol. 11, Page 971* 11(5):971.

Zhang, Y., M. E. Mayorga, J. Ivy, K. Hassmiller Lich et al. 2022, 7. "Modeling the Impact of Nonpharmaceutical Interventions on COVID-19 Transmission in K-12 Schools". *MDM Policy and Practice* 7(2).

Zimmer, Carl 2023. "Health Experts Warily Eye XBB.1.5, the Latest Omicron Subvariant - The New York Times https://www.nytimes.com/2023/01/07/science/covid-omicron-variants-xbb.html".

AUTHOR BIOGRAPHIES

YIWEI ZHANG is an Operations Research Ph.D. student at North Carolina State University. Her research interests include disease model simulation and optimization for public health-related policymaking. She led the simulation implementation and analysis of this project. Her email address is yzhang85@ncsu.edu. Her website is https://www.or.ncsu.edu/people/yzhang85/.

MARIA E. MAYORGA is a professor of personalized medicine in the Dept. of Industrial and Systems Engineering at North Carolina State University. Her research interests include predictive models in health care and health care operations management and humanitarian systems. Her e-mail address is memayorg@ncsu.edu. Her website is http://mayorga.wordpress.ncsu.edu.

JULIE IVY is a professor in the Dept. of Industrial and Systems Engineering at North Carolina State University. Her research interests are mathematical modeling of stochastic dynamic systems with an emphasis on statistics and decision analysis as applied to health systems. Her email address is jsivy@ncsu.edu and her homepage is https://www.ise.ncsu.edu/people/jsivy/.

JULIE SWANN is department head of the Dept. of Industrial and Systems Engineering at North Carolina State University. Her research interests are using analytics and system approaches to enable health care and supply chains to become more efficient, effective, or equitable. Her email address is jlswann@ncsu.edu, and her website is https://www.ise.ncsu.edu/people/jlswann/.