Response to Request for Proposal: Points of Dispense

Kaila Gilbert, Katie Minor, Muriel Pokol, Brian Rhindress, and Matt Tettelbach
April 26, 2020

Summary of Resources

The model implementation was performed using the Gurobi package in Python. Data analysis and cleaning for the demand scenarios were performed using R software. The potential POD sites were a predetermined list of locations in Allegheny County, and populations per zip code were informed by the United States Census. The GIS Network Analyst tool was used to mine the distances traveled. The first of the simulations were run in Excel, using Allegheny County zip code population counts and densities pulled from ACS 2018 5 year data. Staff needed for the POD sites were determined based on the average percentage of the population that are nurses and physicians. Wages for both nurses and physicians in the cost section were determined by the average wage for both occupations in Pennsylvania. Setup costs and ongoing expenses were based on an average popup cost. The overall budget for the POD sites was taken from the Public Health Preparedness section of the Allegheny County 2020 budget.

Background Literature Resources

1. Simchi-Levi, D., Trichakis, N., & Zhang, P. Y. (2019). Designing Response Supply Chain Against Bioattacks. Operations Research, 67(5), 1246–1268.
2. Lee, E. K., Chen, C.-H., Pietz, F., & Benecke, B. (2009). Modeling and Optimizing the Public-Health Infrastructure for Emergency Response. Interfaces, 39(5), 476–490.

Teamwork Plan

Model Formulation - Full Team
Model Implementation - Katie + Full Team
Data Collection - Brian + Muriel
Scenario Creation - Kaila + Matt
Literature Review - Muriel
Project Coordination - Matt + Katie

This team has previously worked together in various capacities and is well versed in the strategies needed to produce a high quality work product. We created an informal chat group to regularly communicate and exchange ideas. In addition, we met virtually as needed to ensure maximal collaboration and collective refinement of our final work product.

Background

Mass dispensing of testing kits requires the rapid establishment of a network of sites and health facilities that are flexible, scalable, and sustainable for the treatment of the general population (Lee et al). Additionally, each Point of Dispense (POD) site must be capable of serving an appropriate proportion, such as 80%, of the affected local population within a specified short time frame (Simchi-Levi et al, Lee et al). Clearly, for very large-scale dispensing, the sophisticated logistical expertise needed to deal with the complexities of selecting an adequate number of strategically well-placed POD locations, and of designing and staffing each POD, is beyond the capability of any human planner or public-health administrator.[2] Limited dispensing capacity along with the limited availability of trained critical staff, such as public-health professionals, further compounds the inherent complexities (Simchi-Levi et al, Lee et al). These challenges form the crux of the problem statement and the model formulations we present in the following sections.

Problem Statement

We are planning the deployment of Points of Dispense in Allegheny County to distribute essential supplies and medicines to the public during large public health emergencies. There will be two location plans for the PODs. The first will summarize the sensitivity between the cost of opening a POD with the "cost to society", being defined as the total travel distance of each individual seeking treatment and a penalty cost associated with each member of society in demand not receiving treatment above an established quota allowable. The second equation will minimize the maximum travel distance while keeping the other costs calculated the same. Regions are defined as zip codes and distance is measured as travel time via car from one zip code centroid to the POD location. One region (zip code) is assigned to only one POD, but a POD can have multiple regions assigned to it.

Any number of PODs will be capable of serving all in need. Only the subpopulations who are "ill" and are deemed as receiving treatment in each zip code per scenario will travel to the PODs. The PODs will only be used once per person and will not be restocked once it reaches capacity in each scenario. The cost for each POD will be the same across scenarios and regions. It will be determined by the number of nurses and doctors to staff as well as number of testing kits stocked. Each site will be capable of serving a number of individuals related to the nearby populations over the course of one month. The cost per POD will be determined by:

$$POD\ Cost = S*H + F$$

Where S is the number of staff needed, H is the hourly rate of the staff members, and F is the fixed cost of opening and operating the PODs. Testing kits are assumed to be provided and funded by the U.S. Federal Government. Each POD will have a set amount of nurses at \$35 per hour and physicians at \$100 per hour for the month. The cost of each POD will total to \$85,000 which includes staff and testing kits. The total Allegheny County budget for the construction and operation of all POD sites for the one-month time period is \$1,200,000.

Formulation (1)

Objective 1: Minimize the cost of opening PODs

Objective 2: Minimize the cost of travel time (Cost to society)

Objective Function:

α = weight given to cost of setting up new POD

(1- α) = weight given to travel cost + untreated individuals to the POD (cost to society)

$$MIN: \ \alpha \left(\sum_{j=1}^{47} x_j * cc_j \right) + (1 - \alpha) \left[\sum_{s=1}^{S} p_s * \left(\sum_{i=1}^{R} \left[g * r_{is} + \sum_{j=1}^{47} t_{is} * d_{ij} * y_{ijs} \right] \right) \right]$$

Decision Variables:

$$x_j \begin{cases} 1 \text{ if POD } j \text{ is selected} \\ 0 \text{ otherwise} \end{cases} \quad for\ all\ j = 1, \dots, 47$$

$$y_{ijs} \begin{cases} 1 \text{ if region } i \text{ is served by POD } j \text{ in scenario } s \\ 0 \text{ otherwise} \end{cases} \quad for\ all\ POD\ j = 1, \dots, 47, region\ i = 1, \dots, R, scenario\ s = 1, \dots, S$$

t_{is} : $number\ of\ treated\ individuals\ from\ region\ i\ in\ scenario\ s$

r_{is} : $number\ of\ people\ in\ demand, yet\ left\ untreated\ over\ a\ quota\ threshold\ from\ region\ i\ in\ scenario\ s$

Parameters:

N_i : $Number\ of\ residents\ in\ Region\ i$

d_{ij} : $distance\ between\ Region\ i\ and\ POD\ j\ for\ all\ i\ and\ j$

C_j : $Assumed\ capacity\ for\ each\ POD\ j$

cc_j : $Assumed\ cost\ of\ opening\ and\ maintaining\ for\ each\ POD\ j$

p_s : $probability\ of\ scenario\ s$

D_{is} : $Demand\ of\ medical\ attention\ in\ region\ i\ of\ scenario\ s$

q : $quota\ of\ required\ Demand\ met$

g : $cost\ of\ not\ meeting\ quota\ per\ individual\ untreated\ person$

s. t.

1. Each region is assigned to exactly one POD in every scenario.

$$\sum_{j=1}^{47} y_{ijs} = 1 \ \ for\ all\ i = 1, \dots, R, s = 1, \dots, S$$

2. The number treated at a POD cannot exceed capacity of the POD AND no one can be treated to a POD if it is not built.

$$\sum_{i=1}^{R} t_{is} * y_{ijs} \le C_i * x_j \ \ for\ all\ j = 1, \dots, 47, s = 1, \dots, S$$

3. The number treated in a region cannot exceed Demand for the region.

$$t_{is} \le D_{is} \ \ for\ all\ i = 1, \dots, R, s = 1, \dots, S$$

4. Shortage r has to be greater than or equal to the gap between the minimum quota allowed and those treated.

$$r_{is} \ge q * D_{is} - t_{is} \ \ for\ all\ i = 1, \dots, R, s = 1, \dots, S$$

5. Non-negativity.

$$x_j, y_{ijs}, t_{is}, r_{is} \ge 0 \ for\ all\ i = 1, \dots, R, j = 1, \dots, 47, s = 1, \dots, S$$

<u>Formulation (2)</u>

Objective 1: Minimize the cost of opening PODs.

Objective 2: Minimize the maximum distance traveled by any one region.

Same Assumptions and variables as in Formulation (1) but with the following changes:

Objective Function:

$$MIN: \alpha \left(\sum_{j=1}^{47} x_j * cc_j \right) + (1 - \alpha) \left[\sum_{s=1}^{S} p_s * \text{ max\_}dist * \left(\sum_{i=1}^{R} g * r_{is} \right) \right]$$

Additional Constraint: Minimize the largest distance:

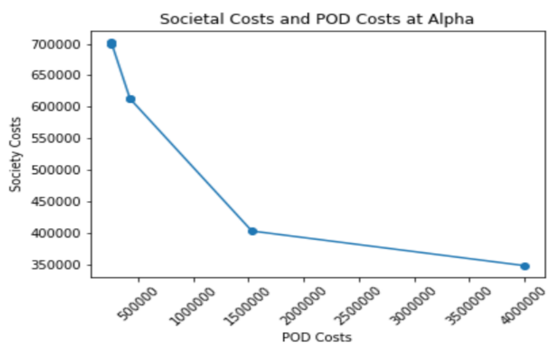
$$max\_dist \geq \sum_{j=1}^{47} d_{ij} * y_{ijs} \quad for\ all\ i = 1, \dots, R, s = 1, \dots, S$$

Data Summary

Capacity	Each site will be stocked with a set amount of tests, determined by sensitivity testing on the basis of the assumption these PODs would be equipped to reasonably serve the population in demand. Capacity is 10,000 test kits per POD.
Staff Needed	Average percentage of the population that are nurses and doctors from here: https://www.truthaboutnursing.org/faq/rn_facts.html#gsc.tab=0 using the assumption that 1 out of every 40 (or 2.5%) nurses and doctors would be assigned to all potential POD sites, so .025(Total Staff) / 47
Nurse Wages	Average hourly wages for PA found here: https://nurse.org/articles/highest-paying-states-for-registered-nurses/
Physician Wages	Average hourly wages for PA found here: https://www.prospectivedoctor.com/what-is-doctors-hourly-wage/
Location	Best approximation for both initial setup costs and one month of expenses of a popup found here: https://www.shopify.com/blog/pop-up-shop
Candidate Sites	The potential POD sites were a predetermined list of locations in Allegheny County provided at the start of this project. There are 47 potential sites.
Distance	The distance traveled was determined by origin-destination pairs from the centroid of each zip code to each POD site location, measured in miles and monetized with the average cost of gas per mile driven, and was mined from the GIS Network Analyst.
Population	Taken per zip code and informed by the United States Census.
Budget	The overall budget for the POD sites was taken from the Public Health Preparedness section of the Allegheny County 2020 budget found here: https://www.alleghenycounty.us/budget-finance/docs/2020-grants-sa.aspx
Demand Scenarios	Allegheny County zip code population counts and densities pulled from ACS 2018 5 year data from the United States Census, and the model assumed the spread of disease for a given zip code was normally distributed with a fixed infection rate mean and standard deviation, adjusting the infection estimate by the normalized population densities to account for increased spread in higher density populations.

Implementation
The above model formulations were implemented using the Gurobi package in Python. For the sake of run time, we allowed our integer programming model tolerance using an optimal gap of 0.1 (m.params.MIPGap=0.1) for Formulation 1 and an optimal gap of 0.2 for Formulation 2. Data related to travel distances, scenarios, and capacity were read-in using CSVs. Vectors were created for alpha values (scaled by .1), scenario probabilities, and sensitivity analyses. Constraints, decision variables, and the objective function were added accordingly using the Gurobi program notations. Linear expressions were utilized when possible.

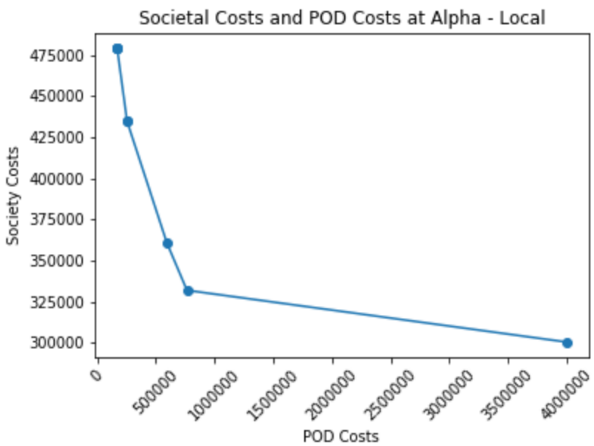
Analysis
Optimal Solution Under Different Alpha Parameters
Weights were used in order to help decision makers balance between costs to society and financial costs of POD setup and operations. When plotting the different optimal solutions values in each case, we saw a constrained distribution of solutions under the current formulation. The calibration of our estimates and outputs could also suggest future investigation. A summary of cost outputs with varying weights can be seen below.

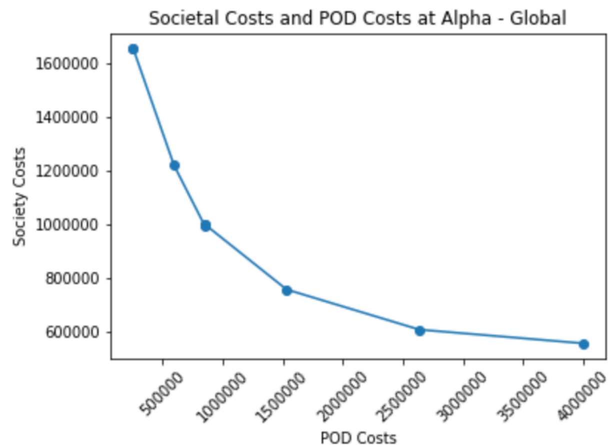


Average Method Formulation 1 vs. Minimax Method Formulation 2

An Average Method (Formulation 1) and a Minimax Method (Formulation 2) were used to analyze the problem. The models were run with a combined scenario set incorporating both global and local scenarios, placing more weight on local outcomes. A minimax approach is generally used when there is great disparity amongst those who are able to access treatment versus those who are not. However, in a comparison between the two models, neither model produced "shortage", meaning individuals in demand not receiving treatment, so we find we do not have disparity in treatment. Also, both tended to result with similar costs to construct PODs. Given these similarities plus the drawback of the minimax method which has a higher optimal tolerance gap, we chose Formulation 1 to draw results from.

Global vs. Local Scenario Results
Using the first formulation we optimized to determine the optimal costs given a local or more global outbreak. We observe that the local outbreak has a much steeper tradeoff between POD and societal costs due to the limited number of zip codes needing service - a small increase in POD expenditure could lead

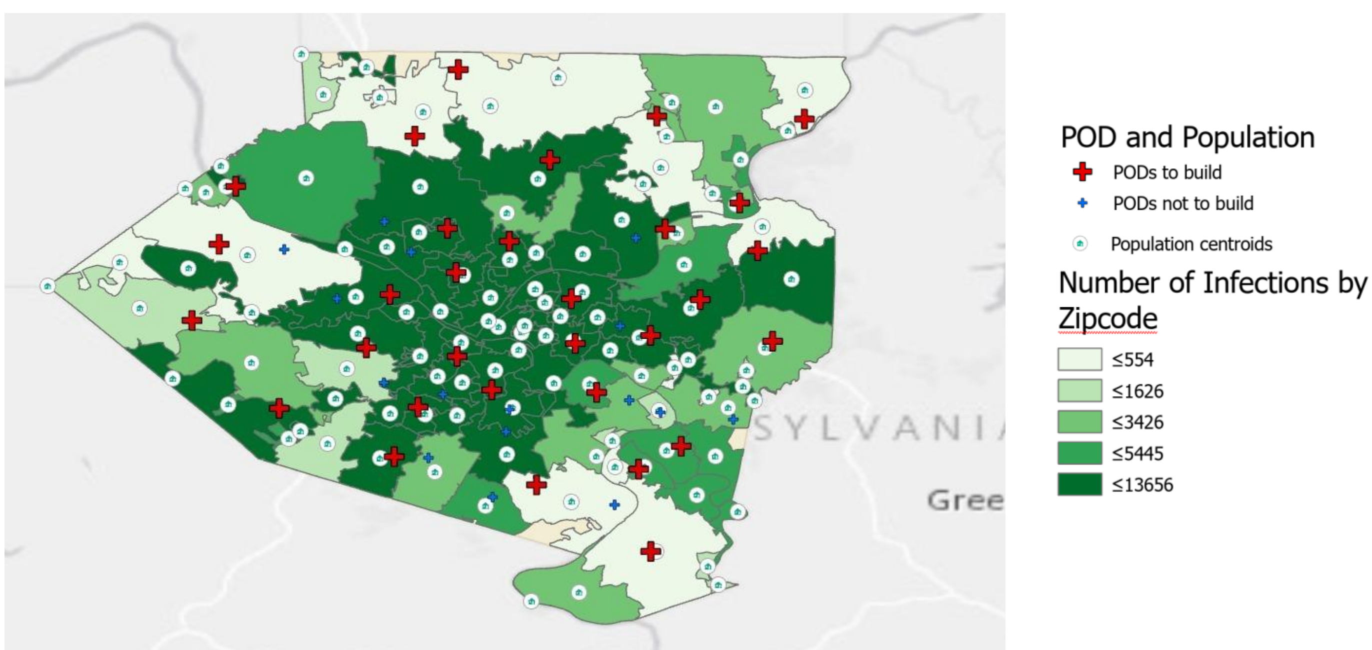




to much lower societal costs. Conversely, the global scenario has a much shallower response surface suggesting that there is a more direct tradeoff between POD and societal costs in the global model.

Recommendations

Our recommendation comes from the weighted goal approach of Formulation 1 mentioned above. Based on the outputs of weights, a 10% weight to the Cost of POD set-up and a 90% weight to Cost of Society emphasizes the importance of minimizing the cost of society while also keeping the POD costs around a reasonable budget for Allegheny County. Allegheny's emergency budget is $1.2 million dollars. The scenario we produced would cost around $1.5 million to set up 18 PODs. A list of the specific 18 PODs is found in the Appendix. An example of a global scenario map of PODs can be seen below.



Allegheny County map of POD facilities in "Global Outbreak Scenario 0" with a dark green choropleth layer representing severity of outbreak, built PODs (red crosses), non-built PODs (blue crosses), and zip code centroids (houses in circles).

Extended Study

The team levied a number of assumptions that simplified both the nature of the disease outbreak as well as the problem formulation. The team recommends further analysis into a locally informed implementation of this model that takes into consideration available resources (at the zip code and county-levels), relevant capacities of PODs, different types of medical and biological crises, and distance costs that can consider alternative forms of transportation. While this model is a proof of concept that can show tradeoffs between two objectives, more exploration is certainly needed to better quantify these costs and resulting solutions. A code and csv containing SIR model generated demand values and parameters is included with this report.

Appendix

ID	POD Name	POD Address
1	AVONWORTH HIGH SCHOOL	304 JOSEPHS LN, PITTSBURGH, PA 15237
2	BRASHEAR HIGH SCHOOL	590 CRANE AVE, PITTSBURGH, PA 15216
6	CARLYNTON SENIOR HIGH	435 KINGS HWY, CARNEGIE, PA 15106
9	CARRICK HIGH SCHOOL	125 PARKFIELD ST, PITTSBURGH, PA 15210
15	HAMPTON HIGH SCHOOL	2929 MCCULLY RD, ALLISON PARK, PA 15101
16	HIGHLANDS HIGH SCHOOL	1500 PACIFIC AVE, NATRONA HEIGHTS, PA 15065
18	GATEWAY HIGH SCHOOL	3000 GATEWAY CAMPUS BLVD, MONROEVILLE, PA 15146
19	ELIZABETH FORWARD SR HIGH	1000 WEIGLES HILL RD, ELIZABETH, PA 15037
20	MCKEESPORT AREA HIGH SCHOOL	1960 EDEN PARK BLVD, MCKEESPORT, PA 15132
21	MT. LEBANON HIGH SCHOOL	155 COCHRAN RD, PITTSBURGH, PA 15228
23	NORTH ALLEGHENY HIGH SCHOOL	10375 PERRY HWY, WEXFORD, PA 15090
24	NORTHGATE JR SR HIGH	589 UNION AVE, PITTSBURGH, PA 15202
28	PERRY HIGH SCHOOL	3875 PERRYSVILLE AVE, PITTSBURGH, PA 15214
29	PITTSBURGH ALLDERDICE HIGH	2409 SHADY AVE, PITTSBURGH, PA 15217
32	PITTSBURGH OBAMA 6-12	515 N HIGHLAND AVE, PITTSBURGH, PA 15206
34	RIVERVIEW SENIOR HIGH SCHOOL	100 HULTON RD, OAKMONT, PA 15139
41	WEST ALLEGHENY HIGH SCHOOL	205 W ALLEGHENY RD, IMPERIAL, PA 15126
44	THOMAS JEFFERSON HIGH SCHOOL	310 OLD CLAIRTON RD, CLAIRTON, PA 15025
47	WOODLAND HILLS HIGH SCHOOL	2550 GREENSBURG PIKE, PITTSBURGH, PA 15221

SIR Model

To understand how the model would respond to simulated outbreaks , the team used a SIR model to track trajectories over a thirty-day span and project how many would be infected under different quarantine and vaccine scenarios. The model assumes an airplane enters Allegheny County with half of the individuals infected. These individuals are randomly assigned to zip codes using probability distributions based on zip code demographics. Following this random initialization, each infected zip code’s disease trajectory is run through the SIR model using its local characteristics. Demand is then modeled as the number infected plus total of individuals recovering by the end of the selected time period. General assumptions are that activity within zip codes are contained, no individuals are naturally immune, and all zip codes share fixed parameters for interaction patterns, vaccine effectiveness, and susceptibility. More details can be found in the associataed R markdown file, although there was insufficient time to fully implement both sets of scenarios.