Outreach Strategies for Vaccine Distribution: A Two-Period Robust Approach

Yuwen Yang, Jayant Rajgopal

Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA

Acknowledgment

This work was partially supported by the National Science Foundation via Award No. CMII-1536430.

1 Corresponding author; e-mail: rajgopal@pitt.edu; Tel. +1 412 624 9840; ORCID 0000-0001-7730-8749
Outreach Strategies for Vaccine Distribution: A Two-Period Robust Approach

Abstract

Vaccination has been proven to be the most effective method to prevent infectious diseases. However, in many low and middle-income countries with geographically dispersed and nomadic populations, last-mile vaccine delivery can be extremely complex. Because newborns in remote locations within these countries often do not have direct access to clinics and hospitals, they face significant risk from diseases and infections. An approach known as outreach is typically utilized to raise immunization rates in these situations. A set of these remote population centers is chosen, and over an appropriate time horizon, teams of clinicians and support personnel are sent from a depot to set up mobile clinics at these locations to vaccinate people there and in the immediate surrounding area. In this paper, we model the problem of optimally designing outreach efforts as a mixed integer program that is a combination of a set covering problem and a vehicle routing problem. In addition, because elements relevant to outreach (such as populations and road conditions) are often unstable and unpredictable, we incorporate uncertainty to study the robustness of the worst-case solutions and the related issue of the value of information.

Keywords: Vaccines; Distribution; Vehicle routing; Mixed integer programming; Robust optimization

1. Introduction

As a biological preparation against infectious disease, vaccines have averted 2 to 3 million deaths annually [1], and coverage rates have improved significantly over the years under the guidance of
the World Health Organization’s Expanded Programme on Immunization (WHO-EPI) and the Global Alliance for Vaccines and Immunization (Gavi) [2, 3]. However, in many of the poorest countries, getting childhood vaccines delivered to their final destinations can be an extremely complex process. Although many low and middle-income countries (LMICs) can often obtain vaccines at low cost, operating a vaccine distribution system can be a challenge. Many vaccines require a narrow temperature range of between 2 and 8°C during storage and transportation, which in turn brings with it high distribution and storage costs. In addition to the challenge of planning for storage devices and transportation capabilities to distribute vaccines throughout the country, geographically dispersed or nomadic populations also present a major challenge. As a result, in many countries significant portions of the population have no direct access to health clinics.

Inadequate infrastructure and geographic barriers such as poor road conditions or limited access to transportation can further compound this problem. For example, in Niger, around 90% of the roads are not paved [4]. A recent study published in The Lancet Global Health estimated that across 48 sub-Saharan countries, 28.2% of women of child-bearing age are more than 2 hours travel time (combined walking and motorized) away from the nearest hospital [5]. The study also found wide variations with the percentage ranging from under 25% in South Sudan to over 90% in several countries including Nigeria, Kenya, Swaziland and Burundi. Another recent study in Uganda [6] concluded that difficulty in access to immunization centers due to poor road terrain has a significant effect and results in low immunization coverage. Thus, people living in remote locations in LMICs often face significant difficulty in obtaining routine vaccinations, and the WHO estimates that almost 20 million infants worldwide are at high risk from vaccine-preventable diseases such as polio, measles, yellow fever and tuberculosis [7].
To supplement the fixed vaccine distribution network, an approach known as outreach is typically utilized to raise immunization rates, especially in remote areas where direct access to clinic services is limited or unavailable. Clinicians and support personnel are sent from an existing clinic location to render services at one or more of these remote population locations. While the exact terminology varies from country to country we will refer to the existing location as a depot and the remote location as a mobile clinic. People at the location and other locations that are within a reasonable distance from the mobile clinic come there to get vaccinated. Note that this service is distinct from a campaign (a one-time attempt) in that outreach is periodic and repeated at regular time intervals. This interval might range from 1 month to 6 months in different countries.

Compared to a fixed clinic, mobile clinics can offer more flexibility and viability when treating vulnerable and isolated populations [8] and avoid unnecessary fixed facility, inventory, and labor cost [9]. Furthermore, outreach is proven to dramatically raise the overall immunization rates in resource-deprived countries that suffer from extremely low coverage rates. An early study in Kenya estimated that outreach increased the coverage rate in the lowest density zone in Kenya from 25% to 57% and from 54% to 82% in the area with greatest population density [10]. With the support of the WHO, outreach activities encompassing 1,982 mobile clinics and 5,964 personnel were able to cover 80% of targeted infants in September 2015 in Yemen; 290,498 children were vaccinated by these actions [8].

While there has been some relatively recent work on the network design phase of the WHO vaccine distribution chains [11, 12, 13, 14], it is somewhat surprising that outreach has not received more attention in the academic literature and that there are almost no quantitative models available to help decision makers create an optimal outreach strategy. Lim et al. were the first to present quantitative models to determine optimal outreach locations and policies to maximize coverage.
rate. The authors contrasted various coverage models using data derived from the state of Bihar in India [15]. In more recent work, Mofrad has proposed a mixed integer programming model to obtain an optimal mix of fixed and outreach vaccination services under demand uncertainty [16].

In this paper we propose a general model for LMICs to build on these early studies. Section 2 provides some background and our assumptions. We then provide an MIP model formulation in Section 3. An extension of the model with uncertainty considerations is presented in Section 4. We introduce the notion of value of information in Section 5 and present numerical results in Section 6. Finally, we discuss the results and summarize our work in Section 7.

2. Problem Description and Literature Review

While outreach has been proven to be effective at increasing vaccination rates in resource-deprived regions of the world, there is no standard structure or process that every country follows. A typical process might be one where a medical team departs from an existing district center or clinic in a van or truck, carrying supplies and vaccines in cold boxes. The team then sets up at one or more mobile clinic location(s) and vaccinates the area’s residents as well as residents from nearby areas. If multiple locations are visited, the team might go to each location sequentially and return to the original depot at the end of the day. However, each country has its own outreach policy and criteria to conduct outreach. For example, an outreach team might consist of clinicians and workers who come from multiple locations. The vaccines might be delivered via a truck by a separate logistical team and stored in refrigerators at mobile clinics before the clinical team arrives there. In some countries, it might be possible for the team to stay in a mobile location overnight. In general, unlike with the operation of fixed clinics and the associated distribution system, there are no clear standards on how outreach should be done.
Despite significant variations in economy, geography, demography, etc., and thus in how outreach is done across all these countries, this paper aims to provide a relatively rigorous process for outreach trips across all countries to meet the WHO’s goal of providing the entire targeted population with the opportunity to be vaccinated. We present a mixed integer programming (MIP) formulation to optimize the outreach strategies.

To retain a tractable MIP model while accounting for the various associated complexities and diversity as best we can, we include three sets of decisions into our consideration. The first is choosing the locations of mobile clinics for outreach as a subset of the existing targeted population centers. The second set of considerations is how to assign population centers to mobile clinics. Note that a population center can be assigned to a mobile clinic only if it is closer than the maximum coverage distance (MCD) to that mobile clinic. The MCD can be defined as the maximum distance that people must travel to get vaccinated and is determined by the planners in the country. A mobile clinic could have the capability to serve multiple population centers, and each population center is assigned to a specific mobile clinic. Third, we determine an optimal set of vehicle trips that ensure that all mobile clinic locations are visited once within some suitable planning horizon (typically, 3 months or 6 months). Each trip would carry the required clinical and support personnel along with the required amount of vaccine for the location(s) served by the trip.

Note that the vaccine regimens are not identical across countries, and based on the demand that is expected at population centers and the vaccine vial volumes, we can estimate the total volume associated with expected demand at each population center at each outreach session. Within the planning horizon, multiple vehicle trips can be undertaken but each vehicle trip must depart from a fixed depot and return to that depot after it visits one or multiple mobile clinic locations. The
vehicles utilized in outreach trips are typically trucks or vans with several coolers or cold boxes and are thus capacititated in terms of how much vaccine can be carried.

To model a realistic process for outreach, we consider time windows on vehicle trips. There is a maximum trip duration (MTD) for each vehicle trip, e.g., 8 to 12 hours if all personnel need to return to the depot on the same day. In the case that they could stay overnight at a mobile clinic location, the MTD could possibly be longer. We also consider service time. This includes time used to set up the mobile clinic and time allocated to vaccinate targeted population members who come to the clinic. Different clinics do not always have identical service time; a clinic at a location that serves a larger population is likely to have longer service times. The service time at the originating depot can be set at zero or to the actual time required to load vaccines and prepare the team on the day of the trip. The travel times between the depot and clinics are obtained by dividing the corresponding distances by the average vehicle speed (e.g., 25 km/h).

We consider two components of cost in our objective function. The first is the direct cost associated with running a mobile clinic at a remote location. This cost includes the setup at the outreach site, the cost of renting or obtaining space, any labor costs for vaccination operations onsite, potential storage and energy consumptions cost, and any other local cost. The second cost component is the trip-related cost that is assumed to be proportional to the duration of the trip. This might include fuel costs, vehicle depreciation, hourly wages/allowances paid to the team and driver, vehicle rental costs, etc. We assume this results in an average cost per hour that is used to compute the cost of the trip based on its planned duration. The total cost is thus determined by the locations of the mobile clinics and the routes taken by the vehicles on their trips. In summary:

1) Our objective is to minimize the sum of direct mobile clinic costs and trip costs.

2) Mobile clinics for outreach are chosen from a set of existing targeted population centers.
3) A population center is said to be covered by a mobile clinic if it is within the specified MCD of the clinic.

4) A population center can be assigned to a mobile clinic only if it is covered by that mobile clinic, and each population center is assigned to one mobile clinic so that the entire population assigned to the depot has the opportunity to be vaccinated.

5) Multiple outreach trips are made within the planning horizon, and every mobile clinic must be visited once within the planning horizon by an outreach trip.

6) In each outreach trip, the vehicle departs from the depot, visits one or more mobile clinic locations, and returns to the depot within the MTD.

7) We consider time windows on outreach trips and assume that a trip cannot be longer than some given maximum duration. There is a service time at each mobile clinic and a travel time between locations.

8) The vehicle is capacitated, and we assume the capacity is more than that what is required at any single population center.

The MCD in assumption 3 is assumed to be set to a value that is acceptable in the country being considered. Assumption 4 captures the WHO policy of ensuring that every child has the opportunity to be vaccinated. Assumptions 5, 6 and 7 are based on the most common practice, and assumption 8 is required to ensure feasibility.

The proposed problem under these assumptions can be viewed as a combination of a set covering problem (SCP) and a vehicle routing problem with time windows (VRPTW): the process of choosing mobile clinics and assigning population centers to each can be viewed as an SCP while the routes to visit these mobile clinics can be viewed as a VRPTW. A typical SCP in this context
would choose the optimal facility locations with the objective of minimizing cost or maximizing the total demand covered [17, 18]. This is a well-studied problem in the operations research community and has been widely applied in the healthcare area [19]. The VRPTW is a variation of the Vehicle Routing Problem (VRP), which due to its wide application and importance in distribution networks, has also been widely studied by researchers. The goal of VRP is to obtain an optimal vehicle trip strategy to serve a set of customers. However, due to its complexity, exact algorithms such as branch-and-cut and branch-and-price usually have a size limit of 50 to 100 nodes; the problem is thus often solved by approximation algorithms and heuristics to find high quality solutions [20, 21]. VRPTW has the additional complication that customers need to be served within predefined time windows and within total trip time durations at minimum total cost, and belongs to the class of NP-hard problems [22]. These problems are often solved by heuristics such as genetic algorithms [23, 24], Tabu search [25], evolutionary algorithms [26], large neighborhood search [27], guided local search [28] and hybrid metaheuristics [29,30]. An iterative route construction and improvement algorithm [31] has been suggested for solving VRPTW with “soft” time windows (a relaxation in the length of the time windows).

3. Model Formulation

Parameters

\( n \): Total number of targeted population centers

\( i \): Index of locations. \( 1 \leq i \leq n \) if \( i \) is a targeted population center; \( i = 0, n+1 \) if \( i \) is depot

\( k \): Index of outreach trips

\( b_i \): Volume of vaccine demanded at population center \( i \) over the planning horizon
$f_i$: Fixed cost of running a mobile clinic at population center $i$

c: Average transportation cost per hour

d$_{ij}$: Distance between location $i$ and location $j$ (with $d_{ii} = 0$)

$D$: Maximal coverage distance (MCD)

$a_{ij} \in \{0,1\}$: 1 if location $i$ is within a distance $D$ from location $j$, 0 otherwise

$K$: Maximum number of outreach trips that can be made within the planning horizon

$t_{ij}$: Travel time from location $i$ to location $j$

$s_i$: Service time at location $i$

$r$: Maximum trip duration (MTD)

$p$: Capacity of vehicle

**Variables**

$X_{ij} \in \{0,1\}$: 1 if population center $j$ is assigned to mobile clinic at location $i$, 0 otherwise

$Y_i \in \{0,1\}$: 1 if there is a mobile clinic at location $i$, 0 otherwise

$Z_{ijk} \in \{0,1\}$: 1 if location $i$ is followed by location $j$ in outreach trip $k$, $k \leq K$

$U_{ik}$: Accumulated vaccine volume already distributed by outreach trip $k$ when arriving at mobile clinic location $i$, $k \leq K$

$W_i$: Total volume of vaccine sent to mobile clinic at location $i$

**MIP-1:**

$$\text{Min} \quad \sum_{1 \leq i \leq n} f_i Y_i + \sum_k \sum_i \sum_j c_{ij} Z_{ijk}$$

subject to
\( X_{ij} \leq a_{ij} \quad \forall i, j \) (2)

\( X_{ij} \leq Y_i \quad 0 \leq i \leq n, 1 \leq j \leq n \) (3)

\[ \sum_{i=1}^{n} X_{ij} = 1 \quad 1 \leq j \leq n \] (4)

\[ W_i = \sum_{j} b_j X_{ij} \quad i \leq n \] (5)

\[ \sum_{j} Z_{0jk} = 1 \quad \forall k \] (6)

\[ \sum_{j} Z_{j0k} = 0 \quad \forall k \] (7)

\[ \sum_{i} Z_{i(n+1)k} = 1 \quad \forall k \] (8)

\[ \sum_{i} Z_{(n+1)ik} = 0 \quad \forall k \] (9)

\[ \sum_{j} Z_{ijk} = \sum_{j} Z_{ji}k \quad \forall k, 1 \leq i \leq n \] (10)

\[ \sum_{j} \sum_{k} Z_{ijk} = Y_i \quad 1 \leq i \leq n \] (11)

\[ U_{ik} - U_{jk} + pZ_{ijk} \leq p - W_j \quad \forall i, j, k \] (12)

\[ W_i \leq U_{ik} \leq p \quad \forall i, k \] (13)

\[ \sum_{i} \sum_{j} (t_{ij} + s_i)Z_{ijk} \leq r \quad \forall k \] (14)

\[ \sum_{i} \sum_{j} \sum_{k} Z_{ijk} \geq \sum_{i} \sum_{j} Z_{ijk} \quad k \geq 2 \] (15)

\[ Z_{ikh} = 0 \quad \forall i, k \] (16)

\( X_{ij} \in \{0, 1\} \quad \forall i, j \) (17)

\( Y_i \in \{0, 1\} \quad \forall i \) (18)

\( Z_{ijk} \in \{0, 1\} \quad \forall i, j, k \) (19)

\( U_{ik} \geq 0 \quad \forall i, k \) (20)

\( W_i \geq 0 \quad \forall i \) (21)
The objective function (1) minimizes the overall cost that has two components: clinic operation costs and outreach trip transportation costs. Constraints (2) ensure that a mobile clinic can only serve population centers within the MCD of the clinic. Constraints (3) ensure that a location can serve other locations only if a mobile clinic is scheduled there. Constraints (4) ensure that each population center is assigned to a mobile clinic. Constraints (5) compute the total vaccine volume handled at a mobile clinic based on the population that the clinic serves. These four sets of constraints define a typical facility location problem.

The next set of constraints relate to the vehicle routing problem. Note that node 0 denotes the origin and node (n+1) is the final node at the end of a trip; both represent the depot. Constraints (6) and (7) imply that each vehicle trip departs from the depot (0) exactly once, while Constraints (8) and (9) imply that each vehicle trip enters back into the depot (n+1) exactly once. Constraints (10) ensure that the flow that enters and departs any population center $i$ is balanced in each outreach trip $k$. Constraints (6) – (10) thus ensure that every vehicle trip is indeed a (0)-(n+1) path.

Constraints (11) state that exactly one vehicle enters and departs each population center during a time horizon if there is a mobile clinic at this location (i.e., $Y_i = 1$). Constraints (12) are the vehicle-specific version of MTZ subtour elimination constraints introduced by Miller, Tucker, and Zemlin [32]. Note that for a particular vehicle route $k$ in which $j$ follows $i$, $Z_{ijk} = 1$ implies $U_{jk} \geq U_{ik} + W_j > U_{ik}$. Suppose there exists a subtour $(i, j, \ldots)$, with $i \neq 0, n$. Then, $U_{jk} > U_{ik} > U_{jk}$ will lead to a contradiction. Constraints (13) ensure that a vehicle carries enough vaccine for each mobile clinic but does not exceed its capacity. Note that if location $i$ is not a part of trip $k$ the values of $U_{ik}$ are irrelevant to the problem as long as they satisfy (13).

Constraints (14) state that the total travel time and service time of a route cannot be larger than the MTD. Constraints (15) are added to avoid degeneracy by ensuring that route $k$ is never
utilized if route $k-1$ is not utilized; with Constraints (15), we reduce the search space by making sure that vehicle routes are chosen in a sequence of 1, 2, 3, etc. In addition, it ensures that vehicle trips with more stops will have a lower index value. Constraints (16) – (21) are self-explanatory.

4. Two-period Stochastic Model

In Sections 2 and 3 we introduced a model that assumed all parameters are constant and deterministic, and the model is solved once. However, conditions in many targeted outreach locations are not always stable and predictable. It can often be difficult to obtain accurate estimates of all problem parameters ahead of time, and these might change as we get closer to the implementation of the outreach trips. For example, because demand is a function of population and birth rate, it can be more accurate to think of it as being stochastic, as both the population and the birth rate within a location could vary from year to year or even within a year. Similarly, in Assumption (7) we estimate the travel time from $i$ to $j$ as a constant based on the distance and the average vehicle speed. However, traffic and road conditions in the targeted zones can be unstable, so that this assumption might also need to be reconsidered. With an extreme event such as a flood or a landslide, a road might even be blocked. Conversely, improvements to infrastructure might actually reduce travel times. Therefore, it would in general be suboptimal to determine a fully fixed strategy ahead of time and simply repeat it in every successive time window.

On the other hand, it can also be problematic if we update all parameters and obtain completely revised plans for each time window. Recall that our problem is to minimize costs while providing the opportunity for 100% coverage; however, for a variety of reasons, in practice not every patient will show up at a clinic. A major goal of the WHO is to make access to vaccinations as easy as possible so as to minimize the number of these lost opportunities, and from this
viewpoint, it is desirable to have a stable set of mobile clinic locations and for the populations assigned to each to be aware of when and where clinics will be conducted a regular basis (e.g., the second Tuesday of every month; or the first Monday in January, April, July and October; or March 15 and September 15). We draw a compromise here by fixing locations but allowing for flexibility in timings. It is undesirable to move mobile clinic locations because it is disruptive and confusing for the populace to be directed to a different location each time for vaccination services. In contrast, it is relatively easy to inform people of a change in the timing of a clinic (because of a change in how we do the vehicle routing), especially if it is only for some clinics and the new times are not too different from those in the previous session.

Suppose we consider our problem using a two-period stochastic model. The two main uncertainties we consider during each time window are (a) with respect to the population (and hence the volume of vaccines required) at each location, and (b) with respect to the travel times between locations \( i \) and \( j \). Suppose that the volume of vaccines demanded at location \( i \) within each time window is stochastic and represented by the random variable \( \tilde{b}_i \), but we can constrain it to lie within some range \((b_i, \bar{b}_i)\). Similarly, the travel time between \( i \) and \( j \) is assumed to be stochastic and given by \( \tilde{t}_{ij} \in (\underline{t}_{ij}, \bar{t}_{ij}) \). This range might in general, be quite wide in order to account for inherent uncertainties and the upper bounds would reflect worst case scenarios. We can then utilize MIP-1 but incorporate this information to now minimize either the expected cost or the maximum cost possible. We choose the latter option as it is more desirable in LMICs if the goal is to follow the WHO guidelines of reaching every child. As we will see, this also has the advantage of not requiring a characterization of the distributions associated with the stochastic variables. The
solution to this yields the optimal mobile clinic locations, the assignment of population centers to these locations, and the associated routes for outreach trips for use within the first time window.

At the end of the first time window we review our estimates of the demand and travel time parameters and update these based on the most current information. For example, estimates of the population in some locations might have changed because of seasonal migrations or because of updated information from public health or other sources. Similarly, we might perhaps know that because of some natural catastrophe certain roads will be unavailable over the next time window, or that driving times along certain routes will be longer or shorter because of changes in the season or changes in road conditions.

In the second period problem, we assume that the locations of mobile clinics and their population center assignments are fixed at the values obtained earlier, but we use updated parameters \((\bar{b}_i', \bar{b}_i'')\) and \((t_{ij}', t_{ij}'')\) to obtain revised routes for outreach trips in the second time window. Typically, the range of parameter values resulting from these updated estimates will be tighter than the ones with the period 1 problem previously solved because of additional information that might now be available.

We illustrate this via a simple example shown in Figure 1. Suppose that we are developing the outreach strategy for an area containing a fixed clinic (which serves as the depot) and 15 population centers that must be served by it via outreach, with a visit every three months. We use our initial estimates of the demand and travel time to obtain the mobile clinic locations, along with the population assignments to each. We also obtain a set of outreach trips with routes as shown in Figure 1, where arrows represent vehicle routes and dotted lines represent assignment of outside population centers to a clinic: we have 8 outreach sessions with mobile clinics at locations 2, 5, 6,
9, 10, 12, 14, 15. While each mobile clinic serves the population at its location, the clinic at location 2 also serves population centers 3 and 4, which are within the MCD of location 2. Similarly location 6 also serves 7; 9 also serves 8; 12 also serves 11 and 13; and people at location 1 are served by the fixed clinic (location 0). We have three separate outreach trips: trip 1 visits and holds outreach clinics at locations 2, 5 and 6 before returning to the depot; trip 2 does the same with locations 10 and 9, and trip 3 with locations 12, 14 and 15.

![Figure 1: Initial Solution](image)

This plan is implemented in the first quarter. At the end of the first quarter we get updated information and learn that the travel time along each edge will be a lot shorter because we have a new better vehicle now, but that the roads connecting locations 2 and 5, as well as 0 and 6 will be closed because of major repairs. Without changing the locations of our mobile clinics, if possible we would like to obtain a better set of outreach trips to cover these same locations during the next
planning period (quarter 2) based on the updated information. This results in the strategy displayed in Figure 2. We still have the same eight clinics but now only have two outreach trips (0-12-14-15-2-0 and 0-5-6-9-10-0).

![Figure 2: Updated Solution](image)

This process can then be repeated for subsequent time windows, and as we obtain new information we can obtain updated solutions for the outreach trips each time. In summary, we solve the following two-period stochastic mixed integer programming models (TS-MIP):

**TS-MIP-1:**

\[
Z_1 = \min_{X,Y,Z,U,V,W} \max \left( \sum_{1 \leq i \leq n} f_i Y_i + \sum_i \sum_j \sum_k c_{ij} Z_{ijk} \right)
\]

subject to

Constraints (2) – (4)
\[ \sum_{i} W_i = \sum_{j} \tilde{b}_j X_{ij} \quad i \leq n \]  

Constraints (6) – (13)

\[ \sum_{i} \sum_{j} (\tilde{t}_{ij} + s_i)Z_{ijk} \leq r \quad \forall k \]  

Constraints (15) – (21)

\[ \tilde{t}_{ij} \in (t_{ij}', \tilde{t}_{ij}'); \tilde{b}_i \in (\tilde{b}_i, \tilde{b}_i) \]

Let \( Y_i^* \) be the optimal value of \( Y_i \) in the solution to TS-MIP-1. Then we have

**TS-MIP-2:**

\[ Z_2 = \left( \sum_{1 \leq i \leq n} f_i Y_i^* + \right) Min_{Z,u,v,w} Max_{\sum_{i} \sum_{j} \sum_{k} c_{ij} Z_{ijk}} \]

subject to

Constraints (2) – (4), with \( Y_i = Y_i^* \) in (3)

\[ \sum_{i} W_i = \sum_{j} \tilde{b}_j X_{ij} \quad i \leq n \]  

Constraints (6) – (13), with \( Y_i = Y_i^* \) in (11)

\[ \sum_{i} \sum_{j} (\tilde{t}_{ij} + s_i)Z_{ijk} \leq r \quad \forall k \]  

Constraints (15) – (17), (19) – (21)

\[ \tilde{t}_{ij} \in (t_{ij}', \tilde{t}_{ij}'); \tilde{b}_i \in (\tilde{b}_i, \tilde{b}_i) \]

Note that TS-MIP-2 is solved with the clinic locations fixed, and optimizes deliveries over an updated range of demands \( \tilde{b}_i \) and travel times \( \tilde{t}_{ij} \).
**Proposition 1**: Assuming feasibility, TS-MIP-1 and TS-MIP-2 are equivalent respectively to (a) solving TS-MIP-1 with $\tilde{t}_{ij} = \overline{t}_{ij}$ in (22), (24) and $\overline{b}_j = \overline{b}_j$ in (23); and (b) solving TS-MIP-2 with $\tilde{t}_{ij} = \overline{t}'_{ij}$ in (25), (27) and $\overline{b}_j = \overline{b}'_j$ in (26).

**Proof**: First, note that from (23) or (26), as the value of $\overline{b}_j$ increases, so does the value of $W_i$. This in turn reduces the size of the feasible regions for TS-MIP-1 and TS-MIP-2 by tightening the constraints defined by (12) and (13). Similarly, an increase in $\tilde{t}_{ij}$ tightens the constraints defined by (24) or (27) while also increasing the cost coefficient for $Z_{ijk}$ in the objective. So with these changes, assuming feasibility, the objective function can only increase from its current value (or at best, stay the same). Its maximum value is thus obtained when each $\overline{b}_j$ and $\tilde{t}_{ij}$ is at its largest possible value.

The above result is intuitive: when the population (demand) increases, it is possible that limitations arising from the vehicle capacity might increase the number of trips required to cover all locations, and when travel times along an arc $i$-$j$ increase, the total travel costs rise; it is also possible that the length of a trip might exceed the trip MTD $(r)$, again causing an increase in the number of trips. Proposition 1 states that if we are conservative and plan for the worst-case scenario with respect to the period 2, then this corresponds to when travel times and populations are as large as they could get. We refer to the solutions for these worst-case scenarios as **robust** solutions.

**5. Robustness and the Value of Information**

In Section 4 we introduced a two-period procedure to address the unstable outreach environment that is typical in practice. In this section we compare, discuss and interpret the costs associated with the robust solutions to the period 1 and period 2 problems.
We can interpret $Z_1$ as the optimal cost associated with the conservative strategy at the beginning of the first period that addresses the worst-case scenario. The optimal value of TS-MIP-2 given by $Z_2$ is also for a conservative strategy but with an updated worst case scenario and with clinic locations fixed based upon the optimal solution to TS-MIP-1 for the first period. Any difference between $Z_1$ and $Z_2$ is a result of possibly updated outreach trips with better vehicle routes. While $Z_2$ could in general be larger or smaller than $Z_1$, if the updated upper bounds are the same or smaller than before, then as the following corollary states, $Z_2$ will be smaller.

**Corollary 1:** If $\overline{b_i} \leq \overline{b_i}'$ and $\overline{t_{ij}} \leq \overline{t_{ij}}'$, then $Z_2 \leq Z_1$.

**Proof:** In proving Proposition 1 we saw that as the values of $\overline{b_i}$ and $\overline{t_{ij}}$ increase, the feasible regions for both problems shrink, and when they decrease the region expands. Therefore, $Z_1$ and $Z_2$ are monotone non-decreasing in both $\overline{b_i}$ and $\overline{t_{ij}}$. Further, TS-MIP-2 has the same locations as the optimal locations in TS-MIP-1 (at $i$ corresponding to $Y_i^* = 1$), and if $\overline{b_i} \leq \overline{b_i}'$ and $\overline{t_{ij}} \leq \overline{t_{ij}}'$ it has an expanded feasible region for choosing the delivery routes; so $Z_2 \leq Z_1$.

**Definition 1:** The percentage improvement in the robust cost that arises from tighter upper bounds is defined as $\Delta Z = 100 \times (Z_1 - Z_2)/Z_1$.

Note that the mobile clinic locations used in TS-MIP-2 were obtained by solving TS-MIP-1, and in general, these need not be optimal with the updated problem parameter estimates. If we had the ability to relocate mobile clinics in each time window, we could design a network and an associated outreach strategy with a possibly lower cost than $Z_2$ for the new time window. To see this, we define the following One-Stage Stochastic Mixed Integer Programming model (OS-MIP):

$$Z_0 = \text{Min}_{X,Y,Z,U,V,W} \text{Max} \sum_{1 \leq i \leq n} f_i Y_i + \sum_{t} \sum_{j} \sum_{k} c_{ij} Z_{ijk}$$

(28)
subject to

Constraints (2) – (4)

\[ \sum_i W_i = \sum_j \tilde{b}_j X_{ij} \quad \text{for } \forall i \leq n \] (29)

Constraints (6) – (13)

\[ \sum_i \sum_j (\tilde{t}_{ij} + s_i)Z_{ijk} \leq r \quad \text{for } \forall k \] (30)

Constraints (15) – (21)

\[ \tilde{t}_{ij} \in (t_{ij}^l, t_{ij}^u); \tilde{b}_i \in (b_i^l, b_i^u) \]

Similar to Proposition 1 we have

**Proposition 2**: Program OS-MIP is equivalent to solving it with \( \tilde{t}_{ij} = \overline{t}_{ij} \) in (28), (30) and \( \tilde{b}_j = \overline{b}_j \) in (29).

The following proposition relates OS-MIP to TS-MIP-2:

**Proposition 3**: \( Z_0 \leq Z_2 \).

**Proof**: It is clear that OS-MIP is a relaxation of TS-MIP-2, with the option of picking locations other than those given by \( Y_l = Y_l^* \). Therefore, \( Z_0 \leq Z_2 \).

Note that the optimal value of OS-MIP (\( Z_0 \)) is yet another conservative cost, and corresponds to the theoretical best robust solution to the outreach problem for the second period. Any difference between \( Z_2 \) and \( Z_0 \) is due to the fact that in OS-MIP we have the freedom to update mobile clinic locations. We may also interpret this reduction as the value of having better information on the parameter bounds at the beginning of the first period, as opposed to having to wait for it until the beginning of the second period.

**Definition 2**: The value of information is defined as \( V = 100 \times (Z_2 - Z_0)/Z_2 \).
Thus $V$ is the percentage savings possible (in the worst-case scenario), from obtaining information in the form of correct bounds on $\bar{t}_{ij}$ and $\bar{b}_i$ at the beginning of the first period.

6. Numerical experiments

We tested the procedure introduced in the previous sections on data that we adapted from four countries in sub-Saharan Africa. Due to issues with data confidentiality we label these countries A through D. Country B is smaller and has a high population density. The other three are larger in area and have some pockets of dense population with others (such as desert areas) where it is much more sparsely populated. To explain our numerical experiments and demonstrate some of the insights to be gained, we will first describe in detail an illustrative example with data derived from Country D. Following this, we analyze and summarize results from a larger set of instances.

Our illustrative example has 9 population centers on a 20 km by 20 km graph, with the depot is in the middle of the graph. Each population center is assumed to have an average of 100 newborns in a year. For implementation in the first period a robust solution (with value $Z_1$) is obtained for problem TS-MIP-1 using initial estimates of upper bounds on demand and travel times. The mobile clinic locations and the population centers assigned to each location in this solution are then fixed, then using the most current information on demands and travel times vehicle routes are updated for the second period by obtaining a robust solution (with value $Z_2$) to problem TS-MIP-2. We also obtain via OS-MIP the theoretical best robust solution to the period 2 problem (with value $Z_0$). We assume that the bounds on the updated estimates are always tighter than the initial ones, and we study upper bound changes (i) only in $b$ (demand), (ii) only in $t$ (travel times), and (iii) in both $b$ & $t$. 
We first generated a base case for the period 2 problem with associated values for $\bar{b}_j$ and $\bar{t}_{ij}$, and obtained $Z_0$ by solving problem OS-MIP. This solution with a value of $Z_0 = 619.17$ represents the best robust solution obtainable for period 2 and is illustrated in Figure 3: it has four clinics (at locations 1, 2, 3, and 5) with two trips (0-1-2-0) and (0-3-5-0). The mobile clinics at locations 2, 3 and 5 also cover the populations at 7, (4,6) and (8,9) respectively. The cost of 619.17 is comprised of 290.50 in facility costs and 328.67 in operating costs.

Next, for each of the three types of parameter changes ($b$, $t$, and $b$ & $t$) we studied (a) small, (b) moderate and (c) large reductions from the initial estimates of the upper bounds before the first period ($\bar{b}_j$ and $\bar{t}_{ij}$). Specifically, for these three cases we assumed that $\bar{b}_j$ and $\bar{t}_{ij}$ were on average 20%, 80% or 150% larger than their values before the second period ($\bar{b}_j'$ and $\bar{t}_{ij}'$). In all cases we first solve problem TS-MIP-1 with the appropriate values of $\bar{b}_j$ and/or $\bar{t}_{ij}$ to obtain the robust
solution for the first period, along with its value $Z_1$. We then fix clinic locations and their allocations to solve problem TS-MIP-2 using $\overline{b_j}$ and $\overline{t_{ij}}$ for the bounds, and obtain the robust solution for the second period, along with its value $Z_2$. The results are listed in Table 1, and we discuss some of the insights that these offer.

Table 1 Example in Country D

| Case | $b$       | $t$        | $b & t$   |
|------|-----------|------------|-----------|
|      | Small     | Moderate   | Large     |
| $Z_1$| 619.17    | 619.17     | 1003.68   |
|      | 651.29    | 714.19     | 778.49    |
|      | 693.39    | 831.13     | 1682.26   |
| $Z_2$| 619.17    | 619.17     | 817.49    |
|      | 630.32    | 630.32     | 642.03    |
|      | 619.35    | 627.30     | 923.71    |
| $\Delta Z$ | 0.00% | 18.48% | 3.22% | 1.77% | 24.26% | 0.03% | 1.29% | 32.97% |
| $V$  | 0.00%     | 24.26%     | 1.77%     | 3.56% | 0.03% | 1.29% | 32.97% |

Note: Theoretical best robust optimum for period 2 = $Z_0$=619.17

First, it may be seen that with the tighter bounds, the robust optimum for TS-MIP-2 ($=Z_2$) shows improvement over that for TS-MIP-1 ($=Z_1$) in seven of the nine cases, with the percentage improvement ($\Delta Z$) being much more significant when the upper bounds get tighter (i.e., large reductions). While these results are intuitive, it is interesting that the improvements are more pronounced with tighter time estimates as compared to tighter demand estimates (and simultaneous reduction of uncertainty in both parameters further magnifies the savings).

Next, we look at the issue of what we could have achieved in the second period if we had been able to re-optimize locations and allocations. That is, we compare the robust optimum $Z_2$ from TS-MIP-2 to its theoretical lowest value of $Z_0=619.17$. In particular, we compute the theoretical maximum percentage improvement possible in $Z_2$, i.e., the value of information ($V$) as given by Definition 2. It may be observed that for this example, the updated robust solution to the second period problem is actually very close to the theoretical best value in seven of the nine cases.
The only instances where the value of information is high is when there are large reductions in the estimated upper bounds for demand, especially when there are simultaneous reductions in the upper bound on travel times. We visually illustrate some of the results in Figures 4 to 7.

![Solution in Period 1 with initial demand estimates](image)

*Figure 4: Solution in Period 1 with initial demand estimates*

First, consider demand estimates. If the revised estimates in the upper bounds are only slightly or moderately tighter (columns 1 and 2 in Table 1), we find the theoretical best solutions at the beginning and this does not change with revised estimates; thus there is no value to these revised estimates. In contrast, if there is a large reduction in the estimate from period 1 to period 2 (column 3 in Table 1) the situation is different. Figure 4 illustrates the solution to problem TS-MIP-1 for the first period, with six clinics scheduled via five outreach trips covering 1, 2, (3, 9), 5, 6 respectively. Note that clinics at 2, 9 and 6 also cover the populations at locations 7, 8 and 4, respectively in this solution. As displayed in Table 1, this solution has a total cost of 1003.68.
Now, consider period 2 where we are constrained to maintain outreach clinics at these same six locations, but use the updated information on the worst-case demand to solve problem TS-MIP-2. This yields the updated solution shown in Figure 5 with two outreach trips \((0, 1, 5, 9, 0)\) and \((0, 3, 6, 2, 0)\) to cover the six clinics, and a total cost of 817.49. Locations 7, 8 and 4 are covered by the clinics at 2, 9 and 6, respectively. Note that the very loose initial upper bound for demand caused the robust solution to the first period to have more trips because a much larger demand could cause vehicle capacity constraints to be violated if the period 2 solution is adopted. In both solutions we have facility costs of 435.75 for the six open mobile clinics, but operating costs of 567.93 in the first period as opposed to 381.74 in the second.

![Figure 5: Updated solution in Period 2: large reductions in demand estimates](image)

Finally, note that if we had had the updated information prior to period 1, we would have obtained the best overall robust solution (shown in Figure 3), which is 24.26% better than the one from TS-MIP-2; this is the value of information in this instance.
Next we look at travel time; Figure 6 is for the case where upper bounds on travel times are tightened slightly or moderately, i.e., when the initial upper bound estimates were either 20% or 80% larger on average than their updated values in period 2 (columns 4 and 5 in Table 1). The robust solutions for both periods are identical in both cases, and the reductions from $Z_1$ to $Z_2$ (651.29 to 630.32, and 714.19 to 630.32, respectively) are only because $\overline{t}_{ij}' < \overline{t}_{ij}$. Also, the only difference between Figure 6 and the theoretical best design for period 2 shown in Figure 3 is that a clinic is located at 8 instead of 5. This is because $\overline{t}_{05}$ happens to be larger than $\overline{t}_{08}$, and thus in period 1, location 8 is preferable to location 5 in the solution to TS-MIP-1. When in period 2 the clinic is fixed at location 8 and we use $\overline{t}_{05}'$ and $\overline{t}_{08}'$ in Problem TS-MIP-2, the cost ($Z_2=630.32$) is 11.15 units higher than it would have been ($Z_0=619.17$) with the optimal location (i.e., $V=1.77\%$).

Figure 6: Solutions in Periods 1 & 2: small or moderate reductions in travel time estimates

Figure 7 depicts the case where there is a large reduction (column 6 in Table 1) in the initial estimates of travel time ($\overline{t}_{ij}$ exceeds $\overline{t}_{ij}'$ by 150% on average). Again, the solutions are identical in
both periods, and in contrast with the case when reductions in the bounds are small or moderate, a clinic is now assigned to location 7 instead of location 2. Using these locations as opposed to the optimal ones in Figure 3 yield a period 2 cost of $Z_2=642.03$ that is 22.86 units higher than the theoretical best value of $Z_0$ (with $V=3.56\%$).

![Figure 7 Solutions in Period 1 & 2: large reductions in travel time estimates](image)

With simultaneous changes in demand and travel times parameter estimates (columns 7, 8 and 9 in Table 1) similar types of results are obtained; figures are omitted in the interest of space.

We now present results from a larger set of instances across all four countries using the same methodology. The resulting values for the value of information ($V$) are summarized in Tables 2, 3 and 4 based upon the demographic characteristics of the region. Table 2 shows results for examples from Country B and regions of Countries A, C, and D with high population densities (e.g., around their capital cities). Table 3 has instances from Countries A, C, and D where populations are moderately distributed, while Table 4 covers larger, often remote regions in the
same countries, with relatively sparse populations. Note that the examples in Table 4 have fewer population centers that are more sparsely distributed on a larger graph, while the examples in Table 2 have more population centers on a relatively small graph; the examples in Table 3 are in between these two extreme cases. Also note that smaller values for $V$ in the tables indicate worst-case solutions that are relatively robust with our approach, while larger values indicate that having tighter, more accurate initial estimates of parameter bounds can result in more significant savings over the solutions from our approach. Insights that can be drawn from the computational results in these tables are discussed in detail in the following section.

*Table 2 Value of information for examples in smaller, densely populated regions*

|       | Small | Moderate | Large | Small | Moderate | Large | Small | Moderate | Large | Small | Moderate | Large |
|-------|-------|----------|-------|-------|----------|-------|-------|----------|-------|-------|----------|-------|
| A1    | 9.72% | 17.71%   | 32.47%| 0.00% | 0.00%    | 0.00% | 9.72% | 17.71%   | 32.47%|       |          |       |
| B1    | 0.00% | 22.87%   | 50.99%| 0.38% | 1.96%    | 1.96% | 0.38% | 23.35%   | 50.99%|       |          |       |
| B2    | 0.00% | 15.47%   | 47.88%| 0.05% | 0.05%    | 0.05% | 0.05% | 15.47%   | 47.88%|       |          |       |
| B3    | 0.00% | 19.76%   | 49.24%| 0.00% | 1.09%    | 2.22% | 0.00% | 20.02%   | 49.24%|       |          |       |
| B4    | 31.22%| 31.70%   | 48.00%| 0.00% | 0.00%    | 31.87%|       |          |       | 31.22%| 31.70%   | 58.58%|
| B5    | 0.00% | 17.45%   | 30.35%| 0.26% | 0.26%    | 0.26% | 0.26% | 17.51%   | 30.87%|       |          |       |
| B6    | 0.00% | 15.54%   | 26.85%| 0.00% | 0.00%    | 1.14% | 0.00% | 15.63%   | 36.13%|       |          |       |
| C1    | 24.57%| 38.75%   | 68.49%| 0.00% | 0.00%    | 0.00% |       | 24.51%   | 38.82%| 68.49%|          |       |
| C2    | 0.58% | 24.08%   | 39.33%| 0.58% | 1.75%    | 2.04% | 0.58% | 25.19%   | 50.24%|       |          |       |
| D1    | 20.05%| 41.38%   | 65.20%| 0.27% | 2.27%    | 2.62% | 20.05%| 42.57%   | 65.20%|       |          |       |
| Mean  | 8.61% | 24.47%   | 45.88%| 0.15% | 0.54%    | 4.22% | 8.68% | 24.80%   | 49.01%|       |          |       |
| Median| 0.29% | 21.32%   | 47.94%| 0.02% | 0.15%    | 1.55% | 0.48% | 21.68%   | 49.74%|       |          |       |
### Table 3 Value of information for examples in moderately populated regions

|      | Small | $b$ | Moderate | Large | Small | $t$ | Moderate | Large | Small | $b$ & $t$ | Moderate | Large |
|------|-------|-----|----------|-------|-------|-----|----------|-------|-------|----------|----------|-------|
| A2   | 0.00% | 17.98% | 26.23% | 0.01% | 30.98% | 26.23% | 0.01% | 26.23% | 26.23% |
| A3   | 0.00% | 0.00%  | 16.21% | 0.00% | 0.00%  | 0.00%  | 0.00% | 0.00%  | 16.21% |
| A4   | 0.00% | 0.14%  | 30.88% | 0.00% | 0.00%  | 0.00%  | 0.00% | 10.83% | 30.88% |
| C3   | 0.30% | 0.30%  | 10.84% | 0.30% | 1.73%  | 2.02%  | 0.30% | 0.30%  | 18.59% |
| C4   | 0.00% | 9.52%  | 9.52%  | 0.00% | 0.00%  | 0.00%  | 0.00% | 0.00%  | 16.21% |
| D2   | 0.00% | 11.69% | 20.94% | 0.00% | 0.00%  | 0.00%  | 0.00% | 0.00%  | 35.22% |
| D3   | 0.00% | 0.00%  | 24.26% | 1.77% | 1.77%  | 3.56%  | 0.03% | 1.29%  | 32.22% |
| D4   | 11.52%| 11.52% | 20.67% | 0.00% | 0.00%  | 20.67% | 11.52%| 11.52% | 39.95% |
| D5   | 0.00% | 13.14% | 22.51% | 0.00% | 0.00%  | 0.00%  | 20.67%| 11.52% | 11.52% |
| D6   | 0.00% | 0.00%  | 12.19% | 1.69% | 0.00%  | 2.58%  | 2.58% | 0.00%  | 34.67% |
| Mean | 1.18% | 6.43%  | 19.42% | 0.38% | 3.45%  | 6.40%  | 1.44% | 8.49%  | 28.60% |
| Median| 0.00% | 4.91%  | 20.80% | 0.00% | 0.00%  | 2.30%  | 0.00% | 10.17% | 31.16% |

### Table 4 Value of information for examples in larger, sparsely populated regions

|      | Small | $b$ | Moderate | Large | Small | $t$ | Moderate | Large | Small | $b$ & $t$ | Moderate | Large |
|------|-------|-----|----------|-------|-------|-----|----------|-------|-------|----------|----------|-------|
| A5   | 0.00% | 0.00% | 0.00%    | 2.42% | 2.02% | 2.02% | 2.42% | 2.02%  | 17.99% |
| A6   | 0.00% | 0.12% | 0.12%    | 0.12% | 8.42% | 8.42% | 0.12% | 8.42%  | 8.42%  |
| A7   | 0.00% | 6.98% | 14.29%   | 0.00% | 14.29%| 14.29%| 0.00% | 14.29% | 14.29% |
| A8   | 0.00% | 0.00% | 16.00%   | 0.00% | 0.00% | 0.00% | 0.00% | 0.00%  | 16.00% |
| C5   | 6.81% | 6.81% | 6.81%    | 0.28% | 0.00% | 0.00% | 6.56% | 6.56%  | 17.78% |
| C6   | 0.00% | 0.00% | 7.97%    | 0.00% | 0.00% | 0.00% | 0.00% | 7.97%  | 15.40% |
| C7   | 0.00% | 5.51% | 5.51%    | 0.00% | 0.00% | 5.51% | 0.00% | 5.51%  | 11.57% |
| C8   | 0.00% | 0.00% | 0.00%    | 0.00% | 0.00% | 0.00% | 0.00% | 0.00%  | 0.00%  |
| D7   | 0.00% | 0.00% | 4.91%    | 2.05% | 4.91% | 4.91% | 2.05% | 4.91%  | 11.25% |
| D8   | 0.00% | 0.00% | 0.00%    | 0.66% | 0.96% | 0.96% | 0.00% | 0.96%  | 0.96%  |
| Mean | 0.68% | 1.94% | 5.56%    | 0.55% | 3.06% | 3.61% | 1.11% | 5.06%  | 11.37% |
| Median| 0.00% | 0.00% | 5.21%    | 0.06% | 0.48% | 1.49% | 0.00% | 5.21%  | 12.93% |
7. Discussion and Conclusions

In Tables 2, 3 and 4, we list the mean as well as median values for each of the nine cases studied for each of the three different demographic characteristics. We separately highlight instances with \( V \) values over 20\% and those with values under 5\% to indicate “large” and “small” values, respectively. The entries not highlighted may be thought of as being in between. First, consider changes in estimates of only \( b \) or only \( t \) (the first two columns in the tables for each). Although some exceptions exist, we can draw the general conclusion that our approach is quite robust (i.e., \( V \) is small) in the following situations:

- Revisions are only in the travel time estimates, regardless of whether they are small, moderate or large: the value of information is under 5\% in 79, and over 20\% in only 4 out of the 90 instances corresponding to this situation (the nine columns under “\( t \)”).
- Revisions are only in the demand estimates and they are small: \( V \) is under 5\% in 24 and over 20\% in only 3 out of the 30 instances for this case (the three columns under “\( b \)” and “Small”).
- Revisions are only in the demand estimates and they are moderate, but we are in larger areas with moderate to sparse population densities: \( V \) is under 5\% in 12 of 20 instances (the two columns under “\( b \)” and “Moderate” in Tables 3 and 4).

Conversely, the costs in the worst-case scenario can be higher with our approach than they would be if we had perfect information in advance (i.e., \( V \) is larger) under the following scenarios:

- There are large revisions in the bounds on demand: \( V \) is over 20\% in 16 and under 5\% in only 5 out of the 30 instances for this case, all of the latter for sparsely populated regions (the three columns under “\( b \)” and “Large”).

• There are moderate revisions in densely populated regions: $V$ is over 20% in 5 of the 10 instances for this case and never under 5% (column under “$b$” and “Moderate” in Table 2).

When we consider simultaneous changes in the estimated bounds for both travel time and demand estimate (the columns under “$b \& t$”), the results are closely correlated with what happens when there are changes in demand alone, leading one to conclude that demand revisions constitute the primary factor and their effects overshadow revisions in travel time estimates.

Next, we conducted a set of separate, nonparametric, Wilcoxon signed-rank tests for each of the three different types of regions ($d$, $m$, $s$) to see if there were differences in the magnitude of the mean effects of the different types of changes ($H_0: \mu_b = \mu_t$ vs. $H_1: \mu_b > \mu_t$ and $H_0: \mu_{b,t} = \mu_b$ vs. $H_1: \mu_{b,t} > \mu_b$ for each of $d, m, s$). Note that in each of the six comparison we have 30 paired instances across which we study difference in means. The null hypotheses were strongly rejected (P-values all under 0.001) for five out of the six tests; the only case where there was no significant difference in the mean value of $V$ was when comparing individual changes in estimates of $b$ and $t$ in sparse regions ($\mu_{stb}$ and $\mu_{st}$), which had a P-value of 0.66. In other words, (i) in dense and moderately populated regions, the value of $V$ with changes in only $b$ is significantly higher than that with changes in only $t$, and (ii) in all types of regions $V$ is significantly higher with changes in both $b$ & $t$ as compared with changes in only $b$.

Finally, we look at the general demographic characteristics to study their effect. When we have large, sparsely populated regions (Table 4), our approach is quite robust: $V$ is under 5% in 62 out of 90 instances, and always under 20%. When the population density is moderate (Table 3), our approach is still reasonably robust unless there are large changes in the estimated demand as observed previously, in which case the value of $V$ starts to increase. Finally, in smaller, densely
populated areas (Table 2) the results can be much more sensitive to changes in demand and there is significant value to obtaining more precise estimates of demand. Given that there are 90 instances for each type of region we conducted two simple one-sided Z-tests for equality of means \((H_0: \mu_d = \mu_m \text{ vs. } H_1: \mu_d > \mu_m \text{ and } H_0: \mu_m = \mu_s \text{ vs. } H_1: \mu_m > \mu_s)\). The null hypothesis is strongly rejected in favor of the alternative in both tests (P-values in both cases are under 0.0002), confirming that small, dense regions tend to have larger value of information than larger, moderately populated regions, and in turn, the latter yield larger value for \(V\) than large, sparsely populated regions.

Based on our computational study, we can draw two main conclusions. First, larger sparsely populated regions tend to have lower value of information, while the opposite is true for smaller more densely populated regions. This is explained by the fact that when there are fewer population centers and they are relatively far apart and can serve relatively fewer neighboring population centers, a larger fraction of the locations are selected for outreach, and capacity is less of an issue with fewer people being served by each outreach trip. Thus, even with perfect information, there is relatively little opportunity to revise the initial plan even when demand and travel times estimates change. Conversely, in smaller, denser regions there are more dependencies between population centers and more people are served in each trip. Thus changes in population estimates, and to a lesser extent, travel times, have a significant effect: often, the best strategy could be different from the plan that we obtain because capacities might be exceeded or alternative solutions to the set covering problem yield shorter vehicle routes. Thus the value of obtaining accurate information is much higher, and the solution with our approach might not be as robust.

Second, it is better to focus more on obtaining more accurate population (demand) estimates than on travel time estimates. The latter have relatively low value of information and our
approach is very robust even in smaller, densely populated regions with approximate estimates of these times. On the other hand, if demand estimates are too conservative we could arrive at a strategy that results in locations that are not cost-effective after we get updated information; thus it is important to be able to get good estimates of demand in order for our approach to be robust.

In summary, this paper presents a systematic way to plan for economical outreach operations by formulating the problem as a mixed integer program. It also studies the issues related to the typical uncertainties associated with estimating demand for vaccines and planning individual outreach trips and provides insights on where to focus attention if we are to follow a robust approach that plans for worst-case scenarios in order to comply with WHO-EPI guidelines to provide universal coverage.

**References**

1. World Health Organization, "Immunization," 2019. http://www.who.int/topics/immunization/en/. [Accessed 15 Aug, 2019].

2. World Health Organization, "The Expanded Programme on Immunization," 2013. http://www.who.int/immunization/programmes_systems/supply_chain/benefits_of_immunization/en/. [Accessed 15 Aug, 2019].

3. Gavi, "Gavi’s Mission," 2019. http://www.gavi.org/about/mission/. [Accessed 15 Aug, 2019].

4. J. I. Blanford, S. Kumar, W. Luo and A. M. MacEachren, "It’s a long, long walk: accessibility to hospitals, maternity and integrated health centers in Niger.," *International journal of health geographics*, vol. 11, no. 1, p. 24, 2012.
5. P. O. Ouma, J. Maina, P. N. Thuranira, P. M. Macharia, V. A. Alegana, M. English, E. A. Okiro, and R. W. Snow, "Access to emergency hospital care provided by the public sector in sub-Saharan Africa in 2015: a geocoded inventory and spatial analysis," *The Lancet Global Health*, vol. 6, no. 3, pp. 342-350, 2018.

6. O. O. Malande, D. Munube, R. N. Afaayo, K. Annet, B. Bodo, A. Bakainaga, E. Ayebare, S. Njunwamukama, E. A. Mworozzi, and A. M. Musyoki, "Barriers to effective uptake and provision of immunization in a rural district in Uganda," *PLOS ONE*, vol. 14, no. 2, 2019.

7. World Health Organization, "Immunization coverage," 2019. http://www.who.int/news-room/fact-sheets/detail/immunization-coverage. [Accessed 15 Aug, 2019].

8. World Health Organization, "Mobile clinics provide health care to people in some of the world’s worst crises," 2016. http://www.who.int/hac/donorinfo/2016/mobile_clinics/en/. [Accessed 15 Aug, 2019].

9. M. S. Daskin and L. K. Dean, "Location of health care facilities," in *Operations Research and Health Care: A Handbook of Methods and Applications*, Springer, Springer, 2004, pp. 43-76.

10. World Health Organization, "Expanded programme on immunization: study of the feasibility, coverage and cost of maintenance immunization for children by district mobile teams in Kenya," *The Weekly Epidemiological Record*, vol. 52, no. 24, pp. 197-199, 1977.

11. S. I. Chen, B. A. Norman, J. Rajgopal, T. M. Assi, B. Y. Lee and a. S. T. Brown, "A planning model for the WHO-EPI vaccine distribution network in developing countries," *IIE Transactions*, vol. 46, no. 8, pp. 853-865, 2014.
12. J. Lim, "Improving the design and operation of WHO-EPI vaccine distribution networks," Doctoral Dissertation, Retrieved from http://d-scholarship.pitt.edu/28696/, University of Pittsburgh, Pittsburgh, PA, 2016.

13. Y. Yang, H. Bidkhori and J. Rajgopal, "Optimizing vaccine distribution networks in low and middle-income countries," arXiv preprint arXiv:1907.13434, 2019.

14. Y. Yang and J. Rajgopal, "An iterative cyclic algorithm for designing vaccine distribution networks in low and middle-income countries," arXiv preprint arXiv:1907.10777, 2019.

15. J. Lim, E. Claypool, B. A. Norman and J. Rajgopal, "Coverage models to determine outreach vaccination center locations in low and middle income countries," Operations Research for Health Care, vol. 9, pp. 40-48, 2016.

16. M. H. Mofrad, "Optimizing vaccine clinic operations in low and middle income countries," Doctoral Dissertation, Retrieved from http://d-scholarship.pitt.edu/27262, University of Pittsburgh, Pittsburgh, PA, 2016.

17. R. Church and C. ReVelle, "The maximal covering location problem," Papers of the Regional Science Association, vol. 32, no. 1, pp. 101-118, 1974.

18. O. Berman and D. Krass, "The generalized maximal covering location problem," Computers & Operations Research, vol. 29, no. 6, pp. 563-581, 2002.

19. M. S. Daskin and L. K. Dean, "Location of health care facilities," in Operations Research & Health Care: A Handbook of Methods & Applications, 2004, pp. 43-76.

20. S. N. Kumar and R. Panneerselvam, "A Survey on the vehicle routing problem and its variants," Intelligent Information Management, vol. 4, no. 3, pp. 66-74, 2012.

21. P. Toth and D. Vigo, "An overview of vehicle routing problems," SIAM Monographs on Discrete Mathematics and Applications, pp. 1-26, 2002.
22. J. K. Lenstra and A. H. G. Rinnoy Kan, "Complexity of vehicle routing and scheduling problems," *Networks*, vol. 11, no. 2, pp. 221-227, 1981.

23. S. R. Thangiah, K. E. Nygard and P. L. Juell, "GIDEON: A genetic algorithm system for vehicle routing with time windows," in *Proceedings of the Seventh IEEE Conference on Artificial Intelligence Applications*, Miami Beach, 1991.

24. J.-Y. Potvin and S. Bengio, "The vehicle routing problem with time windows, part II: genetic search," *INFORMS Journal on Computing*, vol. 8, no. 2, pp. 165-172, 1996.

25. E. Taillard, P. Badeau, M. Gendreau, F. Guertin and J.-Y. Potvin, "A tabu search heuristic for the vehicle routing Problem with soft time windows," *Transportation Science*, vol. 31, no. 2, pp. 170-186, 1997.

26. J. Homberger and H. Gehring, "Two evolutionary meta-heuristics for the vehicle routing problem with time windows: meta-heuristics for location and routing problems," *Information Systems and Operational Research*, vol. 37, no. 3, pp. 297-318, 1999.

27. P. Shaw, "Using constraint programming and local search methods to solve vehicle routing problems," in *Principles and Practice of Constraint Programming*, New York, Springer-Verlag, 1998, pp. 417-431.

28. P. Kilby, P. Prosser and P. Shaw, “Guided local search for vehicle routing problems with time windows: meta-heuristics” Advances and Trends in Local Search for Optimization, Boston: Academic Publishers, 1999, pp. 473-486.

29. J. Homberger and H. Gehring, "A two-phase hybrid meta-heuristic for the vehicle routing problem with time windows," *European Journal of Operational Research*, vol. 162, no. 1, pp. 220-238, 2005.
30. S. Masrom, S. Z. Z. A. M. Nasir and A. S. A. Rahman, "Software framework for vehicle routing problem with hybrid metaheuristic algorithms," in *Proceedings of the Applied Computing Conference*, Angers, 2011.

31. M. A. Figliozzi, "An iterative route construction and improvement algorithm for the vehicle routing problem with soft time windows," *Transportation Research Part C*, vol. 18, no. 5, pp. 668-679, 2010.

32. D. L. Miller, A. W. Tucker and R. A. Zemlin, "Integer programming formulations of traveling salesman problems," *Journal of the ACM*, vol. 7, pp. 326-329, 1960