Proceedings of the 2016 Industrial and Systems Engineering Research Conference
H. Yang, Z. Kong, and MD Sarder, eds.

Stochastic Optimization for Vaccine Vial Replenishment

Zahra Azadi, Harsha Gangammanavar and Sandra Ekşioğlu
Clemson University
Clemson, SC.

Abstract

We present a two-stage stochastic programming model to manage inventory of vaccine vials in developing countries. In these countries immunization programs often involve targeted outreach services in remote locations. An efficient replenishment and utilization plan is necessary to reduce wastage. Wastage incurs when doses are discarded from open vials after their safe use time. Our model allows for selecting an optimal portfolio of different sized vaccine vials, and corresponding vial-opening decisions. These decisions are made so as to minimize the system-wide costs in the face of demand uncertainty. We use the stochastic Benders decomposition (L-shaped) algorithm to solve the model. Computational results on instances built using data from outreach clinic are also presented.

Keywords
Vaccine vial replenishment, two-stage stochastic programming, Benders decomposition.

1. Introduction
One of the challenges faced by health care providers is identifying the inventory of vaccines necessary to ensure a successful immunization program while minimizing purchasing, inventory and wastage costs. Vaccines for preventable diseases — such as Bacillus Calmette-Guérin (BCG), measles, diphtheria-pertussis-tetanus (DTP) — are distributed in single and multi-dose vials. Wastage results when doses are discarded from opened vials after their safe use time. According to a World Health Organization (WHO) report, more than 50% of vaccines are wasted around the world [1]. The Global Alliance for Vaccines and Immunizations (GAVI) has requested countries to take measures to bring down vaccine wastage rates[1]. According to a recent document published by the UNICEF [2], "Highest vaccine wastage occurs at service delivery level (27% for DPT and 61% for BCG at outreach session site) as compared to the supply chain levels (Measles 3.5%, others $< 1\%$)".

Typically, doses from opened vials perish after their safe use time which usually ranges between 6 to 8 hours without refrigeration. Discarded doses contribute to what is known as open vial wastage (OVW). In a recent study in Bangladesh, it was estimated that the OVW for BCG, measles, DTP, and TT was 85%, 71%, 44.2%, and 36.6%, respectively [3]. While single-dose vials have zero OVW, they are more expensive than multi-dose vials. On the other hand, opening multi-dose vials may lead to wastage during low demand periods. Since the decision to open the vials has to be made before the demand has realized, one should expect the wastage to be higher when demand is irregular and unpredictable.

Immunization programs in developing countries have targeted outreach services in which healthcare officers are responsible to go to rural areas and hold free immunization sessions for a specific period of time. For example, in Bangladesh 94% of immunization programs are conducted through such services [3]. Due to their remote location the outreach centers seldom have refrigeration facilities. Therefore, an efficient inventory plan is necessary to minimize OVW and ensure the success of these services. Moreover, such plans should be made while considering the limited available budget, which is often the case in low or middle income countries.

Previously, OVW using historical data collected in Bangladesh was analyzed in [3]. However, they do not prescribe any plan for efficient inventory management which require optimization models. In [4] the total cost is minimized using a mixed integer programming model integrated with Monte-Carlo simulation by choosing optimal ordering decisions. On the other hand, a Markov Decision Process model was proposed in [5] to identify a vial-opening policy which minimized the expected OVW cost over a finite horizon. The authors used the backward induction algorithm to solve their probabilistic model. Both these studies address inventory management of a single vial type, and uncertainty in demand is characterized either probabilistically or through simulation. However, it is necessary to consider both

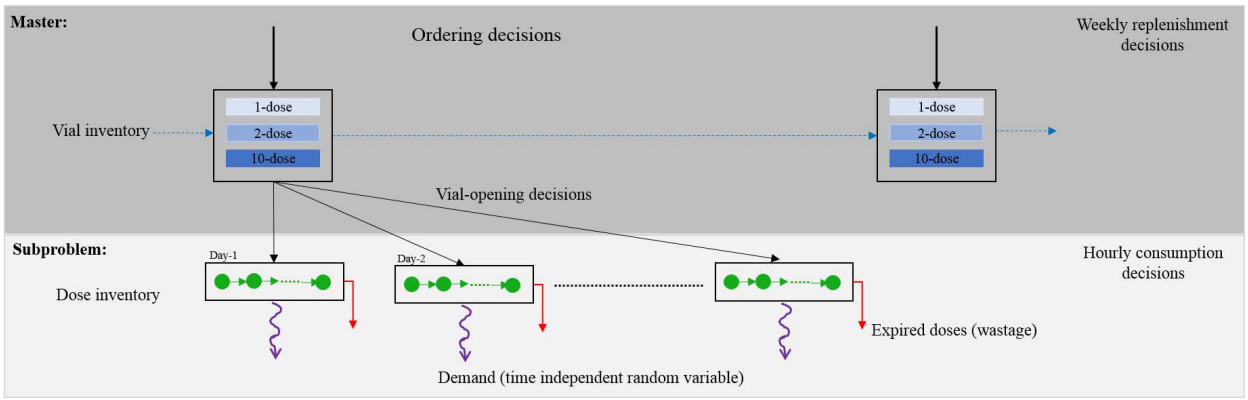


Figure 1: Inventory dynamics in vaccine vial replenishment model.

ordering as well as vial-opening decisions together in a vaccine vial replenishment model. Further, for problems with short horizon, for example the outreach services which last only a few month, an explicit representation of uncertainty is beneficial compared to a probabilistic model. In this regard, the main contributions of this work are the following:

- A stochastic programming model for vaccine vial replenishment which captures (a) the order timing and quantity for different sized vials, (b) corresponding vial-opening decisions, and (c) tradeoff between OVW and inventory management costs. The model considers separate inventory for different vial sizes and an inventory for doses. Uncertainty in demand is represented by scenarios generated using a real dataset.
- A computational framework which provides an analytical support to immunization programs. Experimental results obtained from this framework help us identify (a) an optimal portfolio of different sized vials for outreach centers located in different regions, and (b) necessary subsidies on purchase costs to achieve higher immunization levels.

In what follows, we present our stochastic programming model in §2. We discuss our solution approach in §3 and the computational results in §4. Ideas regarding extensions of the current model and a new solution approach are presented in §5.

2. Vaccine Vial Replenishment Model

We consider a replenishment model for an outreach clinic in a developing country. The model captures the operations of the clinic over a horizon of few month with respect to the inventory of vaccine vials. While replenishment order decisions are made on a weekly basis, consumption decisions have to be made at a much finer timescale. We will use $\mathcal{T} = 1,\ldots,T$ to denote the ordering decision epochs with each ordering period comprising of N consumption decision epochs. Therefore, $\mathcal{N} = 1,\ldots,NT$ will denote all the consumption decision epochs over the horizon. We consider vials of different sizes (single-dose, two dose, and ten dose) which are denoted by the set $\mathcal{V}$. Once the vials are opened they have to be consumed within their safe use time; we will use τ to denote this limit. We will assume that the safe use time of an open vial is the same as the maximum number of working hour per day.

At replenishment decision epoch $t \in \mathcal{T}$ order decisions z_t are made at a fixed cost of f_t. If a replenishment order is placed ($z_t = 1$), then the number of different sized vials is to be determined. We will denote these decisions by $x_{\nu t}$ and the corresponding variable purchase cost by $c_{\nu t}$ $\forall \nu \in \mathcal{V}$. At each consumption decision epoch $n \in \mathcal{N}$, vial-opening decisions $u_{\nu n}$ are made. Replenishment order decisions $x_{\nu t}$ and vial-opening decisions $u_{\nu n}$ together determine the state of vial inventory which we denote as $s_{\nu n}$ $\forall \nu \in \mathcal{V}$. The evolution of inventory for each vial size $\nu \in \mathcal{V}$ is captured by the following flow-balance equations:

$$s_{\nu Nt} = s_{\nu(Nt-1)} + x_{\nu t-1} - u_{\nu Nt} \qquad \forall t \in \mathcal{T}, \tag{1a}$$

$$s_{\nu n} = s_{\nu n-1} - u_{\nu n} \qquad \forall n \in \mathcal{N} \setminus \{N, 2N, \ldots, TN\}. \tag{1b}$$

The vial inventory is associated with a holding cost of $d_{\nu t}$. While (1a) captures the arrival of new orders $x_{\nu t}$ at replenishment decision epoch, (1b) ensures that the flow is balanced at the remaining consumption epochs. Recall that the order quantity for each vial size is determined only if an order is placed in replenishment decision epoch $t \in \mathcal{T}$. This is ensured by the following inequality:

$$\sum_{\nu \in \mathcal{V}} x_{\nu t} \leq M_t z_t \qquad \forall t \in \mathcal{T}, \tag{2}$$

where M_t is the limit on the number of vials ordered. Figure 1 illustrates the interactions between vial and dose inventories in the replenishment model.

The vials opened in consumption period n can be utilized in period $m < n+\tau$. We will use y_{nm} to denote the number of doses obtained from vials opened in period n, and used in period m. Using this notation $y_{n(n+\tau)}$ will represent the number of doses which are past their safe use time, and hence will contribute to OVW. The relationship between opened vials of different sizes and the number of doses available is captured by the following equation:

$$\sum_{m=n}^{n+\tau} y_{nm} = \sum_{\nu\in\mathcal{V}} q_\nu u_{\nu n} \qquad \forall n \in \mathcal{N}, \tag{3}$$

where q_ν is the number of doses in vial ν. The right-hand side of the above expression is the total number of doses obtained by opening vials in period n, while the left-hand side is the sum of utilized doses and expired doses.

The doses are used to serve the uncertain demand over all consumption periods $n \in \mathcal{N}$ which we represent by the random variable $D_n(\tilde{\omega})$. If the demand exceeds the available doses, then the unserved demand (denoted by r_n) is penalized at a uniform cost of p. The expired doses are associated with a uniform wastage cost g. With these, the dose inventory must satisfy the following flow-balance equation:

$$\sum_{m=n-\tau+1}^{n} y_{mn} + r_n = D_n(\omega) \qquad \forall n \in \mathcal{N}. \tag{4}$$

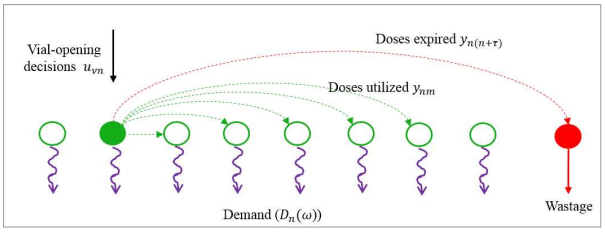


Figure 2: Dose utilization for $n = 2$ with $\tau = 6$.

Here ω is a realization of the random variable $\tilde{\omega}$. Figure 2 demonstrates how the doses obtained from opened vials in period $n = 2$ are utilized over $\tau = 6$ periods. The expired doses are represented by the node in red.

Replenishment decisions z_t and $x_{\nu t}$ as well as the vial-opening decisions $u_{\nu n}$ have to be made prior to demand realization. Therefore, these decisions are non-anticipative in nature [6]. On the other hand, the consumption decisions y_{nm} depend on the demand scenario, and therefore they can be made in an adaptive manner. Due to this we can write the replenishment model as a two-stage stochastic program with recourse.

$$\begin{aligned} \min \sum_{t\in\mathcal{T}} (f_t z_t + \sum_{\nu\in\mathcal{V}} c_\nu x_{\nu t}) + \sum_{\nu\in\mathcal{V}}\sum_{n\in\mathcal{N}} d_\nu s_{\nu n} + \mathbb{E}\{h(u,\tilde{\omega})\} \qquad (5)\\ s.t.\ (1a),(1b),(2)\\ z_t \in \{0,1\},\ \ x_{\nu t}, u_{\nu n}, s_{\nu n}, \in \mathbb{Z}^+ \qquad \forall n \in \mathcal{N}, t \in \mathcal{T}, \nu \in \mathcal{V}, \end{aligned}$$

where the recourse function $h(u,\omega)$ corresponding to the first stage vial-opening decisions $u_{\nu n}$ and realization ω of random variable $\tilde{\omega}$ is given by:

$$\begin{aligned} h(u,\omega) = \min \sum_{n\in\mathcal{N}} (g y_{n(n+\tau)} + p r_n) \qquad (6)\\ s.t.\ (3),(4)\\ y_{nm} \in \mathbb{Z}^+ \qquad \forall n \in \mathcal{N}, m \in \{n,\ldots,n+\tau+1\}. \end{aligned}$$

Notice that the decision in both master (5) and subproblem (6) are discrete variables. Uncertainty effects only the right-hand sides of equation (4) (demand realization). The recourse matrix characterized by left-hand sides in equations (3) and (4) and the transfer matrix characterized by right-hand side of equation (3) are independent of randomness. Therefore, the above two-stage stochastic program is said to have fixed recourse [6].

Moreover, the subproblem (6) can be viewed as a standard network flow problem. Therefore, the recourse matrix satisfies the total unimodularity property [7]. This implies that, when the right-hand side of (3) and (4) are integers (which is the case in our model) the linear programming solutions of subproblem (6) are integers as well, and thereby satisfy the requirement $y_{nm} \in \mathbb{Z}^+$.

3. Solution Approach

Two-stage stochastic programs are well studied in the literature. Decomposition based methods have been used to tackle these programs which construct piecewise linear approximation of expected recourse function [6]. A classic representative of such decomposition based methods is the stochastic Benders decomposition algorithm which is also known as the L-shaped method [8]. This method is motivated by cutting plane methods for deterministic problems.

Traditional mathematical programming constructs (such as dual variables) are used to update the approximation during each iteration of this algorithm. Dantzig-Wolf decomposition [9] and progressive hedging [10] are other methods which have been applied for two-stage programs.

When $\tilde{\omega}$ has finite support then the expectation function in (5) can be written as:

$$\mathbb{E}\{h(u,\tilde{\omega})\} = \sum_{i=1}^{\Omega} p^i h(u,\omega^i), \tag{7}$$

where Ω is the number of scenarios and p^i is the probability of scenario i. For methods using Monte Carlo simulation the number of scenarios is chosen based on computing resources (such as platform and solver) available, and for such instances one uses $p^i = 1/\Omega$. When the expectation function in (5) is written using (7) we obtain a large scale linear problem which is often termed as Extensive Scenario Formulation (ESF)[6].

We solve ESF reformulation of the two-stage program in (5)-(6) using the single-cut version of L-shaped method. Here we only present the computation of optimality cut, and refer the reader to [6] for details of the algorithm. To do this, let u^k_{vn} be the vial-opening decision in the kth iteration of the algorithm. For each $\{\omega^i\}_{i=1}^{\Omega}$, let π^{ki}_n and θ^{ki}_{vn} denote the optimal dual solution to subproblem (6) with (u^k_{vn}, ω^i) as input. The optimality cut is computed using these dual solutions, and has the following affine form:

$$\eta \geq \sum_{i=1}^{\Omega} p^i \left[\sum_{n\in\mathcal{N}} \left(\theta^{ki}_n D_n(\omega^i) + \sum_{v\in\mathcal{V}} \pi^{ki}_n q_v u_{vn} \right) \right], \tag{8}$$

This affine function is added to the master problem to obtain an updated approximation of the expected recourse function. Note that the subproblem in (6) satisfies the complete recourse assumption, and hence there is no need for feasibility cuts. Since there are finitely many extreme points for the subproblem, this algorithm is known to converge in finite number of iterations.

4. Computational Results

In our computational experiments we will verify the effect of arrival rate, variable purchase cost, and the viability of our solution approach.

Dataset: The study was conducted on instances built using dataset obtained from a clinic in Iran. This is an outreach clinic where replenishment decisions are made once a week, and the clinic operates for 6 hours a day and 5 days a week over a horizon of 4 weeks. We considered Pentavalent vaccines which are distributed in 1, 2, and 10 dose vials. We assume the ordering cost to be uniform over the horizon and set it to \$10. Different instances (P1-P8) were created by varying the purchase cost over the range specified in Table 1 of [11]. Inventory holding cost was set to 10% of the purchase cost, and OVW cost was set to a value greater than the highest purchase cost for all instances. The demand (patient arrival) is often assumed to follow Poisson distribution [3, 11]. The dataset included daily demand for vaccines, and hourly scenarios were generated by simulating from a Poisson distribution with daily demand as mean (assuming it to be stationary over the entire day). Experiments with different number of scenarios in the ESF reformulation were conducted. Based on our analysis 100 scenarios were found to appropriately represent the uncertainty in demand. Recall that the unserved demand is penalized in our model, we use a uniform penalty cost of \$5 in our experiments.

Platform: The single-cut version of the L-shaped method was implemented in c programming language on a 64 bit Intel Core i5 processor @ 1.9 GHz with 8 GB RAM. All linear and mixed integer programs were solved using CPLEX callable subroutines.

Setup: Our experiments begin by first predicting an optimal or near optimal (gap < 0.001) first stage solution using the L-shaped method. In the posterior analysis the quality of these first stage solutions is evaluated by fixing them and simulating the subproblem (6) for different demand realizations. The scenarios used to built ESF reformulation used for optimization as well as the scenarios used for evaluation are generated from the same distribution. This is often termed as out-of-sample-evaluation in stochastic programming literature. This evaluation is terminated when $(1-\alpha)\times 100\%$ confidence interval (CI) is built on the recourse function value. Here we summarize our observations:

1. *Effect of demand arrival:* The experiment was conducted on instance P5 with different arrival rates. The arrival rates were chosen to reflect those at outreach centers. The outreach centers see an average demand of less than 5 in remote locations, while the demand exceeds 10 in moderately populated regions. Table 1 presents the objective function value obtained from optimization as well as mean, standard deviation and CI from the evaluation runs. While optimization objective function value provides a lower bound on the problem, the evaluation results are an estimate of its upper bound. The results indicate that the total costs increases as the demand arrival rate

Mean daily arrival rate	Optimization	Evaluation		
	Obj. Val. ($)	Mean ($)	Std. dev. ($)	CI ($
1	192.7	193.5	0.2	[193.1, 1
5	443.1	453.4	0.5	[452.5, 4
10	544.9	573.4	0.6	[572.2, 5
15	766.2	779.5	0.7	[778.1, 7
20	796.4	840.6	0.8	[838.9, 8

Table 1: Results with varying mean arrival rate.

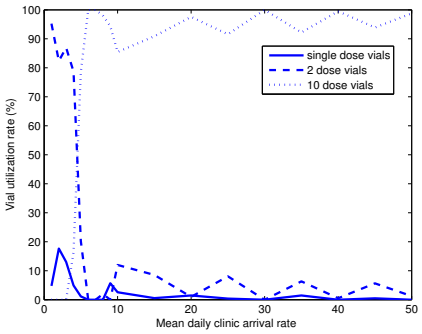


Figure 3: Vial portfolio at varying arrival rate.

Instance	Optimization	Evaluation		
	Obj. Val ($)	Mean ($)	Std. Dev. ($)	CI ($)
P1	195.5	195.8	0.2	[195.4, 196.3]
P2	224.4	226.7	0.2	[226.2, 227.2]
P3	344.4	346.6	0.2	[346.1, 347.1]
P4	355.2	357.8	0.4	[357.1, 358.6]
P5	359.5	362.1	0.4	[361.2, 362.8]
P6	362.2	361.1	0.5	[360.2, 362.3]
P7	365.2	363.2	0.5	[364.3, 366.2]
P8	370.2	372.2	0.5	[371.3,373.1]

Table 2: Results for instances with varying purchase cost.

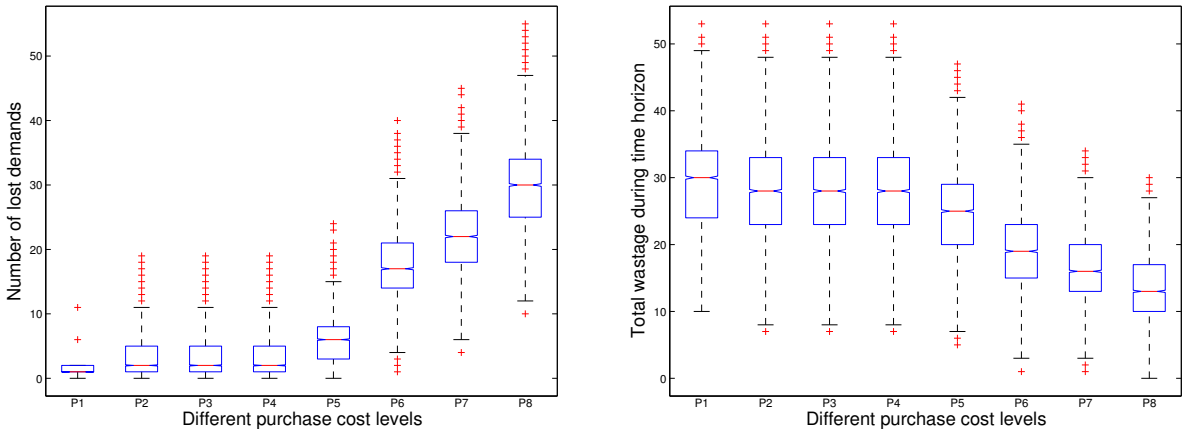
increases. Figure 3 shows how the ordering portfolio changes with patient arrival rate. As the arrival rate increases so does the usage of 10 dose vials. This can be attributed to the low cost per dose of multi-dose vials. At higher arrival rates the single-dose vials are ordered to supplement the multi-dose vials particularly to meet end of the day arrivals. Based on our results in Figure 3 it is economically efficient to procure single and two dose vials for outreach centers in remote locations and larger size vials for moderately populated regions.

2. *Effect of purchase cost:* The experiment was conducted on instances P1-P8 with varying purchase cost and constant mean demand rate of 8. Instance P1 has lowest purchase cost for all vaccine vial sizes and the costs increase as we go from instance P1 to P8, consequently it can be seen in Table 2 the optimal objective function value also increases. Figure 4a and 4b present the lost demand and wastage observed during evaluation. As purchase cost increases, conservative vial opening decisions are made to reduce wastage, and therefore a higher demand is lost. This analysis provides a tool for policy makers to recommend necessary subsidies on vaccine purchases to achieve a certain level of immunization. For example when the cost was set to their lower limit in instance P1 ($c_1 = 1.6, c_2 = 1.5, c_{10} = 1$) the lost demand was less than 2 for 99% (indicated by the top whisker) of scenarios tested, while it increases to 11 for P2-P4, and is as high as 43 for P8.

5. Conclusion and Future Research

In this work we presented a two-stage stochastic programming model for vaccine vial replenishment and usage for targeted immunization outreach centers. To the best of our knowledge this is the first stochastic optimization model which simultaneously captures both replenishment order and vial-opening decisions. The model minimizes costs and wastage due to unused doses at the end of the day. Our computational framework provides valuable analytical insights into the model. Experiments with different demand arrival rates indicate that different portfolios are economically suitable for outreach centers in different region. On the other hand, experiments with varying purchase cost illustrated the achievable immunization level at these costs. Such computational analyses will provide necessary tools to design effective immunization programs.

Since the L-shaped algorithm is an outer linearization method, the objective function values reported in Table 1 and 2 provide a lower bound on the ESF problem. When this objective function value is close to the value obtained from evaluation (upper bound), then one can conclude that the scenarios used in the ESF problem approximates the uncertainty fairly well. This was the case for all instances in Table 2 . However, when the objective function value



(a) Unserved demand **(b)** Wastage

Figure 4: Effect of varying purchase cost on dose utilization.

reported by optimization algorithm and the evaluation value are significantly different (for example instances with mean arrival rate ≥ 10 in 1) the selection of scenarios used for optimization is questionable. Moreover, there is no clear guideline for determining number of scenarios used. We are currently investigating application of sequential sampling algorithms, such as two-stage stochastic decomposition (SD) method [12], for our model. The SD method was originally designed for two-stage stochastic linear programs and does not require *a priori* selection of scenarios. Our work will extend this algorithm to accommodate discrete decision variables in the first stage.

As part of our future work, we will develop the demand prediction model which is based on population size, number of clinics and the birth cohort of developing countries. We will also evaluate the impact of pandemic outbreak. Since the model is intended for application in developing countries we will design simple implementable policies for vaccine vial distribution. These policies will be driven by our understanding of the model through extensive computational analyses.

References

[1] World Health Organization and others, 2005, "Monitoring vaccine wastage at country level," Guidelines for Program Managers.

[2] United Nations Children's Fund, 2010, "Vaccine wastage assessment."

[3] Guichard, S., Hymbaugh, K., Burkholder, B., Diorditsa, S., Navarro, C., Ahmed, S. and Rahman, M.M., 2010, "Vaccine wastage in Bangladesh," Vaccine, 28(3), pp.858-863.

[4] Dhamodharan, A., and Ruben A. P., 2012, "Determining the optimal vaccine vial size in developing countries: a Monte Carlo simulation approach," Health care management science 15, pp.188-196.

[5] Mofrad, M.H., Maillart, L.M., Norman, B.A. and Rajgopal, J., 2014, "Dynamically optimizing the administration of vaccines from multi-dose vials," IIE Transactions, 46(7), pp.623-635.

[6] Birge, J. R., & Louveaux, F., 2011, "Introduction to stochastic programming," Springer Science & Business Media.

[7] Dimitri P. B., 1998 "Network Optimization: continuous and discrete methods," Athena Scientific, 8.

[8] Van Slyke, R.M. and Wets, R., 1969, "L-shaped linear programs with applications to optimal control and stochastic programming," SIAM Journal on Applied Mathematics, 17(4), pp.638-663.

[9] Dantzig G. B. and Wolfe P., 1960, "Decomposition Principle for Linear Programs," Operations Research, 8(1) pp.101-111.

[10] Rockafellar, R. T. and Wets, Roger J.-B., 1991, "Scenarios and Policy Aggregation in Optimization Under Uncertainty," Math. Oper. Res., 16(29) pp.119-147.

[11] Lee, B.Y., Norman, B.A., Assi, T.M., Chen, S.I., Bailey, R.R., Rajgopal, J., Brown, S.T., Wiringa, A.E. and Burke, D.S., 2010, "Single versus multi-dose vaccine vials: an economic computational model," Vaccine, 28(32), pp.5292-5300.

[12] Higle, J. L. and Sen, S., 1996, "Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming."