MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **15**, Number **1**, February **2018**

doi:10.3934/mbe.2018011

pp. **255–273**

THREE-LEVEL GLOBAL RESOURCE ALLOCATION MODEL FOR HIV CONTROL: *A HIERARCHICAL DECISION SYSTEM APPROACH*

Semu Mitiku Kassa

Department of Mathematics and Statistical Sciences
Botswana International University of Science and Technology (BIUST)
P/Bag 16, Palapye, Botswana
and
Department of Mathematics
Addis Ababa University, P.O.Box 1176, Addis Ababa, Ethiopia

Abstract. Funds from various global organizations, such as, The Global Fund, The World Bank, etc. are not directly distributed to the targeted risk groups. Especially in the so-called third-world-countries, the major part of the fund in HIV prevention programs comes from these global funding organizations. The allocations of these funds usually pass through several levels of decision making bodies that have their own specific parameters to control and specific objectives to achieve. However, these decisions are made mostly in a heuristic manner and this may lead to a non-optimal allocation of the scarce resources. In this paper, a hierarchical mathematical optimization model is proposed to solve such a problem. Combining existing epidemiological models with the kind of interventions being on practice, a 3-level hierarchical decision making model in optimally allocating such resources has been developed and analyzed. When the impact of antiretroviral therapy (ART) is included in the model, it has been shown that the objective function of the lower level decision making structure is a non-convex minimization problem in the allocation variables even if all the production functions for the intervention programs are assumed to be linear.

1. Introduction.

Funds have been raised globally to help the fight for HIV and AIDS epidemics in low- and middle-income countries. In the year 2014 alone USD 19.2 billion was made available globally to finance the response to AIDS in such countries. Out of this, USD 8.5 billion was handled by the global organizations like *The Global Fund to fight AIDS, Tuberculosis and Malaria*, and *UNITAID* [1, 14, 25]. The type and intensity of response may vary from region to region and from country to country depending on the type of risk groups, rate of prevalence, social make-ups and economic conditions they have.

Globally over 36.7 million people are leaving with HIV and of these cases about 70.5% live in Sub-Saharan Africa, 13.8% live in Asia, 4.6% in Latin America, and around 3% live in Eastern Europe and central Asia. More than 1.8 million children under the age of 15 live with HIV/AIDS globally and among them more than 83 percent live in Sub-Sahara African region. A further 2.1 million people were infected

2010 *Mathematics Subject Classification.* Primary: 92B05, 92D30, 90B50, 91A80; Secondary: 34B60, 91A10, 90C31, 00A71.

Key words and phrases. Multilevel resource allocation, hierarchical planning, epidemiological modeling, HIV, hierarchical optimization.

in the year 2015 [24]. Every day over 5,700 new HIV infections occur worldwide. Out of these more than 95 percent occur in low- and middle-income countries and above 600 new cases per day are among children under 15 [24]. It is estimated that more than 90% of infant infection is caused by mother-to-child transmission which could have been prevented. In general global HIV prevalence (proportion of people with HIV) is remaining at the same level, although the global number of people with HIV is rising because of new infections and longer survival times, and continuously growing global total population.

In the absence of a curing medicine and a working vaccination, the investment to control the HIV epidemics is mainly to reduce the total number of new infections and the rate of progression to AIDS.

Untreated HIV progress to AIDS and death in most individuals. The antiretroviral (ARV) therapy to HIV infected people limits the viral replication allowing either immune preservation (at the earlier course of the infection) or immune reconstruction, resulting in durable, life-saving effects. The ARV treatment has two fold advantages: an advantage to the infected person as a treatment to increase the healthy life of the infected individual and an advantage to the society at large in decreasing the rate of infectiousness of the infected individual. The use of ART can significantly reduce the plasma viral load (by up to six orders of magnitude [16]) which decreases the rate of infectiousness of a treated person.

The other outcome of the ART intervention both to the individual and the society is that it reduces the costs of illness which may include hospitalization and the use of other expensive therapies. However, only about 17 million people worldwide received ART at the end of 2015 [24] and this amounts to only 46% of the total.

The intervention to reduce, and eventually eliminate the Mother-to-Child Transmission (MTCT), which is also called a vertical transmission, is through the use of ARVs during pregnancy and delivery and to the infants following births and the use of replacement infant feeding [4]. It is known that potent viral suppression greatly reduces perinatal HIV transmission.

To finance the interventions, the fund raised globally is very limited. It is only up to 70% of the required amount even at its pick stage [15, 25] (see Fig. 1).

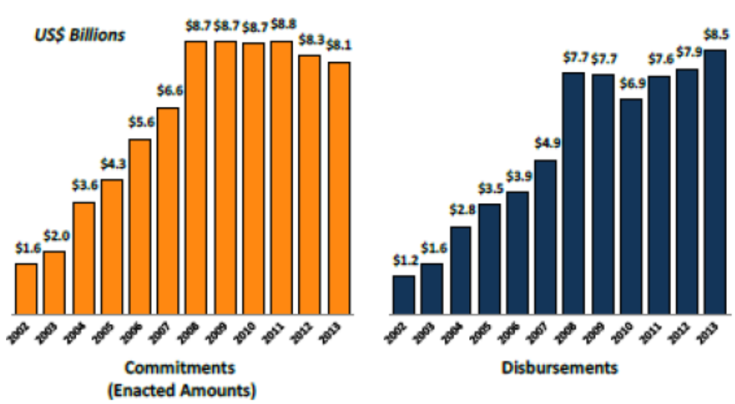

FIGURE 1. Global financing for the fight of HIV from Donor Governments: Commitments & Disbursements, 2002-2013. Taken from [15].

Due to economic difficulties or political reasons international aid may fluctuate significantly and hence the investment on health care could be highly affected, in particular in low- and middle-income countries. Therefore, the scarce resources have to be allocated optimally to get better and tangible achievements, in epidemic terms.

Health economics theory states that *allocating resources to medical interventions in increasing order of their cost-effectiveness ratios until the available budget has been exhausted will result in the optimal allocation of funds* [6, 27]. Actually this leads to a *Give-it-all or nothing* strategy. Moreover, to establish the Cost-Effectiveness criteria one needs to collect adequate data on the population distribution and generate cost-effectiveness data for individual and community-level interventions. Even in the perfect information case, 'Allocation by Cost-Effectiveness' (ACE) does not allow for several important factors such as, increasing or diminishing marginal returns to scale, mutual exclusivity of programmes, and interaction of program outcomes, which are necessary to be taken into consideration [18].

Therefore, several researchers have proposed the method of operations research in allocating resources for HIV interventions (see for instance [2, 3, 7, 8, 30, 31] and the references therein). However, all of them use one level-optimization procedures, which assume that resources are allocated by the same country where the programs are implemented. Such models may work well for resource allocation decisions in developed countries, where the fund is raised only from within the country itself.

A two level resource allocation model has been proposed in [18]. However, this model assumes that the lower level problem has a unique solution, which is not normally the case, especially when the production functions used are not linear (which is true even with the model formulation of the same paper). Therefore, it could be difficult to apply the proposed method in practical situations. Moreover, the effect of ART in reducing new infections was not considered in the model.

In this work a simplified epidemic model with the inclusion of the effect of ART will be used to formulate a global resource allocation model for HIV control. The paper shall concentrate on developing and analyzing the model where the decision for resource allocation passes through various levels. Such levels could be two or more depending on the organizational structure of the global financing agencies and on the mechanism they allocate resources to developing countries.

The remaining part of this paper is organized as follows: the resource allocation problem and the structure of the levels have been formulated and analyzed separately in Section 2, while Section 3 is devoted for the analysis of the model in its entirety. The paper is concluded with a discussion in Section 4.

2. Problem formulation.

The decision for international fund allocation for HIV prevention programs to low- and middle-income countries usually passes through various stages, with all of them aiming in controlling the epidemics and reducing human sufferings. We assume that major decision points in allocating funds can be classified into three hierarchical groupings; the upper level can be taken as the International funding organization, the middle level can be the Regional representatives and the lower level being the Country representatives. Due to the variation in prevalence, the main factors in the transmission mode of the disease, the economic condition and their social set-up, etc. countries with closer features (in addition to their geographic proximities) are classified by many international funding agencies in to some regions, like the Sub-Sahara Africa, Latin America, the Middle East and North Africa, Asia, etc. We take these regional forms as a middle level decision points in the resource allocation process. The diagram in Fig. 2 indicates the structure of the model under study.

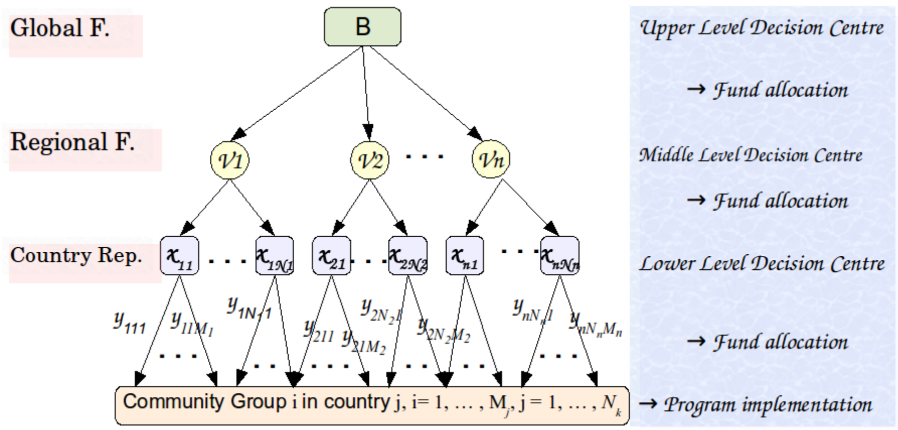

FIGURE 2. Schematic diagram for global resource allocations.

At each stage of the resource allocation process indicated in Fig. 2, the final common goal of investment, which is assumed to be controlling the epidemics and improving the quality of healthy life of the society, will be checked. That means, the investment is targeted to positively affect the dynamics of the epidemics. There are several models in literature which try to describe the dynamics of the HIV infection and spread in the population with ARV treatment. The schematic diagram in Fig. 3 indicates one of such models where interventions have impact on the disease dynamics.

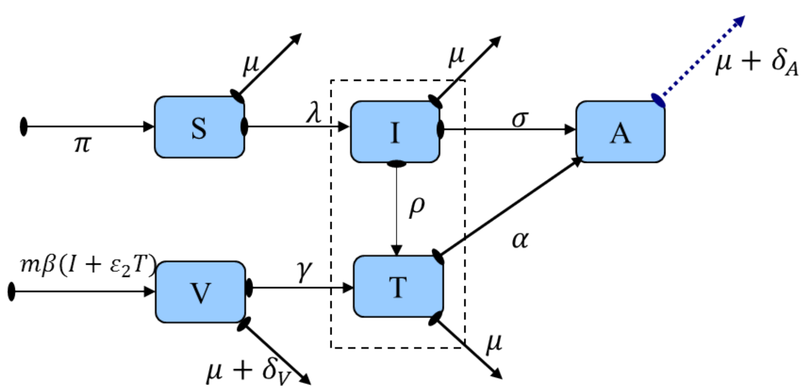

FIGURE 3. Model diagram that show flow of individuals between the compartments.

The dynamical system representing the schematic diagram in Fig. 3 will be given by the following set of differential equations.

$$
\begin{aligned}
\frac{\mathrm{d}V}{\mathrm{d}t} &= m\beta(I+\varepsilon_2 T) - (\gamma\rho + \mu + \delta_V)V \\
\frac{\mathrm{d}S}{\mathrm{d}t} &= \pi - \lambda\frac{S(I+\varepsilon_1 T)}{Q} - \mu S \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \lambda\frac{S(I+\varepsilon_1 T)}{Q} - (\rho+\sigma+\mu)I \\
\frac{\mathrm{d}T}{\mathrm{d}t} &= \rho(\gamma V + I) - (\alpha+\mu)T \\
\frac{\mathrm{d}A}{\mathrm{d}t} &= \sigma I + \alpha T - (\mu+\delta_A)A \\
& S(0)+V(0)+I(0)+T(0)+A(0) = 1,
\end{aligned}
$$

where $Q(t) = S(t) + I(t) + T(t)$ represents the total proportion of the sexually active population. Here, we assume that the population in symptomatic AIDS compartment (A) are not sexually active and do not contribute to the infections.

If we assume that the interventions have direct effect on some of the parameters in the system, allocation of resources should yield in decreasing the rate of infection (λ) and the rate of MTCT (m), in increasing the rate of recruitment to get ART (ρ), and in decreasing the rate of failure in adherence to the proper usage of ARV (α). Let $\phi_\lambda, \phi_m, \phi_\rho, \phi_\alpha$ denote general cost (or production) functions corresponding to each of the rates. Then the above dynamical system where these production functions now replacing the parameters is given by the following set of differential equations:

$$
\begin{aligned}
\frac{\mathrm{d}V}{\mathrm{d}t} &= \phi_m(y^1)\beta(I+\varepsilon_2 T) - (\gamma\phi_\rho(y^3) + \mu + \delta_V)V && (1)\\
\frac{\mathrm{d}S}{\mathrm{d}t} &= \pi - \phi_\lambda(y^2)\frac{S(I+\varepsilon_1 T)}{Q} - \mu S && (2)\\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \phi_\lambda(y^2)\frac{S(I+\varepsilon_1 T)}{Q} - (\rho+\sigma+\mu)I && (3)\\
\frac{\mathrm{d}T}{\mathrm{d}t} &= \phi_\rho(y^3)(\gamma V + I) - (\phi_\alpha(y^4)+\mu)T && (4)\\
\frac{\mathrm{d}A}{\mathrm{d}t} &= \sigma I + \phi_\alpha(y^4)T - (\mu+\delta_A)A && (5)\\
& S(0)+V(0)+I(0)+T(0)+A(0) = 1 && (6)
\end{aligned}
$$

If $N(t)$ is the total population size at time t, then $N(t)S(t)$ gives us the number of people in the susceptible compartment, $N(t)I(t)$ gives us the total number of infected individuals but not under treatment, and so on.

Since non-treated children with HIV have a 10 times higher mortality rate compared to those of other children [21], it is assumed that they will not become sexually active in the population (especially in low-income countries where child mortality rate is significantly high). However, those who are getting proper ARV treatment may mature to adulthood at a rate of γ. Thus, in the model the infected children under the age of 15 (the V) compartment are assumed to have little direct impact on the transmission of the disease. Given the higher rate of child mortality in low- and middle- income countries, this assumption does not deviate much from reality.

2.1. **Lower - level optimal resource allocation.** Assume that there are M_j target groups in country j within region k, where a program is to be implemented. Assume that each one of the program implementers address all the four intervention categories (*i.e.*, reducing the MTCT (m), decreasing the rate of infection due to sexual contacts (λ), increasing the rate of recruitment for the HIV infected people to get ARV (ρ), and decreasing the rate of failure in adherence to the proper use of ARV (α)). It is also assumed that there is no migration from one target group into the other within the given intervention time horizon.

The goal of the interventions program is mainly to minimize the total number of individuals who become newly infected, both in the category of the newly born children with an infected status ($\phi_m(y^1)\beta(I+\varepsilon_2 T)$) as well as in the category of adults that get infection through sexual activities ($\phi_\lambda(y^2)S(I+\varepsilon_1 T)/Q$), over a given time horizon, subject to a budget constraint, some equity constraints, and possibly a limit on attainability of some of the rates or parameters. At the beginning it seems that the effect of investing on treatment of the infected ones may not have a direct impact on reducing the incidence rate. However, giving ARV treatment to the infected individuals has an epidemiological advantage to the community in that the rate of transmission can be reduced by up to 92% [5] and the rate of mother-to-child transmission by at least 98% [17, 20, 22]. Therefore, it is very important in reducing the incidence rate in the population and this impact can be seen through the dynamics of the objective function.

The objective function of the resource optimization problem at the lower level is, therefore, simply taking the total sum of the contribution of the incidence of the newborns and the incidence of adults due to sexual contacts in the given period of time in all the community groups of a country, which is a function of the investment variables.

If a total of x_{kj} is allocated for country j in region k and there are M_j different community groups in country j, the total fraction of new infections within time horizon τ in country j of region k, where $\mathbf{y}_{kji} = (y^1_{kji}, y^2_{kji}, y^3_{kji}, y^4_{kji})$ represents the decision (investment) variables vector, is given by

$$\begin{aligned} f_{kj}(\mathbf{y}_{kji}) \quad = \quad & \sum_{i=1}^{M_j} \int_0^\tau e^{-rt} \Bigg\{ \left(\phi_{\lambda_i}(y^2_{kji}) S_i(t) \left[\frac{I_i(t) + \varepsilon_1 T_i(t)}{Q_i(t)} \right] \right) \\ & + \phi_{m_i}(y^1_{kji}) \beta \left(I_i(t) + \varepsilon_2 T_i(t) \right) \Bigg\} \mathrm{d}t, \end{aligned} \tag{7}$$

where $Q_i(t) = S_i(t) + I_i(t) + T_i(t)$ represents the total proportion of the sexually active population in community group i of country j at time t. In Equation (7), the two other investment variables (y^3 and y^4) do not appear exclusively in the objective function. However, previous investments on ρ and α have an impact on the number of the individuals in the treated cohort and due to reduced infectiousness of these individuals their contribution to the emergence of new infections is significant at the current time. To include this effect in the above objective function, since the closed form solution of the system (1 – 6) is not known, we will approximate the value of T from Equation (4) for a reasonably small change in time δt at any given time t as follows.

$$\begin{aligned} T(t+\delta t) \quad &= \quad T(t) + \left[\phi_\rho(y^3) \left(\gamma V(t) + I(t) \right) - (\phi_\alpha(y^4) + \mu) T(t) \right] \delta t \\ &= \quad \phi_\rho(y^3)[\gamma V(t) + I(t)] \delta t - \left[\left(\phi_\alpha(y^4) + \mu \right) \delta t - 1 \right] T(t) \end{aligned}$$

By reformulating this last equation we can get the value of T at the current time (from the previous value in time) to be

$$T(t) = \phi_\rho(y^3)[\gamma V(t-\delta t) + I(t-\delta t)]\delta t - \left[\left(\phi_\alpha(y^4) + \mu\right)\delta t - 1\right] T(t-\delta t), \tag{8}$$

where, $t - \delta t$ represents the time which is one step before the current time t. In terms of the formulation of (8) the objective function given in Equation (7) now becomes

$$\begin{aligned} f_{kj}(\mathbf{y}_{kji}) \quad = \quad & \sum_{i=1}^{M_j} \int_0^\tau e^{-rt} \Bigg\{ \phi_{\lambda_i}(y^2_{kji}) \frac{S_i(t)}{Q_i(t)} \Bigg[I_i(t) \\ & \quad + \varepsilon_1 \Bigg(\phi_{\rho_i}(y^3_{kji})[\gamma V_i(t-\delta t) + I_i(t-\delta t)]\delta t \\ & \quad - \left[\left(\phi_{\alpha_i}(y^4_{kji}) + \mu\right)\delta t - 1\right] T_i(t-\delta t) \Bigg) \Bigg] \\ & + \phi_{m_i}(y^1_{kji})\beta \Bigg[I_i(t) + \varepsilon_2 \Bigg(\phi_{\rho_i}(y^3_{kji})[\gamma V_i(t-\delta t) + I_i(t-\delta t)]\delta t \\ & \quad - \left[\left(\phi_{\alpha_i}(y^4_{kji}) + \mu\right)\delta t - 1\right] T_i(t-\delta t) \Bigg) \Bigg] \Bigg\} \mathrm{d}t, \end{aligned} \tag{9}$$

which is a nonlinear (non-convex) function of the investment variables. Even in the cases when the cost functions $\phi_{\lambda_i}, \phi_{m_i}, \phi_{\rho_i}$, and ϕ_{α_i} are linear, the criteria function (Equation (9)) is a non-convex function.

In determining the value of $f_{kj}(\mathbf{y}_{kji})$ of Equation (9), in practice we apply numerical integration. Thus, in numerical integration process, one need to take care of the values at time steps t and $t - \delta t$.

In the expression above, the variable vectors y^ℓ_{kji} represent the amount of investment for interventions aimed in reducing one of the rates: $\ell = 1$ for the mother-to-child transmission rate, $\ell = 2$ for the infectious contact rate, $\ell = 3$ for the rate of recruitment to get treatment, and $\ell = 4$ for the rate of failure to adhere in the proper use of the treatment due to various reasons.

If the planning time horizon τ is short enough (in application the actual planning periods could be 1 year or 3 years), then the discounting rate can be approximated by 0. Hence in such cases we can neglect the effect of discounting and set $e^{-rt} = 1$ for the whole planning period.

Therefore, the resource allocation problem at the lower decision point will be given by a non convex optimization problem:

$$\begin{aligned} & \min_{\mathbf{y}_{kji}} N_j f_{kj}(\mathbf{y}_{kji}) \\ & \textbf{Subject to} \\ & \qquad \sum_{i=1}^{M_j} \left(y^1_{kji} + y^2_{kji} + y^3_{kji} + y^4_{kji}\right) \quad \leq \quad x_{kj} \end{aligned} \tag{10}$$

Here, the constraint reflects that the total investment in country j should not exceed the total budget allocated to it from the higher level decision making body.

The solution $\mathbf{y}^*_{kji}$ of the minimization problem (10) which gives the optimal allocation of the total fund x_{kj} for country j in region k that satisfy each of the constraints set by higher-level decision makers. However, since the objective function f

is nonlinear and non-convex even if all the involved cost functions are modeled to be linear, problem (10) is a non-convex optimization problem in investment variables.

2.2. **Middle - and upper - level optimal resource allocations.** On the other hand the Middle-Level decision maker (region k) will solve the following problem in reallocating the total budget v_k received from the upper - level decision maker.

$$\begin{aligned} \max_{x_{k1},\ldots,x_{kN_n}} \quad & F_k = h_{k1}x_{k1} + h_{k2}x_{k2} + \cdots + h_{kN_n}x_{kN_n} \\ & \textbf{Subject to} \\ & x_{k1} + x_{k2} + \cdots + x_{kN_n} \leq v_k \end{aligned} \tag{11}$$

where h_{kj} represents the estimated number of HIV infections prevented in country j over time horizon τ for each money unit invested from region k.

One possible way of estimating the values of h_{kj} could be using the formulation in [31]. This could be done by choosing appropriate weights w_1 and w_2 and calculating the values:

$$h_{kj} = w_1 \frac{r^1_{kj} \times e^1_{kj}}{N_{kj} \times c^1_{kj}} + w_2 \frac{r^2_{kj} \times e^2_{kj}}{N_{kj} \times c^2_{kj}},$$

where,

c^1_{jk}	=	an average estimated cost per person to implement one intervention aiming to avert new infections in country j of region k,
c^2_{jk}	=	an average estimated cost per person to implement one intervention aiming to avert new progression to AIDS in country j of region k
r^1_{kj}	=	the baseline number of new infections that will occur in country j of region k over time τ
r^2_{kj}	=	the baseline number of new AIDS cases that will progress in country j of region k over time τ
e^1_{kj}	=	total number of expected potential infections averted (i.e. the number of infections that would occur in the absence of any investments) per person in country j of region k
e^2_{kj}	=	total number of expected potential AIDS cases prevented from progress per person in country j of region k
N_{kj}	=	The size of population to be addressed during the intervention in country j of region k.

The coefficients h_{kj}, thus calculated, carry the aggregate information about the effectiveness of the intervention in every country. Here, we also assume that the failure to adhere to the proper use of ARV will result in progression to AIDS.

If a solution $x^*_{kj}(v_k)$ is obtained for the middle level problem (11), it results in a solution function $F_k(v_k)$ that gives the total optimal number of HIV infections averted and the total new AIDS cases prevented within the course of time τ in all countries and in all the regions $k = 1, \ldots, n$. Thus $F_k(v_k)$ will take the form [18]:

$$F_k(v_k) = h_{k1}x^*_{k1}(v_k) + h_{k2}x^*_{k2}(v_k) + \cdots + h_{kN_k}x^*_{kN_k}(v_k) \tag{12}$$

The upper level problem will then be finding an optimal allocation $v^*_1, \ldots, v^*_n$ of the fund to each region which maximize the global gain in health. Thus, the decision maker at the upper level will solve the optimization problem:

$$\begin{aligned}
&\max_{v_1,\ldots,v_n} \sum_{k}^{n} F_k(v_k) \\
&\textbf{Subject to} \qquad\qquad (13)\\
&v_1 + v_2 + \cdots + v_n \leq B \\
&x_{kj} \geq d_{kj} v_k \text{ for each } k = 1 \text{ to } n \\
&\sum_{i=1}^{M_j} y^{\ell}_{kji} \geq z^{\ell}_{kj} v_k \text{ for each } \ell = 1,2,3,4,
\end{aligned}$$

where, $0 \leq d_{kj}, z^{\ell}_{kj} \leq 1$ are predetermined parameter values which are usually assigned by the upper level decision maker as equity values described by the intervention policy. The interpretation of these parameters could be: z^{ℓ}_{kj} may represent the minimum proportion of the allocated fund for region k which is to be invested in intervention type ℓ in country j of region k, whereas d_{kj} may represent the minimum proportion of the allocated fund for the region that should be assigned for country j in region k. If such values are considered to be not necessary or are not readily given, one may simply take a value 0 as a minimum threshold for each of them.

3. **Analysis of the model.** When we combine all the three problems described in the previous section in sequential order the general resource allocation problem takes the form,

$$\begin{aligned}
&\max_{v_1,\ldots,v_n} \sum_{k}^{n} F_k(v_k) \\
&\quad \textbf{Subject to} \\
&\quad v_1 + v_2 + \cdots + v_n \leq B \\
&\qquad x_{kj} \geq d_{kj} v_k \text{ for each } k = 1 \text{ to } n \\
&\qquad \sum_{i=1}^{M_j} y^{\ell}_{kji} \geq z^{\ell}_{kj} v_k \text{ for each } \ell = 1,2,3,4 \qquad \text{where } x_{kj}, y^{\ell}_{kji} \text{ solve} \\
&\max_{x_{k1},\ldots,x_{kN_n}} F_k = h_{k1}x_{k1} + h_{k2}x_{k2} + \cdots + h_{kN_n}x_{kN_n} \\
&\quad \textbf{Subject to} \qquad\qquad (14)\\
&\quad x_{k1} + x_{k2} + \cdots + x_{kN_n} \leq v_k \qquad \text{where, } \mathbf{y}_{kji} \text{ solves} \\
&\qquad \min_{\mathbf{y}_{kji}} N_j f_{kj}(\mathbf{y}_{kji}) \\
&\qquad \textbf{Subject to} \\
&\qquad \sum_{i=1}^{M_j} \left(y^1_{kji} + y^2_{kji} + y^3_{kji} + y^4_{kji}\right) \leq x_{kj}
\end{aligned}$$

where all the variables are nonnegative. *i.e.*, $v_k, x_{kj}, y^{\ell}_{kji} \geq 0$ for all the indices, and where all the inequalities hold for the corresponding indices as well.

Given a total budget B allocated for the global interventions in period τ, the upper level decision maker decides on the optimal distribution v_k of the budget for each region k. Then, considering this given value and the equity constraints, the second level decision maker (region k) decides in how to optimally distribute its available funding (v_k) for the intervention investments in each of the countries in its region and obtains the value x_{kj} for country j. The planners in country j are assumed to be responsible for the consideration of the total number of infections averted (and also in reducing the rate of failure in adherence to the proper use of HIV medication) as a criteria in their allocation of resources for various community groups within the country.

Accepting the values v_k and x_{kj} from the upper level planners, the planner in country j (the lower level decision maker) solves its resource allocation problem (10) and obtains an optimal vector $\mathbf{y}_{kj}(v_k, x_{kj})$, the exact value of which may depend on the choices of v_k and x_{kj} by the upper level decision makers.

Therefore, the optimization problem solved at each stage is a (multi-)parametric problem, where the parameters are decision variables at other decision levels. Moreover, the decision process is hierarchical as the global planner should declare its optimal allocation of funds first for the middle level to act. However, a one level solution procedures may not help to arrive at the optimal solution of such kind of problems as the parameters need to be rechecked again for their agreement with the optimality requirements at each other levels. That means, we need to apply solution techniques for multilevel programming problems to eventually arrive at the optimal allocation of resources for HIV interventions. The techniques that we may choose to solve the problem, however, depend on the types of the production functions for investment of each of the 4 types of the interventions.

Since all the constraints in problem (14) are linear inequalities in investment variables at all levels, the constraint set is polyhedral. But since the objective function of the lower level problem is non-convex, even when all the production functions are chosen to be linear, unique solutions may not be obtained at lower levels and we do not also expect that only full utilization of funds give optimal solutions. Proper definition of the production functions for each of the intervention types requires a closer study on the nature of the investment and its impact on the corresponding parameter values. Although linear functions are easy to work with, such linear approximations may not take important behaviors of the parameters into consideration. Therefore, it is advisable to use a combination of models depending on the nature of interventions.

In some of the intervention programs each incremental unit of money spent may produce the same incremental reduction or upsurge in epidemic component. In such type of programs the production function could be taken to be linear as a function of investment variables [2, 3]. In our model, for example, to formulate the proportion of investment in increasing the rate of recruitment for treatment (ρ) we may use such linear production function. i.e, we may use $\phi_{\rho_i}(y) = a_i + b_i y$ for some constants a_i, b_i. Similarly, the production function to model the proportion of investment in decreasing the rate m of MTCT can also be formulated as $\phi_{m_i}(y) = \alpha_i - \beta_i y$, for some constants α_i, β_i.

On the other hand, production functions that are related to a change in the risky behavior of group of individuals, an incremental investment in such programs may not usually produce the same incremental reduction in the risky behavior of individuals [12]. This is because, as the program expands the individuals reached are increasingly less likely to change their risky behavior. In such cases we may better model the corresponding production function by a function which is convex with respect to the investment variables. In our model, the investment in reducing the rate of infection due to unsafe sexual contacts (λ) and the rate of failure in adherence to the proper medication of ARV (α) require efforts in convincing individuals to change their risky behavior. Therefore, following the model in [3] one may define function of the form

$$\phi_\lambda(y) = \begin{cases} \lambda_{\min}, & \text{if } 0 \le y \le F; \\ \lambda_{\min}\left(a + be^{c(F-y)}\right), & \text{if } y > F, \end{cases}$$

where $\lambda_{\min}$ represents the minimum possible value of the parameter to be decreased, $a, b \in (0, 1]$ with $a + b = 1$, c can be derived from the cost-effectiveness ratios of programs, whereas F is called the startup cost to exhibit reduction. In some cases non-convex and non-concave type production functions can be applied, which have a variable monotonicity of returns to scale. For instance, a production function can be modeled to have a functional structure that, for low level of investment it could be assumed to have an increasing returns to scale (convex like) and then as the program extends and when large enough investment is used for the program it may have a decreasing returns to scale (concave like).

Once the model is properly formulated the next step will be to choose the best solution procedure that can give us an optimal solution. Since, all the constraint sets are polyhedral in investment variables, the solution approach depends on the type of definition of the production functions which appear in the objective function of the lower level problem. If all the production functions are twice continuously differentiable functions, $f_{kj}(\mathbf{y}_{kji})$ in Equation (10) is also continuously differentiable with respect to the investment vector $\mathbf{y}_{kji}$. Hence, in such cases we may apply multi-parametric programming algorithms (for instance, [9, 10]) together with numerical integrations of the lower level objective function. However, if any of the production functions are not smooth enough or if it is costly to calculate the derivatives of their combinations in the objective functions of the lower level problem, then it will be necessary to apply some derivative free or heuristic algorithms (like the method in [29]) to solve the problem.

4. Solution of the tri-level programming problem - *with a hypothetical example.*

To show the procedure for solving the global resource allocation problem using a tri-level hierarchical model, we formulated a hypothetical example. In this situation, assume there are 3 regions globally and a total of $B = \$\ 8.5$ billion budget is available for the disease control in the coming 10 years.[1] Then, this budget is to be distributed to these three regions as v_1, v_2, v_3, where $v_1 + v_2 + v_3 \le B$. Again to simplify the process, we also assumed that there are only 2 countries in each region. So, for each of the countries in any of the regions, say region k, x_{k1} and x_{k2} amount is to be allocated, where $x_{k1} + x_{k2} \le v_k$, for each $k = 1, 2, 3$. Moreover, within each country it is assumed that there are two community (or risk) groups, to which a budget (or resource) of y_{kj1} and y_{kj2} is to be allocated to fight the epidemic in their community groups ($k = 1, 2, 3,\ j = 1, 2$). With these assumptions, totally there are 12 community groups for global allocation of resources. Each of the allocated budget for a single community group has four components, $(y^1_{kji}, y^2_{kji}, y^3_{kji}, y^4_{kji}) = y^\ell_{kji}$, corresponding to each of the intervention categories.

[1] In this example, the allocated budget B does not include the administrative expenses at each level of coordination. The entire resource is assumed to be applied directly for the epidemic control.

For the formulation of the upper level (global) and the middle level (regional) resource allocation problems, the following parameters values are used.

$$h_{kj} = \begin{pmatrix} 0.2 & 0.3 \\ 0.31 & 0.39 \\ 0.25 & 0.4 \end{pmatrix}, \qquad d_{kj} = \begin{pmatrix} 20\% & 25\% \\ 30\% & 15\% \\ 25\% & 20\% \end{pmatrix},$$

where, the rows represent the regions and the columns represent the countries. The entries in the matrix d_{kj} give the minimum percentage of the region's budget that should be allocated to a country in that region. Moreover, assume that the minimum percentage of the total allocation for any one of the intervention categories is $z^{\ell}_{kji} = 7.5\%$ of the region's budget, and it is assumed to be the same for all the indices.

In the lower level problem there are various parameters to be used. Some of them are disease parameters that remain constant across the glob and some others are country or risk group specific. They are summarized in Tables 1 and 2. (Some of these parameters are from [5, 13, 19, 23, 26, 28] and the rest are estimated.)

Parameter value	Description
$\delta_V = 0.15$	mortality rate for infected children
$\delta_A = 0.2$	additional death rate due to AIDS
$\sigma = 0.1$	rate of progression to AIDS, if not treated
$\gamma = 1.3$	preferential rate of recruitment for children to receive ART
$\beta = 0.12$	transmission probability
$\epsilon_1 = 0.08$	factor of reduction on rate of disease transmission due to ART
$\epsilon_2 = 1/6$	factor of reduction on rate of MTCT due to treatment

Table 1. Disease parameters – Assumed to be Constant across regions and countries

Parameter values	Description (each is for the 6 countries)
$\pi = [0.032, 0.036, 0.028, 0.021, 0.015, 0.012]$	Birth rates for the six countries, respectively
$\mu = [0.029, 0.025, 0.025, 0.019, 0.013, 0.010]$	Death rates for the six countries, respectively
$\lambda_H = [0.16, 0.31, 0.13, 0.12, 0.12, 0.11]$	Initial unsafe contact rates for high risk groups
$\lambda_L = [0.11, 0.14, 0.09, 0.075, 0.07, 0.05]$	Initial unsafe contact rates for low risk groups
$m_H = [0.31, 0.37, 0.28, 0.25, 0.22, 0.20]$	Initial rate of MTCT for high risk groups
$m_L = [0.19, 0.22, 0.15, 0.14, 0.12, 0.11]$	Initial rate of MTCT for low risk groups
$\alpha_A = [0.15, 0.32, 0.11, 0.12, 0.11, 0.10]$	Initial rate of defaulting in the use of ART – assumed to be the same for both risk groups in each of the countries
$\rho_H = [0.15, 0.15, 0.20, 0.21, 0.23, 0.24]$	Initial rate of recruitment for ART in High risk groups
$\rho_L = [0.075, 0.075, 0.10, 0.11, 0.15, 0.15]$	Initial rate of recruitment for ART in Low risk groups

Table 2. Parameters that are assumed to vary from region to region or from risk groups to risk groups

Initial size for the Low Risk population group in the 6 countries is assumed to be as in the table below,

Pop. Size	R1C1	R1C2	R2C1	R2C2	R3C1	R3C2
S(0)	15652504	870885.6	35863200	7341600	28122000	16815000
V(0)	87640	12854.4	652800	151200	548250	285000
I(0)	1016624	107120	1632000	361200	1322250	627000
T(0)	438200	47132.8	1836000	411600	1451250	855000
A(0)	333032	33207.2	816000	134400	806250	418000

,

and the initial size for High Risk population group in the 6 countries is taken to be as in the following table.

Pop. Size	R1C1	R1C2	R2C1	R2C2	R3C1	R3C2
S(0)	11706200	611078.4	13785600	2656800	7998000	4512000
V(0)	110176	25708.8	499200	57600	193500	90000
I(0)	1170620	208636.8	2515200	417600	1236250	666000
T(0)	454476	84048	1632000	342000	1053500	600000
A(0)	330528	59328	768000	126000	268750	132000

The production functions for the model are chosen to be

$$\begin{aligned} \phi_{m_i}(y^1) &= m(1+a_2y^1), \qquad \phi_{\lambda_i}(y^2)=\lambda(1+e^{-(c_1y^2)}) \\ \phi_{\rho_i}(y^3) &= \rho(1+a_1y^3), \qquad \phi_{\alpha_i}(y^4)=\alpha(1+e^{-(c_2y^4)}), \end{aligned}$$

with $a_1 = 1.65\times10^{-8}$, $a_2 = 5.43\times10^{-8}$, $c_1 = 4.56\times10^{-1.55}$, and $c_2 = 3.35\times10^{-2}$ (partly taken from [3]). These functions and parameter values are assumed to be the same for all regions and countries. Moreover, since the planning time is considered to be 10 years, the discounting rate r is assumed to be 5%.

Then, given the leader's variable v_k for region k and the investment variable x_{kj} allocated by region k for country j, the decision maker at each of the country levels optimizes its resource allocation problem for the desired epidemic control. This will lead to a multi-follower problem in hierarchical optimization. However, since all the countries are at the same level and they optimize their resource allocation problems independently with their objective functions separable, we can use the equivalent formulation as the sum of all the terms at the same level (see [11] for the details of this approach). Hence, the objective function of the lower level optimization problem is the sum of the costs of all the 6 countries, and the optimization is over all the y-variables. In general, the lower level problem is a 48-variables problem with integral form objective function. Using similar argument, even if the regions operate concurrently and independently, we optimize the sum of all the objectives in the three regions for all the variables x_{kj} at the middle level.

In choosing a solution approach, it is considered that since the lower level objective of the problem is defined in terms of an integral form, it will be expensive to calculate the first and second order derivatives. Hence, the algorithm in [29] is used to get the solution. After reformulation and running the SEAMSP algorithm [29] for the above tri-level problem, we obtain the solutions as in Table 3. The values in this table are given in millions of USD.

When these optimal investment values are used in the dynamical system it can be seen from the simulation graphs (Fig. 4 - 6) that the prevalence of the disease decreases significantly in each of the countries in the region.

5. **Discussion.** In this paper a hierarchical model for global resource allocation of the funds that are raised to fight HIV/AIDS is formulated and analyzed. It has been shown that when the effect of treatment in the aversion of new infection is employed in the model, the objective function of the lower level decision making

	$v_1 = 2,056.9$		$v_2 = 2,461.6$		$v_3 = 3,452.1$	
	$x_{11} = 617.1$	$x_{12} = 1,439.7$	$x_{21} = 1,723$	$x_{22} = 738.5$	$x_{31} = 1,187.9$	$x_{32} = 2,264.1$
Low Risk						
Com. Grp						
y^1	125.46	234.63	310.64	107.62	172.46	423.69
y^2	89.14	193.46	154.29	98.947	193.42	244.36
y^3	97.432	102.46	196.98	69.442	207.16	219.80
y^4	49.864	299.87	130.26	124.67	154.32	501.64
High Risk						
Com. Grp						
y^1	28.808	125.01	120.09	77.004	124.50	142.29
y^2	65.128	166.29	276.41	85.677	103.49	321.64
y^3	56.836	257.30	233.75	115.18	89.809	346.20
y^4	104.41	60.049	300.46	59.954	142.64	64.384

TABLE 3. Solution of the three level resource allocation

structure is neither convex nor concave in the allocation variables even if all the production functions are assumed to be linear. This made the resulting multilevel optimization problem difficult to solve with classical operations research methods.

Here, it is assumed that there are three levels of decisions in the global allocation of resources. However, a two level decision system can also be analyzed using a similar approach by simply merging the middle level with the upper level decision making structure. Since the main difficulty (in terms of solution approaches) arises due to the properties of the lower level objective function, the two level version of

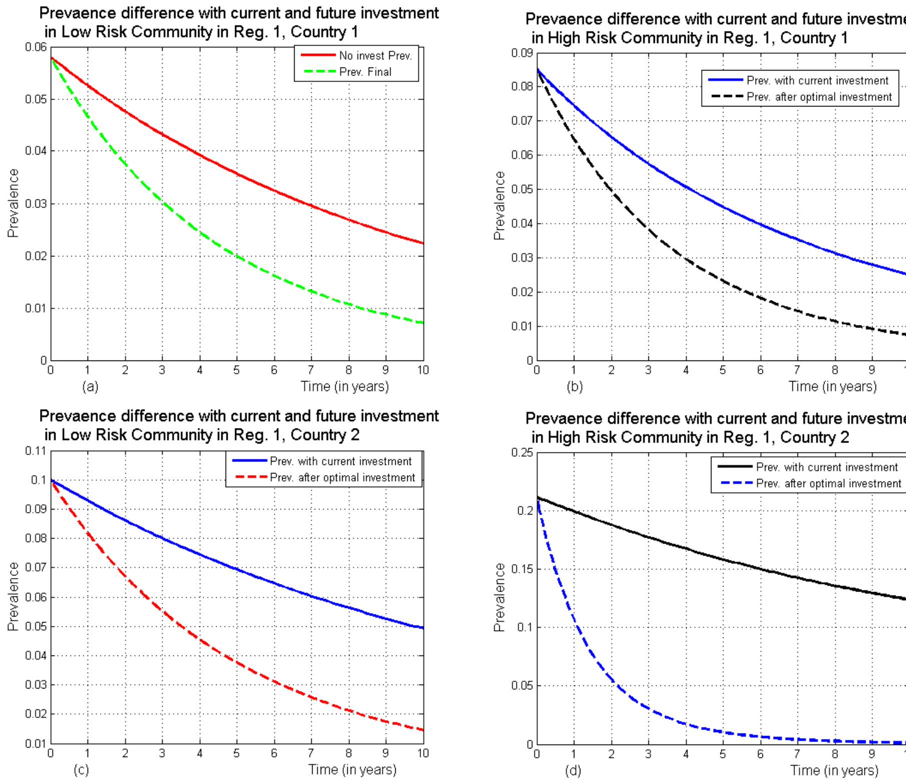

FIGURE 4. Prevalence graphs for Countries in Region 1. The figures to the left indicate the prevalence for Low Risk population groups while those in the right indicate the prevalence for High Risk population groups with in the same country. The broken lines indicate the prevalence if the optimal resource is invested in the planned 10 years period for each of the community groups

the problem also requires the same solution techniques recommended in this paper. On the other hand if the number of levels (or hierarchies) is increased from 3 to any possible number in the formulation of the model, the analysis of the model again follows a similar argument and structure. The solution techniques of any such multilevel problems also depend on the type of constraints and criteria functions formulated at each level of decision. The effectiveness of this model is not yet proved by taking actual data from the field. Nonetheless, unlike the conclusion in [18] (which recommends the use of equity – optimal approach), the author recommends the use of optimal – optimal approach over the other possible decision approaches between the levels as equity parameters are already included in the model constraints.

This study considered the allocation of resources to control the HIV epidemic in a multiple but independent communities of the population in any country. Though the model relies upon the epidemic behavior of HIV/AIDS, the model structure and the analysis given in this paper are likely to be applied to other types of diseases provided some of the interventions also result in reducing the infectiousness of the infected individuals. Moreover, one can also use a similar model to allocate resources

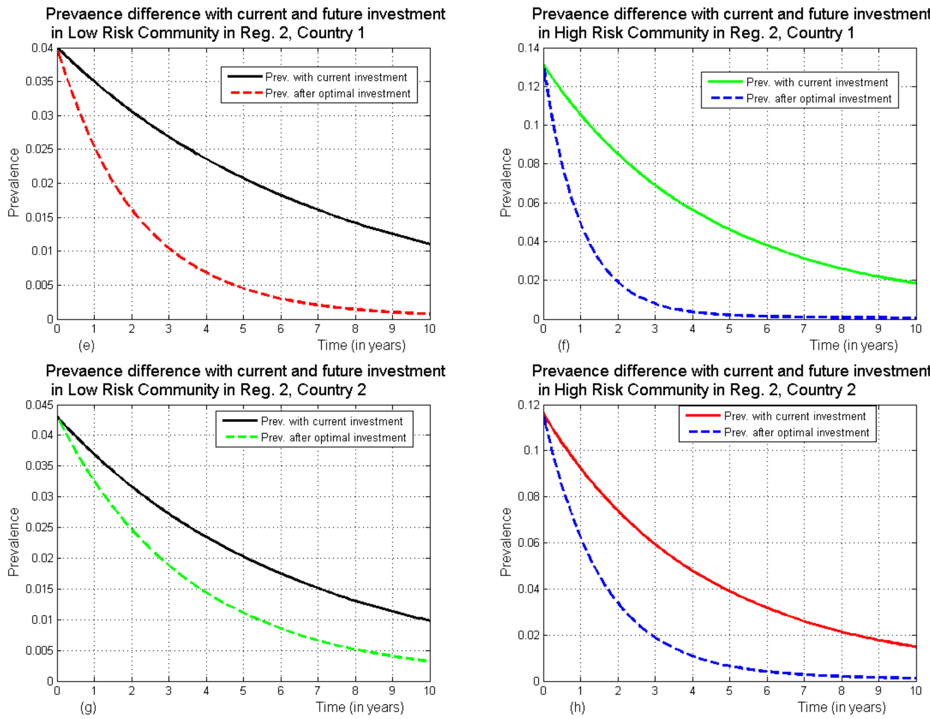

FIGURE 5. Prevalence graphs for Countries in Region 2. The broken lines indicate the prevalence if the optimal resource is invested in the planned 10 years period for each of the community groups

to control multiple diseases in the same population (like, for example, Tuberculosis (TB) and Malaria) if the epidemics are assumed to be independent. But since some infections are not independent, one need to apply a different modeling structure for non-independent diseases (like, for example, HIV and TB, HIV and other sexually transmitted diseases) because the same intervention may have an impact on the incidence of the other diseases as well.

In this paper it is also assumed that countries within a region, or regions within the global settings do not compete each other to receive higher resources from higher level resource allocation body for any of the intervention programs. If the resource allocation structure allows for such a competition, then one has to apply a multi-level multi-follower method to get an optimal solution. However, such techniques are not well developed yet to handle various formulations of lower level problems. Therefore, this requires a further research especially when the objective functions of the lower level decision makers (in our case the country level resource allocation structure) need to solve non-convex and non-concave problems with resources are shared among them.

Acknowledgments. This work is partly supported by the Swedish International Science Program (ISP), through the project at the Department of Mathematics, Addis Ababa University.

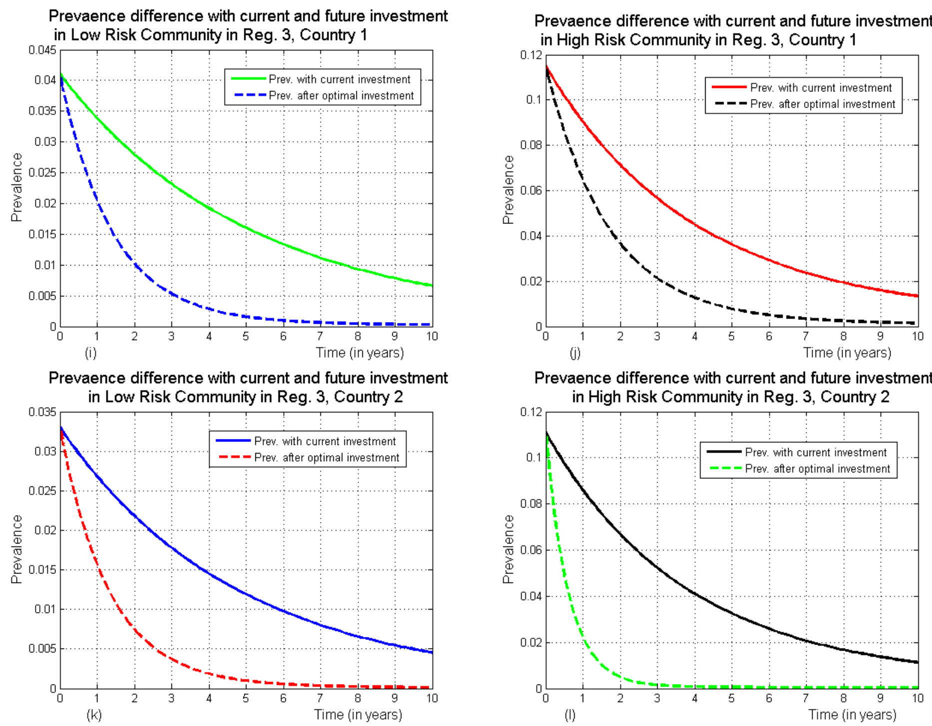

FIGURE 6. Prevalence graphs for Countries in Region 3. The broken lines indicate the prevalence if the optimal resource is invested in the planned 10 years period for each of the community groups

REFERENCES

[1] Avert, *Funding for HIV and AIDS*, 2016. Available from http://www.avert.org/node/353/pdf

[2] M. L. Brandeau, G. S. Zaric and A. Richter, Resource allocation for control of infectious diseases in multiple independent populations: Beyond cost-effectiveness analysis, *Journal of Health Economics*, **22** (2003), 575–598.

[3] M. L. Brandeau, G. S. Zaric and V. De Angelis, Improved allocation of HIV prevention resources: Using information about prevention program production functions, *Health Care Management Science,* **8** (2005), 19–28.

[4] Centers for Disease Control and Prevention (CDC), Achievements in public health, reduction in perinatal transmission of HIV infection – United States, 1985 - 2005, *MMWR Morb Mortal Wkly Rep*, **55** (2006), 592–597.

[5] D. Donnell, J. M. Baeten, J. Kiarie, K. K. Thomas, W. Stevens, C. R. Cohen, J. Mclntyre, J. R. Lingappa and C. Celum, Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: A prospective cohort analysis, *Lancet,* **375** (2010), 2092–2098.

[6] M. Drummond, B. O'Brien, G. L. Stoddart and G. J. Torrance (Eds.), *Methods for the Economic Evaluation of Health Care Programs*, Oxford University Press, New York, 2000.

[7] S. R. Earnshaw, K. Hicks, A. Richter and A. Honeycutt, A linear programming model for allocating HIV prevention funds with state agencies: A pilot study, *Health Care Manage Sci*, **10** (2007), 239–252.

[8] S. Flessa, Where efficiency saves lives: A linear programme for the optimal allocation of health care resources in developing countries, *Health Care Management Science,* **3** (2000), 249–267.

[9] A. M. Kassa and S. M. Kassa, A multi-parametric programming algorithm for special classes of non-convex multilevel optimization problems, *An International Journal of Optimization and Control: Theories & Applications (IJOCTA),* **3** (2013), 133–144.

[10] A. M. Kassa and S. M. Kassa, A branch-and-bound multi-parametric programming approach for non-convex multilevel optimization with polyhedral constraints, *Journal of Global Optimization*, **64** (2016), 745–764.

[11] A. M. Kassa and S. M. Kassa, Deterministic solution approach for some classes of nonlinear multilevel programs with multiple followers, *J Glob Optim,* (2017), 1–19.

[12] S. M. Kassa and A. Ouhinou, The impact of self-protective measures in the optimal interventions for controlling infectious diseases of human population, *Journal of Mathematical Biology*, **70** (2015), 213–236.

[13] S. M. Kassa and A. Ouhinou, Epidemiological models with prevalence dependent endogenous self-protection measure, *Mathematical Biosciences,* **229** (2011), 41–49.

[14] J. Kates, J. A. Izazola and E. Lief, Financing the response to HIV in low- and middle-income countries:, International assistance from donor governments in 2015, 2015. Available from: http://files.kff.org/attachment/Financing-the-Response-to-HIV-in-Low-and-Middle-Income-Countries-International-Assistance-from-Donor-Governments-in-2015

[15] J. Kates, A. Wexler and E. Lief, Financing the response to HIV in low- and middle-income countries: International assistance from donor governments in 2013, UNAIDS Report, July 2014. Available from: https://kaiserfamilyfoundation.files.wordpress.com/2014/07/7347-10-financing-the-response-to-hiv-in-low-and-middle-income-countries.pdf

[16] J. M. Kilby, H. Y. Lee, J. D. Hazelwood, A. Bansal, R. P. Bucy, M. S. Saag, G. M. Shaw, E. P. Acosta, V. A. Johnson, A. S. Perelson and P. A. Goepfert, Treatment response in acute/early infection versus advanced AIDS: Equivalent first and second phase of HIV RNA decline, *AIDS,* **22** (2008), 957–962.

[17] A. P. Kourtis, C. H. Schmid, D. J. Jamieson and J. Lau, Use of Antiretroviral therapy in HIV-infected pregnant women and the risk of premature delivery: A meta-analysis, *AIDS,* **21** (2007), 607–615.

[18] A. Lasry, G. S. Zaric and M. W. Carter, Multi-level resource allocation for HIV prevention: A model for developing countries, *European Journal of Operational Research,* **180** (2007), 786–799.

[19] F. J. Palella, K. M. Delaney, A. C. Moorman, M. O. Loveless, J. Fuhrer and G. A. Satten *et al.*, Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection, *The New England Journal of Medicine,* **338** (1998), 853–860.

[20] L. Palombi, M. C. Marazzi, A. Voetberg and N. A. Magid, Treatment acceleration program and the experience of the DREAM program in prevention of mother-to-child transmission of HIV, *AIDS,* **21** (2007), S65–S71.

[21] A. Prendergast, G. Tudor-Williams, S. Burchett and P. Goulder, International perspectives, progress, and future challenges of paediatric HIV infection, *Lancet,* **370** (2007), 68–80.

[22] N. Siegfried, L. van der Merwe, P. Brocklehurst and T. T. Sint, Antiretrovirals for reducing the risk of mother-to-child transmission of HIV infection, *Cochrane Database of Systematic Reviews 2011,* **7** (2011), Art.CD003510.

[23] J. A. C. Sterne, M. A. Henán, B. Ledergerber, K. Tilling, R. Weber and P. Sendi et al., Long-term effectiveness of potent antiretroviral therapy in preventing AIDS and death: *A prospective cohort study*, *Lancet,* **366** (2005), 378–384.

[24] UNAIDS, AIDS by the numbers, 2016. Available from: http://www.unaids.org/sites/default/files/media_asset/AIDS-by-the-numbers-2016_en.pdf.

[25] UNAIDS, Fast-track update on investments needed in the AIDS response, UNAIDS Reference, 2016. Available from: http://www.unaids.org/sites/default/files/media_asset/UNAIDS_Reference_FastTrack_Update_on_investments_en.pdf

[26] R. Vardavas and S. Blower, The emergence of HIV transmitted resistance in Botswana: When will the WHO detection threshold be exceeded?, *PLoS ONE,* **2** (2007), e152.

[27] M. C. Weinstein, From cos-effectiveness ratios to resource allocation: where to draw the line?, in *Valuing Healthcare: Costs, Benefits, Effectiveness of phramaceuticals and other medical technologies (eds. F. A. Sloan),* Cambridge University Press, New York (1995), 77–97.

[28] World Health Organization, *Towards Universal Access: Scaling up Priority HIV/AIDS Interventions in the Health Sector: Progress Report 2009*, WHO, 2009.

[29] A. T. Woldemariam and S. M. Kassa, Systematic evolutionary algorithm for general multilevel Stackelberg problems with bounded decision variables (SEAMSP), *Annals of Operations Research,* **229** (2015), 771–790.

[30] G. S. Zaric and M. L. Brandeau, Resource allocation for epidemic control over short time horizons, *Mathematical Biosciences,* **171** (2001), 33–58.

[31] G. S. Zaric and M. L. Brandeau, A little planning goes a long way: Multilevel allocation of HIV prevention resources, *Medical Decision Making,* **27** (2007), 71–81.

Received August 24, 2016; Accepted March 3, 2017.

E-mail address: semu.mtk@gmail.com
E-mail address: kassas@biust.ac.bw