International Journal of Drug Policy 38 (2016) 73–80

Contents lists available at ScienceDirect

International Journal of Drug Policy

journal homepage: www.elsevier.com/locate/drugpo

Editors' Choice

Allocative and implementation efficiency in HIV prevention and treatment for people who inject drugs

Clemens Benedikt[a,*], Sherrie L. Kelly[b], David Wilson[a], David P. Wilson[b], on behalf of the Optima Consortium

[a] World Bank, Washington, DC, USA
[b] Burnet Institute, Melbourne, Australia

ARTICLE INFO

Article history:
Received 19 July 2016
Received in revised form 28 September 2016
Accepted 19 October 2016

Keywords:
HIV
Allocative efficiency
Implementation efficiency
Mathematical modelling
Optimization
Harm reduction

ABSTRACT

Background: Estimated global new HIV infections among people who inject drugs (PWID) remained stable over the 2010–2015 period and the target of a 50% reduction over this period was missed. To achieve the 2020 UNAIDS target of reducing adult HIV infections by 75% compared to 2010, accelerated action in scaling up HIV programs for PWID is required. In a context of diminishing external support to HIV programs in countries where most HIV-affected PWID live, it is essential that available resources are allocated and used as efficiently as possible.

Methods: Allocative and implementation efficiency analysis methods were applied. Optima, a dynamic, population-based HIV model with an integrated program and economic analysis framework was applied in eight countries in Eastern Europe and Central Asia (EECA). Mathematical analyses established optimized allocations of resources. An implementation efficiency analysis focused on examining technical efficiency, unit costs, and heterogeneity of service delivery models and practices.

Results: Findings from the latest reported data revealed that countries allocated between 4% (Bulgaria) and 40% (Georgia) of total HIV resources to programs targeting PWID — with a median of 13% for the eight countries. When distributing the same amount of HIV funding optimally, between 9% and 25% of available HIV resources would be allocated to PWID programs with a median allocation of 16% and, in addition, antiretroviral therapy would be scaled up including for PWID. As a result of optimized allocations, new HIV infections are projected to decline by 3–28% and AIDS-related deaths by 7–53% in the eight countries. Implementation efficiencies identified involve potential reductions in drug procurement costs, service delivery models, and practices and scale of service delivery influencing cost and outcome. A high level of implementation efficiency was associated with high volumes of PWID clients accessing a drug harm reduction facility.

Conclusion: A combination of optimized allocation of resources, improved implementation efficiency and increased investment of non-HIV resources is required to enhance coverage and improve outcomes of programs for PWID. Increasing efficiency of HIV programs for PWID is a key step towards avoiding implicit rationing and ensuring transparent allocation of resources where and how they would have the largest impact on the health of PWID, and thereby ensuring that funding spent on PWID becomes a global best buy in public health.

Introduction

Current HIV programs are faced with the necessity to scale-up prevention while also providing treatment to a larger number of people living with HIV (PLHIV) than ever before. In the 2011 United Nations Political Declaration on HIV and AIDS, countries agreed to reduce sexual and injection-related HIV transmission by 50% by 2015 (UNGASS, 2011). Estimated new HIV infections among PWID remained stable at around 140,000 per year over the 2010–15 period and the 2015 target of a 50% reduction was missed (UNAIDS, 2016). This suggests that there is need for accelerated action if 2020 targets of reducing new HIV infections and deaths by 75% compared to 2010 levels are to be achieved (UNAIDS, 2014a). At the same time, after a decade of rapid growth, international HIV financing stabilized around 2010 (UNAIDS, 2015) and is projected to decline in middle-income countries (Cook,

* Corresponding author.
E-mail address: cbenedikt@worldbank.org (C. Benedikt).

http://dx.doi.org/10.1016/j.drugpo.2016.10.011
0955-3959/

Bridge, McLean, Phelan, & Barrett, 2014), in which a large proportion of HIV-affected PWID live. Programs for PWID have been supported by the Global Fund and other international partners and have been particularly dependent on external funding in many countries.

In this environment, there are two main options for how global HIV impact targets can be achieved: (1) increased domestic financing of HIV programs for PWID and (2) greater efficiency in program design and delivery to ensure that programs can do more with what is available. Previous research on cost and impact of HIV programs for PWID has focused on estimating cost-effectiveness of harm reduction (Wilson, Donald, Shattock, Wilson, & Fraser-Hur, 2015) overall or for specific interventions such as opioid substitution therapy (OST) and needle-syringe programs (NSP) (Cipriano et al., 2012; Kim, Pulkki-Brannstrom, & Skordis-Worrall, 2014). Less attention has been paid to the question of how to make HIV programs for PWID more efficient, as concepts of allocative and implementation efficiency have not been widely applied to programs for PWID. Given the large gaps in coverage of HIV prevention and treatment programs for PWID at a time of limited resources, the question of efficiency will become essential not only for impact, but also sustainability of the response.

The concept of allocative efficiency refers to the maximization of health outcomes using the least costly mix of health interventions. HIV allocative efficiency analysis addresses the question "How can HIV funding be optimally allocated to the combination of HIV response interventions that will yield the highest impact?" Technically, allocative efficiency can be accomplished either within a fixed budget envelope to achieve maximal impact with given amount of money or within defined impact targets to achieve a given impact with minimal cost. In both cases, allocative efficiency is achieved by optimizing the *mix* of interventions to achieve specific impact-level goals. Implementation efficiency can be defined as a set of measures to ensure that programs are delivered in a way that achieves outputs with the lowest input of resources.

Methods

In this paper, we synthesize evidence in relation to HIV prevention and treatment programs for PWID from published and unpublished studies conducted by the Optima Consortium for Decision Science. This consortium of researchers and public health experts collaborate by using data on epidemics and health responses including through mathematical modelling in order to facilitate decision-making in the health sector. The results presented here were generated by applying Optima-HIV, the most widely used model within the Optima suite. Other applications of Optima are currently being used and developed for nutrition, child health, tuberculosis, malaria, hepatitis C, and non-communicable diseases.

Allocative efficiency analysis

In the allocative efficiency analyses presented in this paper, impact goals have been defined as minimizing new HIV infections and AIDS-related deaths. We have conducted allocative efficiency analyses in numerous country contexts in which drug injecting behaviours are common. For reasons of consistency and comparability, we have focused on findings from eight countries in Eastern Europe and Central Asia.

Application of the Optima model

Optima is a dynamic, population-based HIV model with an integrated program and economic analysis framework. In the Optima-HIV epidemic model the total population was partitioned in two ways: by population group and by HIV health state including by whether an individual is diagnosed with HIV and by levels of CD4 count. All individuals for a given country were assigned to a population group based on their dominant risk, for example, injecting drug practices or paid sex. New HIV infections occur through interactions among different populations from regular, casual, or commercial sexual partnerships, through sharing of injecting equipment; or through mother-to-child transmission. In the present analysis, male and female PWID were tracked in relation to all three modes of transmission and across the different HIV related health states including five clinical categories (uninfected, undiagnosed, diagnosed, on ART with unsuppressed virus, and on ART with suppressed virus) and six disease states (primary HIV infection, $CD4 > 500$ cells/mm^3, $350 < CD4 < 500$, $200 < CD4 < 350$, $50 < CD4 < 200$, and $CD4 < 50$). The model uses a linked system of ordinary differential equations to track the movement of PLHIV among HIV health states.

The key assumptions of resource optimization are the relationships among (1) the cost of HIV programs for people who inject drugs and other populations, (2) the resulting coverage levels of populations reached with these HIV programs, and (3) how these coverage levels of HIV programs for priority populations influence behavioural and clinical outcomes. Cost assumptions were based on expenditure data collected by in-country experts in the context of global AIDS progress reporting using the National AIDS Spending Assessment (NASA) definitions of cost categories. Coverage of programs was established using data from national databases and implementation program records. Data from integrated bio-behavioural and other population-based surveys were used to establish behavioural and epidemic trends. As the relationships between cost, coverage, and outcome used to generate future projections and optimizations were based on empirical data, it was implicitly assumed that program quality in translating investment into coverage and outcome was constant over time. To perform optimization analysis, Optima uses an adaptive stochastic descent algorithm to identify the optimal mix of programs within a very large number of possible combinations of programs (Kerr, Smolinski, Dura-Bernal, & Wilson, 2016). The optimization algorithm starts at the current allocation of resources and randomly changes allocations towards a specific HIV intervention. If these changes improve the outcome, the algorithm increases the step size of changes towards that program. This process is repeated until no further improvements are possible and the optimized mix of investments is identified.

The Optima-based studies summarized in this paper were partially conducted using Matlab and partially in a web-based user interface developed in Python. The full Optima model is described in detail in a methods paper with the set of ordinary differential equations used provided in the supplementary material (Kerr et al., 2015). The Optima model has also been compared to other models commonly applied for HIV epidemic and allocative efficiency analysis (Kahn, Bollinger, Stover, & Marseille, 2016; World Bank, 2016).

Allocative analysis efficiency beyond HIV

All HIV interventions have some direct or indirect non-HIV benefits. As these benefits extend into different areas including contraception for condom use and prevention of other infectious diseases for needle and syringe programs, secondary benefits of HIV programs on additional outcomes were not included within the optimization analysis. For opiate substitution therapy, a special approach was applied given the proven benefits on HIV and across different sectors (MacArthur et al., 2012). Such additional benefits were reflected using a cross-sectoral financing model to effectively