The global Optima HIV allocative efficiency model: targeting resources in efforts to end AIDS


Summary

Background To move towards ending AIDS by 2030, HIV resources should be allocated cost-effectively. We used the Optima HIV model to estimate how global HIV resources could be retargeted for greatest epidemiological effect and how many additional new infections could be averted by 2030.

Methods We collated standard data used in country modelling exercises (including demographic, epidemiological, behavioural, programmatic, and expenditure data from Jan 1, 2000, to Dec 31, 2015) for 44 countries, capturing 80% of people living with HIV worldwide. These data were used to parameterise separate subnational and national models within the Optima HIV framework. To estimate optimal resource allocation at subnational, national, regional, and global levels, we used an adaptive stochastic descent optimisation algorithm in combination with the epidemic models and cost functions for each programme in each country. Optimal allocation analyses were done with international HIV funds remaining the same to each country and by redistributing these funds between countries.

Findings Without additional funding, if countries were to optimally allocate their HIV resources from 2016 to 2030, we estimate that an additional 7·4 million (uncertainty range 3·9 million–14·0 million) new infections could be averted, representing a 26% (uncertainty range 13–50%) incidence reduction. Redistribution of international funds between countries could avert a further 1·9 million infections, a 33% (uncertainty range 20–58%) incidence reduction overall. To reduce HIV incidence by 90% relative to 2010, we estimate that more than a three-fold increase of current annual funds will be necessary until 2030. The most common priorities for optimal resource reallocation are to scale up treatment and prevention programmes targeting key populations at greatest risk in each setting. Prioritisation of other HIV programmes depends on the epidemiology and cost-effectiveness of service delivery in each setting as well as resource availability.

Interpretation Further reductions in global HIV incidence are possible through improved targeting of international and national HIV resources.

Funding World Bank and Australian NHMRC.

Introduction

The global community is committed to reducing new HIV infections by 90% from 2010 levels by 2030 to end the AIDS epidemic as a public health threat.\(^1,2\) To help reach this goal, UNAIDS has set ambitious diagnosis, treatment, and viral suppression targets supplemented with high coverage of prevention as a roadmap to achieving this goal. However, because international funding decreased by 7% in 2016,\(^3\) national governments are being urged to mobilise new domestic HIV resources to cover the billions of dollars in additional funds anticipated to be needed to achieve these targets.\(^4\)

As part of their Investment Framework for the Global HIV Response,\(^5\) UNAIDS directs national governments to invest strategically in HIV programmes. Siapka and colleagues\(^6\) did a systematic review of cost-effectiveness of the six most essential HIV programmes included in this framework and showed that further evidence was needed to better understand how to best achieve efficiency gains in HIV programmes. One type of gain is known as allocative efficiency, whereby funding is allocated across a mix of interventions in the right combination to yield the greatest health outcomes. The objective of this modelling study is to estimate how to minimise the number of HIV infections by 2030 by targeting global resources to the most cost-effective combination of interventions and locations worldwide.

International funding organisations, including The Global Fund to Fight AIDS, Tuberculosis and Malaria, now require applicants to provide evidence that their proposed budget will be invested cost-effectively. Many countries have used HIV modelling tools such as Goals,\(^7\) the AIDS Epidemic Model,\(^8\) and Optima HIV\(^9\) to assist in developing their investment strategy. Since 2011, more than 40 national governments have requested Optima HIV modelling analysis support,\(^10,11\) led by the World Bank, UN agencies, or the US Centers for Disease Control and Prevention, to improve the allocative efficiency of their HIV responses. This study incorporates expansions of previously generated Optima HIV country models as well as the generation of new subnational and national models (55 Optima HIV models in total) to...
represent more than 80% of the population living with HIV worldwide. We combined subnational and national models to generate a global Optima HIV model. We then projected the potential epidemiological gains that could be achieved through the most cost-effective investment in HIV programmes to minimise new HIV infections by 2030 at the subnational, national, regional, and global level if international funding remained the same or if these funds were redistributed across countries.

Methods

Model design

Using Optima HIV version 2.3.6, which is available for use without restriction, we generated a global HIV model to estimate the optimal resource allocation across HIV programmes to minimise new infections between Jan 1, 2016, and Dec 31, 2029 (ie, by 2030) at the subnational, national, regional, and global level. Optima HIV is a population-based compartmental model that uses a linked system of ordinary difference equations to track the transmission of HIV within and between context-specific population groups (appendix p 2). We used demographic, epidemiological, and behavioural data by population group, along with expenditure and coverage levels of HIV programmes from Jan 1, 2000, to Dec 31, 2015, to inform the model. The model captures elements such as sexual and injecting risk behaviour and mother-to-child transmission, and it dynamically tracks people as they move across clinical categories, disease states, and age and risk groups. Following the 2016 WHO guidelines, we specified in the model that all people living with HIV are eligible for treatment irrespective of CD4 T-cell count.

We selected 44 countries with the greatest numbers of people living with HIV in their respective regions so as to include 80% of the population living with HIV worldwide (an estimated 29.5 million of the 36.7 million people; appendix p 2). Regional representation of people living with HIV varied on the basis of data availability (86% in Asia and the Pacific, 70% in eastern Europe and central Asia, 75% in Latin America and the Caribbean, 56% in the Middle East and north Africa, and 88% in sub-Saharan Africa). Countries with low prevalence of HIV infections (including all upper-middle-income and high-income countries) were not included in this study.

Data sources

We generated one Optima HIV modelling file for each country and subnational models for countries with
particularly heterogeneous epidemics and for which data were readily available. Country Optima HIV models that were previously generated to inform national strategic planning or funding proposals with country governments were also included. Data to inform these models, originally collated and endorsed in partnership with country governments, were updated where possible. For newly created national and subnational models, data were collated from UNAIDS Global AIDS Monitoring and National AIDS Spending Assessment, US President's Emergency Plan for AIDS Relief (PEPFAR), Demographic and Health Surveys, and Integrated HIV Bio-behavioral Surveillance reports as well as from the Avenir Health Unit Cost, Vital Statistics, and the World Bank databases supplemented with data from context-specific sources such as national annual reports and strategic plans for HIV and AIDS. Input data and assumptions for Optima HIV country models are available upon request, with sharing of non-public, country-owned data subject to approval from the respective country. Models were representative of national or subnational areas with generalised, concentrated, or mixed HIV epidemics and with diverse HIV spending patterns and responses. Regional projects are available within the online tool.

Model calibration and cost functions
We calibrated the epidemic model with UNAIDS11 or locally provided estimates (or both) for HIV prevalence per population and age group, number of people living with HIV, number of people receiving antiretroviral therapy (ART), incidence of HIV infections, and AIDS-related mortality (calibrations are shown in the appendix p 2).

Uncertainty estimates were generated around the model projections with an approximate Bayesian computation algorithm, with prior distributions defined for the prevalence of HIV infections in each population, transmission probabilities, and the key parameter values needed to define each projection. For cost functions, these parameters are the average cost of reaching someone with the programme at the current level of operations, the estimated maximal attainable coverage of the programme, and the programme effect in terms of behavioural or clinical outcomes. The cost function parameter values were allowed to vary uniformly over ranges within 10% of best assumptions. For each analysis, we calculated IQRs around the estimated cumulative number of infections and deaths expected from the model outputs on the basis of 100 simulations, with parameters sampled from the joint prior distributions (appendix pp 102–139), with cost function curves provided in the appendix (pp 140–198). We considered past expenditures for all services and components of the HIV response as representative of the costs needed to implement these responses in the future. The latest reported unit costs for each HIV programme were applied and did not vary over time. Estimated costs are reported in 2016 US$.

Optimisation algorithm
A unique feature of the Optima HIV model is its optimisation algorithm. Kerr and colleagues14 developed an adaptive stochastic descent algorithm to calculate the optimal resource allocation against defined constrained objective functions. The algorithm forms probabilistic assumptions about which parameters (changes in spending on programmes that affect changes in programmatic coverage levels, which influence prevention, treatment, and other outcomes) have the greatest effect on minimising new infections and uses optimal step sizes for each parameter. We used Monte Carlo initialisations for the optimisations to minimise the possibility of finding a local optimum. The default for optimisations is that they start ten times from the initial allocation and ten times from random allocations. We applied this algorithm to estimate the optimal allocation of HIV resources across available HIV interventions for every jurisdiction and across jurisdictions to minimise new infections from 2016 to 2030, compared with last reported budget allocations in each jurisdiction.

We differentiated between targeted and non-targeted HIV programmes (appendix p 4). Targeted programmes include treatment and prevention programmes with a clear, potential effect on reducing HIV transmission, morbidity, or mortality. Non-targeted programmes are those that might be essential in an HIV response but do not have a direct effect on health or cannot be attributed to population-specific outcomes. As non-targeted HIV programme expenditures do not have a direct effect on outcomes, they were considered to be fixed, remaining in the HIV response, but were not included in the optimisation. To reflect ethical treatment approaches, the optimisation was subject to constraints such that funding to treatment programmes (ART, opiate substitution therapy, and prevention of mother-to-child transmission) could not be decreased.

Optimisation analyses were based on reallocation of last reported HIV funding amounts to redistribute funding between programmes within each country, with international resources15 remaining the same to each country, or to redistribute funding within each country and redistribute international funding between countries. International funds were considered from funding organisations such as The Global Fund, PEPFAR, and bilateral and multilateral agencies, whereas domestic HIV resources were not redistributed between countries in our analyses.

Using different combinations of coverage for HIV programmes, the change in risk behaviour and morbidity and mortality outcomes were modelled. New HIV infections and AIDS-related deaths were projected to 2030 under different funding amounts and programmatic allocations across every possible combination of
allocation. The optimisation algorithm within the Optima HIV model was then used to estimate global minima for the optimal allocation of resources to minimise new HIV infections. Each programme had defined effectiveness assumptions with justifications from available international evidence (eg, ART was assumed to reduce CD4-stratified mortality and, when viral suppression is achieved, to reduce infectiousness by 96% [uncertainty range 73–99%]). Model output was aggregated from subnational projects to the national level, and national output was aggregated to the regional and then the global level. Our global optimisation analyses were assessed not only at currently available global HIV resources but for all levels of HIV funding varying from 0% to 200% of last reported spending levels in 20% increments. We measured the effect of these funding changes, with optimal resource allocation, on the cumulative new HIV infections and AIDS-related deaths by 2030.

Role of the funding source

The funders of the study had no role in study design. Staff at the World Bank contributed to data collation and writing of this Article. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

If the last reported HIV expenditure amounts and allocations are held constant from Jan 1, 2016, to Dec 31, 2029, we projected a gradual increase in new HIV infections worldwide (figure 1), thus moving further from the target of reducing infections by 90% relative to 2010 and of ending AIDS by 2030. However, if the same global budget is invested in the optimal mix of HIV programmes, we estimated that the annual number of new infections would decrease by by 26% (uncertainty range 13–50%) relative to 2010. To achieve these improved global health outcomes, the highest priority is to scale up ART funding from 40% of annual global HIV spending in 2015 (US$5·1 billion) to 48% (US$6·1 billion) until 2030 (figure 2). This would represent a cumulative shift of $14·5 billion in global HIV spending towards ART by 2030, effecting a 9% increase in coverage of people receiving ART. We also estimated that it is possible to reduce incidence by 90% globally by 2030, but an approximate increase to $40 billion in total annual funds will be needed for the countries modelled. Cost-effective allocation of resources is therefore important, as is exploration of other types of efficiencies, such as delivering high-quality HIV services at reduced costs and reducing commodity costs. In existing funding channels, the next priorities are to increase funding towards prevention of mother-to-child transmission and programmes targeting key populations, including people who inject drugs and men who have sex with men.

In sub-Saharan Africa, a region that accounts for nearly 70% of all people living with HIV and with mainly generalised or mixed HIV epidemics, we estimated that new HIV infections could be reduced by 23% (uncertainty range 3–59%) between 2016 and 2030 (equivalent to 4·6 million infections [670 000–11·6 million] averted if resources were optimally allocated; appendix p 7). With no increase in the annual funding available during this period, the incidence reductions would largely be accomplished by shifting about $550 million towards ART, thereby increasing total budget allotment on ART from 38% to 46% (figure 2). An optimised allocation for this region would also see investments shifting towards prevention of mother-to-child transmission (a 57% relative increase from about $340 million to $530 million) and key population prevention programmes, including programmes for female sex workers (124% relative programme budget increase from $26 million to $58 million) and, under the last reported budget, away from less contextually cost-effective programmes targeted at the general population, such as condoms and social behaviour communication change.

At the national level, the most common prioritisation of resource reallocation is towards ART, as shown in 33 (75%) of 44 countries modelled (figure 3). For 34 (77%) of 44 countries, the next priority is to scale up one or more prevention and testing programmes targeting key populations.

Another potential opportunity to consider for achieving further reductions in HIV infections and moving closer to global HIV targets is to evaluate the effect of redistributing HIV funds from international sources...
between countries. In this analysis, only international funds exceeding non-targeted and treatment programme (ART, prevention of mother-to-child transmission, and opiate substitution therapy constrained for ethical reasons) amounts were considered for redistribution between countries. International HIV funding accounted for $4·6 billion (36%) of the $12·8 billion budget for countries modelled, with $1·3 billion available for redistribution according to our imposed constraints for the analysis. The proportion of international to domestic funds varies widely between countries, with 64% of international funds invested in countries in sub-Saharan Africa.13 We found that an optimal redistribution of international funds between countries would see the largest share shifted to countries in sub-Saharan Africa, primarily to countries in central and western Africa, which otherwise do not receive equitable donor funding compared with eastern and southern African countries. Specifically, we found that country allocations are generally already well distributed, but our estimated optimal allocation had a very modest increase in funding for sub-Saharan Africa, from 55% ($7·0 billion) to 56% ($7·2 billion) of the total global budget. To a lesser extent, there were shifts towards countries in eastern Europe and central Asia and in the Middle East and north Africa (figure 2). In sub-Saharan Africa and in the Middle East and north Africa, the gains in international funds are prioritised towards scaling up ART, whereas in eastern Europe and central Asia, opiate substitution therapy, needle and syringe programmes, and prevention and testing programmes targeting sex workers are of highest priority. At the global level, we estimate that this shift could lead to an additional 7% incidence reduction compared with optimal allocation within countries alone, representing an overall 33% reduction in new HIV infections worldwide. This could avert an additional 9·3 million (uncertainty range 4·2 million–18·3 million) additional new infections compared with maintaining the latest reported allocation over this period, with a reduction to 1·3 million infections annually by 2030 (figure 1; appendix p 7). By better targeting HIV resources towards the most cost-effective mix of programmes in the right locations, it is therefore possible to substantially reduce global infections by 2030 without additional resources.

If total global HIV funding were increased, then it would become affordable to include several programmes in the most cost-effective programmatic mix that were not found to be part of the optimisation under the last reported budget. Were more funding to become available, there would be more opportunity to shift funds towards the next most cost-effective programme (or programmes), including HIV testing, and more towards programmes targeting key populations (ie, people who inject drugs, men who have sex with men, and female sex workers; figure 4A; appendix p 15). However, if global HIV resources were reduced by 20% (to 80% of the last reported amount), then several crucial programmes would fall out of the most cost-effectiveness mix.
Effective mix; however, as budget varies, so do the priorities for cost-effective allocation (appendix p 15). We estimated that if funding was reduced by 20% from 2016 to 2030, new infections would increase worldwide by 41% (uncertainty range 22–62%) or by 29.0 million cumulative new infections (23.6 million–35.1 million) compared with an estimated 20.6 million infections (17.0 million–27.0 million) if 100% of global funds were optimised during this period (figure 4B; appendix p 15). If the last reported global spending amounts were doubled, distribution towards prevention programmes targeting key populations at greatest risk would increase from 14% of the last reported budget allocation to an estimated 21% of the optimised budget, which is better aligned with the 25% advocated by UNAIDS to be spent on prevention.17 If the budget were doubled, the proportion of optimised funds for HIV testing would also increase from what is less than 1%, as last reported, to 9%. Lastly, to achieve a nearly 90% reduction in global HIV incidence from 2010 to 2030 with optimal allocation, an increase to approximately $40 billion in annual budget is estimated to be necessary for countries modelled.

Discussion
Limited HIV resources must be invested cost-effectively. We have shown that by optimising global HIV resources from 2016, about 30% more new infections and AIDS-related deaths could be averted by 2030 compared with existing allocations. In all countries and regions included in our analysis, the most common first priority towards achieving these reductions should be to scale up ART and one or more prevention and testing programmes targeting key populations.

Further reductions in incidence and deaths could be achieved if international HIV resources were redistributed, with a priority shift in funds to countries in sub-Saharan Africa. International funding organisations like The Global Fund and PEPFAR might choose to consider enhancing their strategic investment from the global perspective towards countries or settings where the greatest health outcome could be achieved.18 Should additional funds become available, certain programmes that are not prioritised at existing funding levels will become a higher priority for funding (for example, HIV testing programmes within the allocations for sub-Saharan Africa and for Latin America and the Caribbean).

These findings are consistent with previous modelling studies in specific countries19–21 and for the sub-Saharan Africa region.22 With a constant HIV budget, optimal reallocation towards ART and programmes targeting key populations would necessitate that funds be shifted.

Figure 3: HIV resource allocations for targeted and non-targeted prevention programmes by country
Stacked bars show the last reported (LR) and optimal (O) HIV programme resource allocations to minimise new HIV infections from 2016 to 2030 for each country modelled. ART=antiretroviral therapy. Other KP=other key population prevention. FSW=female sex workers. MSM=men who have sex with men. NSP=needle and syringe programmes. OST=opiate substitution therapy. OVC=orphans and vulnerable children. PMTCT=prevention of mother-to-child transmission. PWID=people who inject drugs. SBCC=social behaviour change communication. VMMC=voluntary medical male circumcision. Supporting data are available in the appendix (p 15).

Figure 4: Global HIV resource allocations and new HIV infections with varying budget
(A) Optimal allocation of annual global HIV programme resources to minimise new HIV infections by 2030 with 20% incremental budget increases from 0% to 200% compared with the last reported (LR) allocation with 100% budget. (B) Cumulative new HIV infections by population group from 2016 to 2030 at variable budget levels. We did not consider non-targeted programme spending within the optimal allocation, so these programmes were excluded here. Other KP=other key population prevention. ART=antiretroviral therapy. FSW=female sex workers. MSM=men who have sex with men, including people who are transgendered. NSP=needle and syringe programmes. OST=opiate substitution therapy. PMTCT=prevention of mother-to-child transmission. PWID=people who inject drugs. SBCC=social behaviour change communication. VMMC=voluntary medical male circumcision. Supporting data are available in the appendix (p 15).
away from low-impact programmes. Although it is not surprising that recommendations are to increase ART funding, the optimal allocation of remaining resources was context-specific. Funding to other programmes in a prioritised HIV strategy depends on local epidemiology and cost-effectiveness of local service delivery. Moreover, with varying budget, different programmes are prioritised for funding. Lastly, we reaffirm the ongoing gap in global HIV resource needs. More resources will be needed to achieve HIV incidence targets by 2030.

Although not included in the optimisation, almost half of total HIV funds from modelled countries are being spent on non-targeted programmes, with wide variation in spending by region and at the national level. This probably reflects different accounting and administrative frameworks and suggests an opportunity to capitalise on reducing spending on non-targeted programmes and to optimally reinvest any savings in targeted programmes to further improve health outcomes.

As with any modelling study, there are limitations to this global HIV model analysis. First, the model only includes countries with the greatest numbers of people living with HIV by region, capturing 80% of all people living with HIV worldwide, and only $12·8 billion of the roughly $19·2 billion in annual HIV spending reported for all low-income and middle-income countries. Second, limitations in data availability and reliability can lead to uncertainty about projected results. Although the model optimisation algorithm accounts for inherent uncertainty, it might not be possible to account for all aspects of uncertainty because of poor quality or insufficient data, particularly for important cost values. Cost functions, which were applied to every HIV programme in every country, were the primary driver (coupled with epidemic burden) of optimal resource estimations. Third, we used evidence from systematic reviews of clinical and research studies to inform model assumptions. These assumptions might be conservative; more optimistic values might exist in certain settings, for example in the level of programme efficacy, which would lead to even further projected health gains. We did a sensitivity analysis by varying key parameters and showed the effect on model outcomes. Fourth, we did not capture the effect of migration of people living with HIV between countries but instead model countries in isolation. Fifth, because of limited data availability, we did not include the potential effect of pre-exposure prophylaxis and cash transfers within the optimisation analysis, but we are working to include these interventions in future. Sixth, we did not incorporate time-varying optimisation where it might be optimal to scale up or to scale down programmes over time. We anticipated that this approach would have more appropriately prioritised funding to programmes for which health gains from early investment will be realised only in later years. We expected this limitation to mostly affect funding for voluntary medical male circumcision in generalised epidemic settings, as shown for South Africa by Shattock and colleagues. Seventh, for the optimisation scenario whereby international funds were permitted to be shifted between countries, we assumed that redistribution would not be limited to investment in select programme (or programmes), as is often specified by some funding organisations. Finally, these findings are only modelling analysis projections and have not been confirmed in practical settings. The models used in this study have been calibrated to reflect country-endorsed and UNAIDS-endorsed epidemiological estimates, but validation of results showing that optimal allocations are indeed more efficient in practice has not been possible. Shifting resources on the basis of evidence from resource optimisation studies is not always feasible and is not necessarily politically favourable, but it should be considered if there is the will to make a greater impact.

Resource redistribution towards programme combinations identified as more cost-effective in allocative efficiency studies, including ART and key population HIV prevention and testing, has been shown from Optima HIV modelling cases (eg, in Sudan and Belarus) and for many other countries that have used Optima HIV in Global Fund Concept Notes and National Strategic Plan development and target setting. The epidemiological effects of these programmatic changes are anticipated to be realised; however, rigorous impact evaluations have not been established, and they would not be simple at national levels because of the absence of an empirical counterfactual. Our choice of objective to minimise new HIV infections resulted in optimal reallocation of funds towards treatment, which would in turn lead to reductions in AIDS-related deaths. However, different objectives, for example to minimise AIDS-related deaths or disability-adjusted life-years, will result in somewhat different optimal allocations and outcomes. Finally, to examine accrual and spread of genotypic resistance was outside the scope of this model and study.

Using allocative efficiency analyses, we estimated where shifts in resource allocation could lead to improved effect with existing funding or to a similar effect with less funding. These findings have been used at the national level and could also be used at the regional and global levels to guide programme planners, policy makers, and donors in their decisions for improving population health outcomes. However, allocative efficiency will only improve the HIV response to a certain extent. Innovations must also be realised to deliver treatment and prevention services at reduced costs by revising policies to allow procurement of more affordable antiretroviral drugs, to deliver services at quality at reduced prices, and to aim to support essential health environments at appropriate cost. Ultimately, resources must be invested in the most cost-effective HIV programmes that target populations and locations where they will have the strongest health effect.
Contributors
SLK and DPW conceived the study. SLK, RM-H, DPW, RMS, CCX, DJK, and KLG developed the global HIV modelling approach. SLK, RM-H, XFY, RMS, KLG, SAH, IR, AJJ, LG, HH-B, ZB, OK, JP, RTG, CJB, NF, and CCK compiled the data and did the original country analysis. RM-H, SLK, RMS, and CCK did the global analyses. SLK wrote the first version of the Article. JS-W, MG and DW provided guidance. All authors reviewed and approved the final version of the Article.

Declaration of interests
We declare no competing interests.

Acknowledgments
The World Bank provided funding to support country applications, and the Australian NHMRC (APP1086540) provided modelling research funding for this study. SLK received a Graduate Scholarship and International Postgraduate Research Scholarship from Monash University. DPW received a Senior Research Fellowship from the Australian NHMRC. OK received a professorship grant from the Swiss National Science Foundation (163878). We thank Andrew Craig for help with data collation, Zara Shubber for contribution to in-country applications, and national representatives for their input.

References