

The Highest Attainable Standard of Evidence (HASTE) for HIV/AIDS Interventions: Toward a Public Health Approach to Defining Evidence

STEFAN D. BARAL, MD, MPH^a
ANDREA WIRTZ, MHS^a
FRANGISCOS SIFAKIS, PhD, MPH^a
BENJAMIN JOHNS, PhD^b
DAMIAN WALKER, PhD^c
CHRIS BEYRER, MD, MPH^a

ABSTRACT

Objectives. Evidence-driven decisions have become a standard for health interventions, policy, and programs. While randomized controlled trials (RCTs) are encouraged for public health interventions, there are limitations with RCTs as the gold standard of evidence for HIV interventions. We developed a novel system of evaluating evidence for assessing HIV preventive interventions termed the Highest Attainable Standard of Evidence (HASTE).

Methods. The HASTE system focuses on triangulation of three distinct categories of evidence: efficacy data, implementation data, and plausibility. We conducted systematic reviews, including experimental and observational data, to assess all available interventions for men who have sex with men (MSM). We collected implementation and programmatic data using a global electronic consultation, Internet searches, and in-person consultations. We assessed plausibility with expert analyses of both biological and public health evidence.

Results. HASTE includes four grades of evidence: Strong (Grade 1), Conditional (Grade 2), Insufficient (Grade 3), and Inappropriate (Grade 4). We used the HASTE system to evaluate the evidence for HIV interventions for MSM in low- and middle-income countries. Several differences emerged in the strength of recommendation with the use of the HASTE system, including strong recommendations for voluntary counseling and testing and for structural interventions.

Conclusions. The HASTE system addresses a need for an evidence evaluation tool that is specific for HIV interventions and facilitates an evaluation of biomedical, behavioral, and structural approaches using the highest standard of attainable evidence. HASTE represents a tool that balances scientific integrity and practicality in assessing the quality of evidence of preventive interventions targeting the most-at-risk populations for HIV.

^aJohns Hopkins School of Public Health, Department of Epidemiology, Center for Public Health and Human Rights, Baltimore, MD

^bJohns Hopkins School of Public Health, Department of International Health, Baltimore, MD

^cThe Bill and Melinda Gates Foundation, Seattle, WA

Address correspondence to: Stefan D. Baral, MD, MPH, Johns Hopkins School of Public Health, Department of Epidemiology, Center for Public Health and Human Rights, E7146, 615 N. Wolfe St., Baltimore, MD 21205; tel. 410-502-8975; fax 410-614-8371; e-mail <sbaral@jhsph.edu>.

©2012 Association of Schools of Public Health

Evidence-based decision-making has become a global standard for health interventions, policy, and programs. In the human immunodeficiency virus (HIV) arena, this methodology has been a welcome trend. While randomized controlled trials (RCTs) with HIV endpoints are encouraged for public health interventions, there are limitations with using RCT as the gold standard of evidence for HIV interventions.¹⁻³ These limitations include the very high cost of efficacy trials, the relative scarcity of populations with sufficient incidence in which to mount trials, and the ethical imperative to compare experimental treatments with ever more potent control conditions, which can mitigate the ability to assess efficacy. In addition, some interventions have never been formally evaluated with RCTs but have demonstrated significant impact from implementation and observational research, making RCTs now either unfeasible or ethically challenging. An example of this limitation is the use of needle and syringe exchange programs for injecting drug users (IDUs), for which many would argue the window of opportunity to conduct an RCT has long passed.⁴

Within the realm of clinical medicine, evidence-based medicine (EBM) is now considered the basis by which to define standards of clinical care. Defining packages of clinical services has been predicated on systematic reviews of individually randomized double-blinded placebo-controlled trials of medications and/or services for patients with varying clinical conditions. The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system has been widely endorsed as the most effective method with which to grade the current state of evidence for a variety of clinical interventions.^{5,6} The GRADE system was designed for individual-level clinical interventions in which the traditional hierarchy of evidence is applied. Specifically, the highest-quality evidence is derived from double-blinded placebo-controlled RCTs, followed by unblinded RCTs, prospective cohort studies, case-control studies, clinical case series, and consensus among experts. Additional weight is given to appropriately executed systematic reviews and meta-analyses of studies with little heterogeneity among participants, methods, and results.

To further standardize the presentation of evidence in clinical interventions and meta-analyses, criteria including the Consolidated Standards of Reporting Trials (CONSORT) and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) were developed, which have facilitated the grading of evidence.^{7,8} Given the nature of individual-level clinical interventions, the GRADE system has facilitated the development of clinical practice guidelines and other

clinical practice tools to promote the practice of EBM.⁹ The GRADE system is also relevant, as it integrates the potential for a separation between quality of evidence and strength of recommendations based on extenuating circumstances, such as cost-efficacy, risk-benefit, and contextual factors.¹⁰

While there is general acceptance of the use of the GRADE system in clinical medicine, there has been no widely accepted standard for grading public health interventions, although a series of different algorithms and hierarchies of evidence have been proposed.^{1,11-13} Varying algorithms of evaluating public health interventions are also in use by the International Agency for Research on Cancer, the United States Preventive Services Task Force (USPSTF), the newly revived Canadian Task Force on Preventive Health Care (CTFPHC), and the National Institute for Health and Clinical Excellence in the United Kingdom, among others. While Hill's criteria for causality, including strength of the relationship, dose-response, temporality, experimental evidence, analogy, and biologic plausibility, still apply to public health interventions, it is more difficult to demonstrate efficacy using traditional evaluation strategies.¹⁴ One reason for this difficulty is that public health interventions tend to be context-specific, including geographic and socioeconomic contexts, and are generally multifaceted. Moreover, primary prevention strategies targeting at-risk populations may be subject to the prevention paradox first described by Rose.¹⁵ The prevention paradox describes a situation in which an effective population-level public health intervention may provide only little benefit at the individual level while being significant at the population level, thereby complicating the measurement of effectiveness or efficacy.¹⁶

There is an emerging awareness of the significant roles the key or most-at-risk populations (MARPS) may play in HIV epidemics. Many country epidemics began as concentrated epidemics among MARPS including gay men and other men who have sex with men (MSM), sex workers, and IDUs, and then transitioned to more generalized epidemics. The role of MARPS in concentrated epidemics is relatively uncontroversial. However, in generalized epidemic settings, the initial presentation of numerous HIV epidemics was among MSM.^{3,4} With the emergence of generalized epidemics, the role of these three MARPS and other country-specific MARPS, such as truckers, internally displaced people, and victims of gender-based violence, has been given less attention. However, there is a growing evidence base of disproportionate risk among MARPS including sex workers, MSM, and IDUs in these settings.^{17,18} Recent assessments of global HIV prevention suggest that few

HIV/acquired immunodeficiency syndrome (AIDS) prevention, treatment, and care programs include targeted programming for these populations.¹⁹ Given this lack of funding, infections continue to increase in the context of slowing epidemics in the general population and, with the exception of IDUs, there has been limited progress in defining the optimal package of services for MARPS in low- and middle-income settings.

In response to this lack of progress, our multidisciplinary team worked to develop a novel system of evaluating evidence for HIV interventions targeting decreasing HIV risk specifically among MSM. We propose the use of the term the “Highest Attainable Standard of Evidence” (HASTE). HASTE was initially used to define a package of services for preventing HIV infection among MSM in low- and middle-income countries (LMIC). Other derivatives of the GRADE system have been suggested, including by Tang et al.¹³ When stigma affecting MARPS is compounded with the difficulties and limitations of RCTs evaluating public health intervention with biological endpoints, the evidence base for any HIV intervention supporting these vulnerable populations is limited. Thus, while our initial intent was to use the GRADE methodology to evaluate individual interventions, it became clear that the system required modifications to include assessment of what we termed HASTE. HASTE deliberately echoes the language of human rights conventions on the right to health, which accept that what can realistically be attained in resource-constrained environments can still serve as life-saving aspirational goals.^{20,21} The aspiration was to develop a system that balances scientific integrity with the need to make recommendations for prevention programming and policies for an understudied and underserved population.

METHODS

The HASTE process

The HASTE system proposed in this article is specific to HIV interventions among MSM and focuses on the triangulation of three main characteristics: efficacy data, implementation data, and biological and public health plausibility, although it may be adaptable to other public health areas and disciplines. The inclusion of a review of efficacy data is a common denominator across all systems evaluating levels of evidence. However, the response to preventing HIV infection has differed compared with most other clinical conditions in that implementation of these preventive measures is generally managed by civil society and not-for-profit nongovernmental institutions rather than health-care facilities. Although RCT data from interventions are

often unattainable, these interventions constitute a predominant majority of HIV prevention services in LMIC. Thus, our group considered capturing these implementation data that are crucial in informing the optimal package of services for MSM.

Hill’s criteria for causality remain as the most relevant set of determinants of whether a risk factor causes disease or an intervention causes prevention or mitigation of a disease. One of these criteria is plausibility, and our group considered this criterion vital, similar to others before us.¹³ However, in considering the public health plausibility of an intervention, we assessed whether a preventive intervention with limited demonstrated efficacy in preventing HIV as a biological endpoint was within the causal pathway of another intervention that does have demonstrable effectiveness in preventing HIV infection. We will use the example of the role of voluntary counseling and testing to illustrate the importance of public health plausibility for the HASTE system (Figure 1).

To review the evidence for effective interventions to prevent HIV among MSM, we used a process first proposed by the Centers for Disease Control and Prevention’s (CDC’s) Community Preventive Services Task Force when developing an evidence-based guide to public health interventions.¹⁰ This process included forming a multidisciplinary team, developing an approach to assessing the evidence for interventions, selecting which interventions would be included in the review, developing systematic search strategies for each intervention, implementing these search strategies, assessing and summarizing the quality of evidence using a standardized abstraction tool, and then translating this evidence into a set of recommendations for including interventions into a combination HIV prevention/interventions package. Finally, through the global consultation for implementation data describing interventions supporting the needs of MSM in LMIC, we considered information and data outside the realm of effectiveness, and have identified and summarized research gaps (Figure 2). Two reviewers independently abstracted data from the identified reports using a standardized abstraction tool. Conclusions were presented to a multidisciplinary group for review before final adoption of the recommendation.

Efficacy data

We used a scoping review as defined by Arksey and O’Malley to determine individual components of the package of preventive services for MSM.²² In brief, the scoping methodology is focused on mapping literature relevant to HIV preventive interventions for MSM without assessing the quality of included studies. These

Figure 1. A HASTE assessment of the importance of VCT programs for combination HIV prevention interventions for MSM in low- and middle-income countries

| <i>HASTE characteristics</i> | <i>Assessment</i> |
|------------------------------|--|
| Efficacy data | The demonstrated efficacy of VCT interventions in preventing HIV infection has been limited. ^a |
| Implementation data | A large body of consistent implementation data highlighting the importance of VCT as a point of interaction with the health-care system that may result in the provision of a series of reproductive health and preventive services, including sexually transmitted infection diagnosis and care as well as mental health counseling, among other preventive services. A critical step in the pathway for preventive interventions based on knowledge of HIV status. |
| Plausibility | HIV prevention technologies are increasingly dependent on knowledge of HIV status, including oral and topical ART chemoprophylaxis for those not living with HIV and early ART for those living with HIV/AIDS to decrease onward transmission and community viral load. There is evidence of behavior change post-VCT for women living with HIV and also for men living with HIV. ^{a-c} Normalizing testing can also reduce stigma associated with HIV status, which can increase uptake of ART. ^{d,e} Interventions to increase awareness of HIV serostatus have public health plausibility. ^{f,i} |
| Rating | Strong recommendation |

^aHoltgrave D, McGuire J. Impact of counseling in voluntary counseling and testing programs for persons at risk for or living with HIV infection. *Clin Infect Dis* 2007;45 Suppl 4:S240-3.

^bHayes R, Sabapathy K, Fidler S. Universal testing and treatment as an HIV prevention strategy: research questions and methods. *Curr HIV Res* 2011;9:429-45.

^cOjo O, Verbeek JH, Rasanen K, Heikkinen J, Isotalo LK, Mngoma N, et al. Interventions to reduce risky sexual behaviour for preventing HIV infection in workers in occupational settings. *Cochrane Database Syst Rev* 2011;(12):CD005274.

^dKawichai S, Celentano D, Srithanaviboonchai K, Wichajarn M, Pancharoen K, Chariyalertsak C, et al. NIMH Project Accept (HPTN 043) HIV/AIDS Community Mobilization (CM) to Promote Mobile HIV Voluntary Counseling and Testing (MVCT) in Rural Communities in Northern Thailand: modifications by experience. *AIDS Behav* 2012;16:1227-37.

^eSweat M, Morin S, Celentano D, Mulawa M, Singh B, Mbwambo J, et al. Community-based intervention to increase HIV testing and case detection in people aged 16–32 years in Tanzania, Zimbabwe, and Thailand (NIMH Project Accept, HPTN 043): a randomised study. *Lancet Infect Dis* 2011;11:525-32.

^fGrant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *N Engl J Med* 2010;363:2587-99.

^gCohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med* 2011;365:493-505.

^hMcGowan I. Rectal microbicides: can we make them and will people use them? *AIDS Behav* 2011;15 Suppl 1:S66-71.

ⁱDas M, Chu PL, Santos GM, Scheer S, Vittinghoff E, McFarland W, et al. Decreases in community viral load are accompanied by reductions in new HIV infections in San Francisco. *PLoS One* 2010;5:e11068.

HASTE = Highest Attainable Standard of Excellence

VCT = voluntary counseling and testing

HIV = human immunodeficiency virus

MSM = men who have sex with men

ART = antiretroviral therapy

AIDS = acquired immunodeficiency syndrome

reviews included the development of a specific search protocol with selection criteria, but did not include tracking the numbers of included studies. We used sensitive search terms to facilitate the development of an overview of the state of prevention sciences to help facilitate decisions on which prevention methods were to be further assessed. Once we defined the potential components of the package of services, we completed an umbrella review of systematic reviews if systematic reviews of the efficacy of individual components had

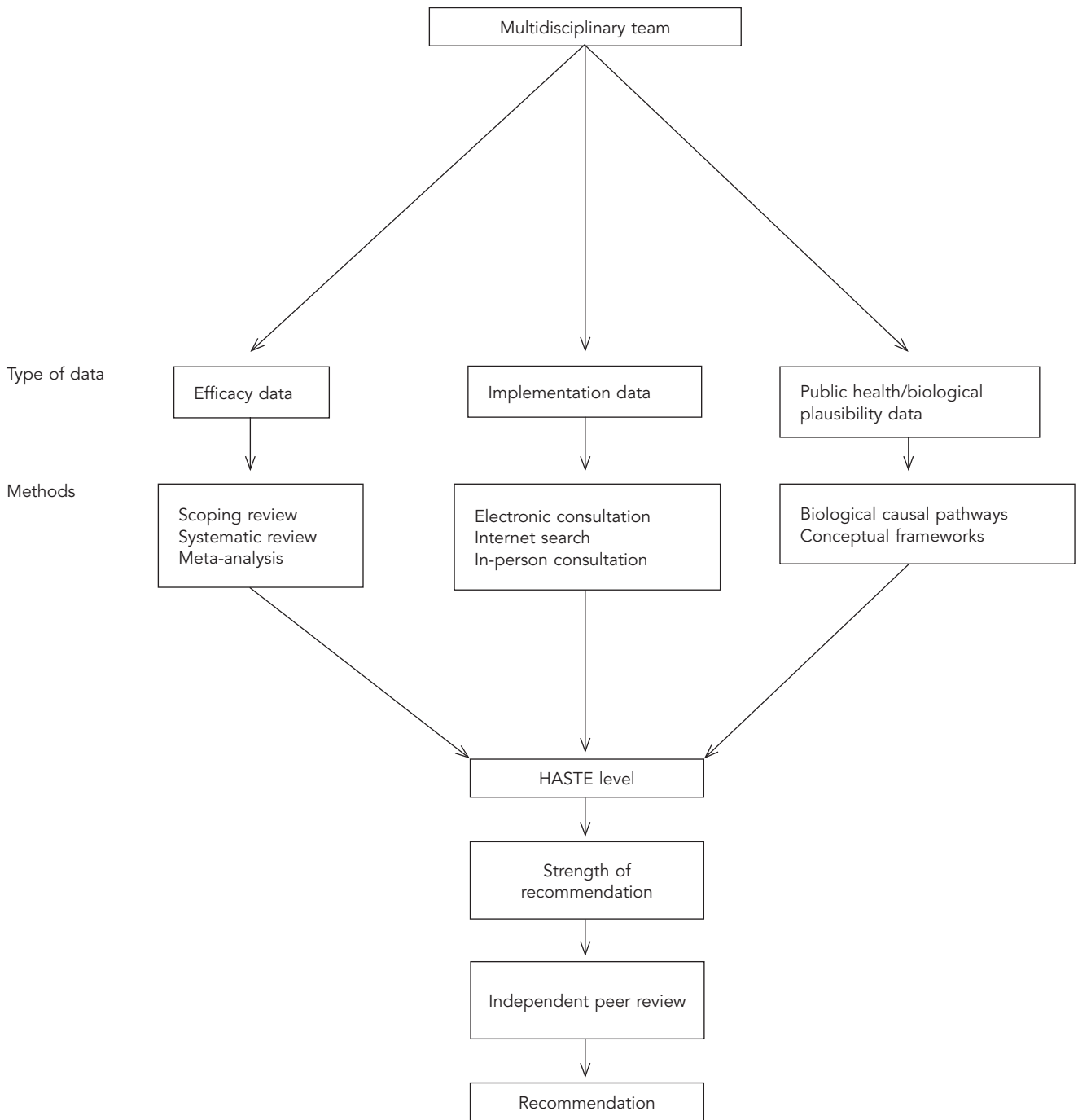
been previously completed, such as assessing the efficacy of behavioral interventions to decrease HIV risk among MSM.²³ However, each individual article used in the systematic reviews was collected to assess specific inclusion and exclusion criteria.

We completed novel systematic reviews with defined search protocols when previous systematic reviews were not harnessed on individual components of HIV prevention for MSM. When quantitative outcomes were available, such as changes in self-reported unprotected

anal intercourse secondary to targeted interventions, we completed a meta-analysis. If meta-analysis was not possible, either because there was only one study or there was significant intra-study variation in design,

relevant data were abstracted with the aforementioned tool, conclusions on efficacy were made by three different investigators, and results were compared until consensus was reached.

Figure 2. HASTE system for HIV interventions process



HASTE = Highest Attainable Standard of Evidence
 HIV = human immunodeficiency virus

Implementation data

We used several different methods to harness implementation data.²⁴ The majority of small implementers with limited scope do not have websites to post these data, but they do produce digital reports of their programs for their funders. As such, an electronic global consultation was completed in October 2009. Letters requesting information on epidemiology, rights contexts, and programming for MSM were sent out using dedicated listservs in Asia, Africa, Latin America and the Caribbean, and Eastern Europe. These letters were also sent out by key funders of related initiatives including the American Foundation for AIDS Research (known as amfAR) with its MSM Initiative to its grantees, the MSM Global Forum for HIV, and key United Nations (UN) agencies including UN Development Programme (UNDP) and UNAIDS. In addition, we contacted key informants in 28 countries requesting information specific to their country. In total, we retrieved implementation data from 68 LMIC pertaining to MSM. To attain implementation data from larger implementers, we searched Google and Bing using similar keywords to what was used for the systematic searches. In addition, we individually searched the websites of large international HIV prevention implementers known to provide services for MARPS for further documentation of programs and outcomes.

Separate from this search, a global consultation was held after the umbrella review, separate systematic reviews, and further review of electronic documents describing implementation data. This consultation included representation from several UN agencies including the UNDP, UNAIDS, World Health Organization, and the UN Populations Fund, in addition to key informants from 15 countries where a member-checking and results-validation session was completed. We presented the package of services using the framework of biomedical, behavioral, and structural interventions along with the evaluation approach. As this consultation was held early in the process of developing this set of recommendations, the feedback for the process and content of this package facilitated appropriate changes being made to the package of recommendations and evaluation system. One of the key changes was the need to review the evidence supporting conversion or reparative therapies, whose focus is to try and change sexual orientation, as these interventions form the base for many interventions targeting MSM in LMIC with a focus on regions in the Middle East and North Africa.²⁵

While there is no common evaluation framework used by HIV implementers large or small, our group used the Reach, Efficacy, Adoption, Implementation,

and Maintenance (RE-AIM) framework, which has five key components in assessing programmatic reach; efficacy or effectiveness; and measures of adoption, implementation, and maintenance to assess the relative importance of key interventions.²⁶ The RE-AIM framework is well established for non-randomized processes and program evaluations.^{27–29}

Reach functions as a measure of program scale and can be measured in several ways, including the number or percentage of the target audience being reached. Because the denominator or true population size of MSM is rarely known, programs describe the absolute number of people reached as the most common characterization of scale. Generalizability of the population reached in the particular program is also included under reach.

Efficacy or effectiveness is measured by assessing the changes in a set of appropriate outcomes, the impact on the quality of life of participants, and any potential adverse effects of the program. Adoption is measured by assessing the proportion of those targeted by the intervention who actually participated in the intervention. It can also be considered as a measure of acceptability of this program with impact on ultimate program uptake. Implementation-related issues include a feasibility assessment of the program with key components, including a proprietary assessment, program evaluability, resource needs for the program (e.g., fixed and marginal costs), and legality. The assessment of program implementation also evaluates whether the program was consistently delivered episodically or in different geographic regions.

Finally, maintenance of the program represents sustainability of the programmatic outcomes generally assessed with cutoffs of either six months or one year. The non-randomized study or program data abstraction tool focused on extracting quantitative and qualitative data from programmatic studies under each of the realms of RE-AIM. While the abstraction tool and framework for evaluation included the potential for assessing all components of RE-AIM, most programmatic descriptions did not report each of these elements.

Plausibility

We applied the biological plausibility criterion to assess biomedical interventions using current levels of knowledge of biological causal pathways. We assessed public health plausibility using conceptual frameworks of HIV risk among MSM developed by members of the team akin to the analytic frameworks used by the Guide to Community Preventive Services.³⁰

Independent peer review

The aforementioned global consultation meeting served as an interim validation and review of the recommendations suggested in the package of services. A separate process for ensuring the appropriateness, akin to the GRADE values and preferences process, of the conclusions included peer review by HIV/AIDS prevention experts not connected with the study.⁶ The external reviewers included both academic experts in HIV/AIDS epidemiology and targeted interventions for MSM, as well as experts in implementation of interventions for MSM in high-, middle-, and low-income settings.

RESULTS

HASTE for HIV interventions

The HASTE process for HIV/AIDS interventions was modified from several tools used to evaluate public health interventions including ones developed by USPSTF, CTFPHC, and GRADE to evaluate the quality of evidence of HIV interventions for MSM. The HASTE system gives the highest weight of efficacy evidence to RCTs, systematic reviews, and meta-analyses of RCT studies where attainable. The differentiating factors for this system are that other experimental studies, such as non-randomized controlled studies and pre- and post-assessments of interventions, were also included in RCTs pending an assessment using the RE-AIM framework.

This approach generated four grades of evidence for HASTE—Strong (Grade 1), Conditional (Grade 2), Insufficient (Grade 3), and Inappropriate (Grade 4)—depending on the amount of efficacy trial and implementation data available for a given intervention (Figure 3). Grade 1 (i.e., strong) recommendations are given when there are available efficacy or implementation data that the benefit of an intervention clearly outweighs the potential risks, or there are data that the intervention addresses a known causal risk factor of HIV risk among MSM using the approach of assessing plausibility.

Grade 2 (i.e., conditional) recommendations are given when there is potential efficacy for an intervention, but further research or implementation data are required to make a strong recommendation. Conditional recommendations may be given when there is not a large body of implementation data and limited quality of experimental research evaluating a particular intervention. Specifically, for Grade 2a (i.e., probable) interventions, there may be limited or no efficacy data, but there are plausibility and consistent implementation data highlighting the importance of this intervention. For Grade 2b (i.e., possible) interventions, there are limited or inconsistent efficacy data, plausibility, but only limited implementation data from non-randomized studies or programmatic data. The most common difference between interventions receiving probable or possible recommendations is the amount of non-randomized data available for

Figure 3. Highest Attainable Standard of Evidence (HASTE) system for HIV interventions—four grades of evidence

| <i>Grade level</i> | <i>Strength of recommendation</i> | <i>Explanation</i> |
|--------------------|-----------------------------------|--|
| Grade 1 | Strong | <ul style="list-style-type: none"> • Efficacy is consistent • High plausibility • Large body of consistent implementation data |
| Grade 2 | Conditional | |
| Grade 2a | Probable | <ul style="list-style-type: none"> • Limited efficacy data • Plausibility • Consistently effective from implementation data |
| Grade 2b | Possible | <ul style="list-style-type: none"> • Limited or inconsistent efficacy data • Plausibility • Limited but consistent implementation data |
| Grade 2c | Pending | <ul style="list-style-type: none"> • Ongoing efficacy trials • Plausibility |
| Grade 3 | Insufficient | <ul style="list-style-type: none"> • Inconsistent data • Undefined plausibility • Inconsistent or paucity of implementation data |
| Grade 4 | Inappropriate | <ul style="list-style-type: none"> • Consistent data demonstrating lack of efficacy • Consensus from implementation data of inappropriate intervention |

HIV = human immunodeficiency virus

assessment using the RE-AIM framework. In addition, for some interventions such as circumcision, Grade 2b recommendations were given if the intervention had the potential to be beneficial for a small subset of MSM, such as those who have high insertivity ratios and multiple female partners.³¹ Separately, for some interventions, Grade 2c (i.e., pending) recommendations are given when there are ongoing efficacy trials among MSM.

Grade 3 (i.e., insufficient) is applied when there is inconsistent evidence for the efficacy of a particular intervention, the causal pathway is unclear or not well defined, and there are limited non-randomized study or programmatic data about a particular intervention. Grade 4 (i.e., inappropriate) is given when there is evidence of no efficacy or effectiveness of an intervention, where there are efficacy or consistent implementation data suggesting harm, where risks outweigh potential benefits, or where there is consensus from implementation data that this is an inappropriate intervention.

Combination HIV prevention interventions for MSM

We used the HASTE system to evaluate the evidence for individual HIV-related interventions for MSM in LMIC. The results of this process are presented in Figure 4, although the complete information and rationale for selection are available in the World Bank report by Beyrer et al.³² Figure 4 also presents an analysis of which interventions are supported by RCTs in both high-income countries and LMIC, as defined by the World Bank Atlas Method.³³ The only interventions evaluated using RCTs in LMIC are individualized risk-reduction counseling, brief client-focused counseling, community-level interventions, and preexposure prophylaxis. Of these interventions, only a single HIV intervention, preexposure prophylaxis, has been evaluated as an RCT among MSM in a generalized HIV epidemic. Thus, these are the only interventions that would be awarded four points as a starting point based on type of evidence under the GRADE system, whereas all observational studies would be given a starting value of two. Moreover, only the oral chemoprophylaxis study included an objective outcome of HIV seroincidence, involved blinding, and had high retention rates; thus, the others would lose between two and three points based on quality of evidence. Directedness of generalizability of all studies but oral chemoprophylaxis is limited given that it was the only one with efficacy completed in an LMIC. Finally, effect size as graded by having a measure of association greater than two or five would also be a limitation, as no HIV intervention has been shown to decrease HIV incidence with this magnitude. With this scoring framework, where a

numerical value of four or higher is associated with a high strength of recommendation—three being associated with medium, two being associated with low, and one or lower being associated with a very low-strength recommendation—only oral chemoprophylaxis would receive a strong recommendation for MSM in LMIC.

Several differences emerge in the strength of recommendation with the use of the HASTE system compared with a classical RCT-driven evaluation system. These differences were strong recommendations for voluntary counseling and testing and for structural interventions. An example of the abbreviated HASTE analysis for voluntary counseling and testing is shown in Figure 1. Structural-level risk interventions among MSM have not been evaluated with an RCT, partly because of the complexity of the study designs required to characterize the efficacy and effectiveness of these interventions, and partly due to logistical and cost considerations. However, there are convincing public health plausibility and implementation data that highlight the importance of these interventions to facilitate the development and delivery of HIV services for MSM in LMIC.³²

DISCUSSION

Among MSM in LMIC, there is a confluence of high HIV-related risk practices at the individual level with structural barriers to HIV preventive services including stigma, discrimination, and criminalization. Drivers of HIV risk among MSM occur at multiple levels including the individual, social, and sexual networks; community; and public policy realms. Together, these factors have resulted in a disproportionate burden of HIV disease among MSM wherever studied, including generalized epidemics.¹⁸ The imperative to mount evidence-based and comprehensive responses to this disproportionate risk is urgent, as more than three decades into the epidemic most LMIC have no dedicated services for MSM. The initial systems for evaluating evidence to define recommendations, including the GRADE system, were initially designed for individual-level clinical interventions. Moreover, the available systems in the literature evaluating public health interventions did not include detailed assessments of implementation data, which comprise a large component of the evidence base in the field of HIV/AIDS prevention. There is significant variability in organizations focused on implementing HIV prevention service delivery, ranging from small community-based organizations to large not-for-profit institutions active in a number of countries spanning multiple continents. Many of the smaller community-based organizations serving the needs of MARPS tend to be led by members of these same populations that

Figure 4. Strength of recommendations for inclusion of specific prevention components in combination HIV prevention intervention packages for MSM

| <i>Intervention</i> | <i>Components</i> | <i>HASTE level</i> | <i>Strength of recommendation</i> | <i>RCT in HIC</i> | <i>RCT in LMIC</i> |
|---|--|--------------------|-----------------------------------|----------------------|----------------------|
| Condoms and water- or silicone-based lubricants | Effective distribution | Grade 1 | Strong | | |
| | Counseling and education about condoms and lubricants | Grade 1 | Strong | | |
| | Thicker condoms | Grade 2b | Possible | | |
| | Female condom | Grade 2b | Possible | | |
| Behavioral interventions | Individualized risk-reduction counseling | Grade 2a | Probable | Yes ^{a-o} | Yes ^{p,q} |
| | Conversion or reparative therapy | Grade 4 | Inappropriate | | |
| | Didactic teaching | Grade 2b | Possible | Yes ^{b,r,s} | |
| | Brief client-focused counseling | Grade 2a | Probable | Yes ^{t-z} | Yes ^{aa} |
| | Community-level interventions | Grade 1 | Strong | Yes ^{bb-mm} | Yes ^{nn,oo} |
| Antiretroviral therapy | Method of secondary prevention | Grade 1 | Strong | | |
| | Primary prevention | Grade 2a | Probable | | |
| | Address substance use and mental health disorders to increase adherence | Grade 2a | Probable | Yes ^{pp-ss} | |
| | Rectal microbicides | Grade 2a | Pending | | |
| | Postexposure prophylaxis | Grade 2a | Probable | | |
| Voluntary counseling and testing | Preexposure prophylaxis | Grade 2a | Probable | Yes ^{tt} | Yes ^{tt} |
| | HIV testing | Grade 1 | Strong | | |
| | Rapid testing for HIV | Grade 2a | Probable | | |
| Circumcision | LGBT sensitization | Grade 1 | Strong | | |
| | Medical male circumcision | Grade 2b | Possible | | |
| Syndromic STI treatment | HSV-2 suppression | Grade 4 | Inappropriate | Yes ^{uu} | Yes ^{uu} |
| | Screening and treatment of all genital ulcerative diseases | Grade 2a | Probable | | |
| Vaccination | HPV | Grade 2a | Probable | | |
| Structural interventions | Government-sponsored anti-homophobia policy, community systems strengthening, and health sector training | Grade 1 | Strong | | |

^aImrie J, Stephenson JM, Cowan FM, Wanigaratne S, Billington AJP, Copas AJ, et al. A cognitive behavioural intervention to reduce sexually transmitted infections among gay men: randomised trial. *BMJ* 2001;322:1451-6.

^bCarballo-Diéguez A, Dolezal C, Leu CS, Nieves L, Díaz F, Decena C, et al. A randomized controlled trial to test an HIV-prevention intervention for Latino gay and bisexual men: lessons learned. *AIDS Care* 2005;17:314-28.

^cCleary PD, Van Devanter N, Steilen M, Stuart A, Shipton-Levy R, McMullen W, et al. A randomized trial of an education and support program for HIV-infected individuals. *AIDS* 1995;9:1271-8.

^dValdiserri RO, Lyter DW, Leviton LC, Callahan CM, Kingsley LA, Rinaldo CR. AIDS prevention in homosexual and bisexual men: results of a randomized trial evaluating two risk reduction interventions. *AIDS* 1989;3:21-6.

^eKelly JA, St. Lawrence JS, Hood HV, Brasfield TL. Behavioral intervention to reduce AIDS risk activities. *J Consult Clin Psychol* 1989;57:60-7.

^fDilley JW, Woods WJ, Sabatino J, Lihath T, Adler B, Casey S, et al. Changing sexual behavior among gay male repeat testers for HIV: a randomized controlled trial of a single-session intervention. *J Acquir Immune Defic Syndr* 2002;30:177-86.

^gChesney MA, Chambers DB, Taylor JM, Johnson LM, Folkman S. Coping effectiveness training for men living with HIV: results from a randomized clinical trial testing a group-based intervention. *Psychosom Med* 2003;65:1038-46.

^hKoblin B, Chesney M, Coates T. Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: the EXPLORE randomised controlled study. *Lancet* 2004;364:41-50.

ⁱWolitski RJ, Gomez CA, Parsons JT. Effects of a peer-led behavioral intervention to reduce HIV transmission and promote serostatus disclosure among HIV-seropositive gay and bisexual men. *AIDS* 2005;19 Suppl 1:S99-109.

^jRotheram-Borus MJ, Swendeman D, Comulada WS, Weiss RE, Lee M, Lightfoot M. Prevention for substance-using HIV-positive young people: telephone and in-person delivery. *J Acquir Immune Defic Syndr* 2004;37 Suppl 2:S68-77.

^kWolitski RJ, Parsons JT, Gomez CA, Purcell DW, Hoff CC, Halkitis PN. Prevention with gay and bisexual men living with HIV: rationale and methods of the Seropositive Urban Men's Intervention Trial (SUMIT). *AIDS* 2005;19 Suppl 1:S1-11.

^lPatterson TL, Shaw WS, Semple SJ. Reducing the sexual risk behaviors of HIV+ individuals: outcome of a randomized controlled trial. *Ann Behav Med* 2003;25:137-45.

continued on p. 581

Figure 4 (continued). Strength of recommendations for inclusion of specific prevention components in combination HIV prevention intervention packages for MSM

- ^mBerg R. The effectiveness of behavioural and psychosocial HIV/STI prevention interventions for MSM in Europe: a systematic review. *Euro Surveill* 2009;14:19430.
- ⁿGore-Felton C, Rotheram-Borus MJ, Weinhardt LS, Kelly JA, Lightfoot M, Kirshenbaum SB, et al. The Healthy Living Project: an individually tailored, multidimensional intervention for HIV-infected persons. *AIDS Educ Prev* 2005;17(1 Suppl A):21-39.
- ^oMyers T, Tudiver FG, Kurtz RG, Jackson EA, Orr KW, Rowe CJ, et al. The Talking Sex Project: descriptions of the study population and correlates of sexual practices at baseline. *Can J Public Health* 1992;83:47-52.
- ^pZimmerman MA, Ramirez-Valles J, Suarez E, de la Rosa G, Castro MA. An HIV/AIDS prevention project for Mexican homosexual men: an empowerment approach. *Health Educ Behav* 1997;24:177-90.
- ^qToro-Alfonso J, Varas-Díaz N, Andujar-Bello I. Evaluation of an HIV/AIDS prevention intervention targeting Latino gay men and men who have sex with men in Puerto Rico. *AIDS Educ Prev* 2002;14:445-56.
- ^rPeterson JL, Coates TJ, Catania J, Hauck WW, Acree M, Daigle D, et al. Evaluation of an HIV risk reduction intervention among African-American homosexual and bisexual men. *AIDS* 1996;10:319-25.
- ^sRobert B, Rosser S. Evaluation of the efficacy of AIDS education interventions for homosexually active men. *Health Educ Res* 1990;5:299-308.
- ^tPicciano JF, Roffman RA, Kalichman SC, Rutledge SE, Berghuis JP. A telephone based brief intervention using motivational enhancement to facilitate HIV risk reduction among MSM: a pilot study. *AIDS Behav* 2001;5:251-62.
- ^uRichardson JL, Milam J, McCutchan A, Stoyanoff S, Bolan R, Weiss J, et al. Effect of brief safer-sex counseling by medical providers to HIV-1 seropositive patients: a multi-clinic assessment. *AIDS* 2004;18:1179-86.
- ^vHospers HJ, Debets W, Ross MW, Kok G. Evaluation of an HIV prevention intervention for men who have sex with men at cruising areas in the Netherlands. *AIDS Behav* 1999;3:359-66.
- ^wRoffman RA, Picciano JF, Ryan R, Beadnell B, Fisher D, Downey L, et al. HIV-prevention group counseling delivered by telephone: an efficacy trial with gay and bisexual men. *AIDS Behav* 1997;1:137-54.
- ^xPicciano JF, Roffman RA, Kalichman SC, Walker DD. Lowering obstacles to HIV prevention services: effects of a brief, telephone-based intervention using motivational enhancement therapy. *Ann Behav Med* 2007;34:177-87.
- ^yCoates TJ, McKusick L, Kuno R, Stites DP. Stress reduction training changed number of sexual partners but not immune function in men with HIV. *Am J Public Health* 1989;79:885-7.
- ^zChoi KH, Lew S, Vittinghoff E, Catania JA, Barrett DC, Coates TJ. The efficacy of brief group counseling in HIV risk reduction among homosexual Asian and Pacific Islander men. *AIDS* 1996;10:81-7.
- ^{aa}Sampaio M, Brites C, Stall R, Hudes ES, Hearst N. Reducing AIDS risk among men who have sex with men in Salvador, Brazil. *AIDS Behav* 2002;6:173-81.
- ^{ab}Community-level HIV intervention in 5 cities: final outcome data from the CDC AIDS Community Demonstration Projects. *Am J Public Health* 1999;89:336-45.
- ^{ac}AIDS community demonstration projects: implementation of volunteer networks for HIV-prevention programs—selected sites, 1991–1992. *MMWR Morb Mortal Wkly Rep* 1992;41(46):868-9, 875-6.
- ^{ad}Hays RB, Rebchook GM, Kegeles SM. The Mpowerment Project: community-building with young gay and bisexual men to prevent HIV1. *Am J Community Psychol* 2003;31:301-12.
- ^{ae}Flowers P, Hart GJ, Williamson LM, Frankis JS, Der GJ. Does bar-based, peer-led sexual health promotion have a community-level effect amongst gay men in Scotland? *Int J STD AIDS* 2002;13:102-8.
- ^{af}Jones KT, Gray P, Whiteside YO, Wang T, Bost D, Dunbar E, et al. Evaluation of an HIV prevention intervention adapted for black men who have sex with men. *Am J Public Health* 2008;98:1043-50.
- ^{ag}Miller RL, Klotz D, Eckholdt HM. HIV prevention with male prostitutes and patrons of hustler bars: replication of an HIV preventive intervention. *Am J Community Psychol* 1998;26:97-131.
- ^{ah}Kelly JA, St Lawrence JS, Diaz YE, Stevenson LY, Hauth AC, Brasfield TL, et al. HIV risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. *Am J Public Health* 1991;81:168-71.
- ^{ai}Kegeles SM, Hays RB, Pollack LM, Coates TJ. Mobilizing young gay and bisexual men for HIV prevention: a two-community study. *AIDS* 1999;13:1753-62.
- ^{aj}Harding R, Bensley J, Corrigan N, Franks L, Stratman J, Waller Z, et al. Outcomes and lessons from a pilot RCT of a community-based HIV prevention multi-session group intervention for gay men. *AIDS Care* 2004;16:581-5.
- ^{ak}Kelly JA, Murphy DA, Sikkema KJ, McAuliffe TL, Roffman RA, Solomon LJ, et al. Randomised, controlled, community-level HIV-prevention intervention for sexual-risk behaviour among homosexual men in US cities. *Community HIV Prevention Research Collaborative. Lancet* 1997;350:1500-5.
- ^{al}Operario D, Smith CD, Arnold E, Kegeles S. The Bruthas Project: evaluation of a community-based HIV prevention intervention for African American men who have sex with men and women. *AIDS Educ Prev* 2010;22:37-48.
- ^{am}Williamson LM, Hart GJ, Flowers P, Frankis JS, Der GJ. The Gay Men's Task Force: the impact of peer education on the sexual health behaviour of homosexual men in Glasgow. *Sex Transm Infect* 2001;77:427-32.
- ^{an}Amirkhanian YA, Kelly JA, Kabakchieva E, McAuliffe TL, Vassileva S. Evaluation of a social network HIV prevention intervention program for young men who have sex with men in Russia and Bulgaria. *AIDS Educ Prev* 2003;15:205-20.

continued on p. 582

Figure 4 (continued). Strength of recommendations for inclusion of specific prevention components in combination HIV prevention intervention packages for MSM

^{oo}Amirkhanian YA, Kelly JA, Kabakchieva E, Kirsanova AV, Vassileva S, Takacs J, et al. A randomized social network HIV prevention trial with young men who have sex with men in Russia and Bulgaria. *AIDS* 2005;19:1897-905.

^{pp}Stall RD, Paul JP, Barrett DC, Crosby GM, Bein E. An outcome evaluation to measure changes in sexual risk-taking among gay men undergoing substance use disorder treatment. *J Stud Alcohol* 1999;60:837-45.

^{qq}Shoptaw S, Reback CJ, Peck JA, Yang X, Rotheram-Fuller E, Larkins S, et al. Behavioral treatment approaches for methamphetamine dependence and HIV-related sexual risk behaviors among urban gay and bisexual men. *Drug Alcohol Depend* 2005;78:125-34.

^{rr}Sorensen JL, Dilley J, London J, Okin RL, Delucchi KL, Phibbs CS. Case management for substance abusers with HIV/AIDS: a randomized clinical trial. *Am J Drug Alcohol Abuse* 2003;29:133-50.

^{ss}Kelly JA, Murphy DA, Bahr GR, Kalichman SC, Morgan MG, Stevenson LY, et al. Outcome of cognitive-behavioral and support group brief therapies for depressed, HIV-infected persons. *Am J Psychiatry* 1993;150:1679-86.

^{tt}Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *N Engl J Med* 2010;363:2587-99.

^{uu}Celum C, Wald A, Hughes J, Sanchez J, Reid S, Delany-Moretlwe S, et al. Effect of aciclovir on HIV-1 acquisition in herpes simplex virus 2 seropositive women and men who have sex with men: a randomised, double-blind, placebo-controlled trial. *Lancet* 2008;371:2109-19.

HIV = human immunodeficiency virus

MSM = men who have sex with men

HASTE = Highest Attainable Standard of Evidence

RCT = randomized controlled trial

HIC = high-income country

LMIC = low- and middle-income countries

LGBT = lesbian, gay, bisexual, and transgender

STI = sexually transmitted infection

HSV-2 = herpes simplex virus-2

HPV = human papillomavirus

recognize the vulnerable state and disproportionate burden of HIV risk and disease among their peers.

Given limited resources, human and financial resources are focused on service delivery rather than sophisticated evaluation strategies. Larger HIV prevention service implementers often evaluate interventions using process indicators of HIV prevention rather than the measurement of biological endpoints. Moreover, HIV implementers rarely submit the results of their programs for publication in peer-reviewed journals. This publication dearth is likely because smaller-scale implementers lack the inherent technical capacity to do so and because larger implementers do not allocate resources to this activity. The HASTE system presented in this article was developed to be able to comprehensively assess all forms of evidence relating to HIV prevention for MARPS and was not developed as a tool to evaluate evidence relating to all public health interventions. Rather, this tool was used to develop recommendations for combination HIV preventive interventions for MSM in LMIC. The full HASTE analysis supporting Figure 4 is included in the World Bank report, "The Global HIV Epidemics Among Men Who Have Sex with Men."³² Figure 4 summarizes the interventions assessed, demonstrating that few of these interventions have been tested with RCT, although many are considered the standard of care rendering

RCT level of evidence unattainable. Even fewer of these interventions have been tested in LMIC, where the risk environment is often heightened because of stigma and discrimination. Consequently, if RCT-level evidence is held as the standard by which one can make strong assertions of evidence efficacy, there would be few interventions that would receive a strong recommendation. As such, we believe that the HASTE tool can also be used to define an appropriate package of HIV services for other vulnerable populations including sex workers, IDUs, victims of gender-based violence, and internally displaced or other migrant populations.

Limitations

There were clear limitations to the HASTE system presented in this article. For one, there is no common evaluation framework for HIV/AIDS implementers, and the evaluation of these programs tends to be specific to the identified needs of the funders focusing on process rather than outcome indicators. While we used the RE-AIM system that has been previously used to evaluate health sector programming, a standardized system of evaluating and reporting non-randomized research studies was suggested by CDC and termed Transparent Reporting of Evaluations with Non-Randomized Designs (TREND).³⁴ However, this system has not yet been widely adopted by HIV implementing

organizations. Standardizing the evaluation and reporting of HIV interventions would likely increase the availability of evidence in the peer-reviewed literature. In addition, when including implementation data from a series of organizations, there is inconsistency in the quality of these data, given that the primary purpose of these programs is to provide services rather than to conduct research.

To compensate for these limitations, the HASTE system combines implementation data with efficacy data and plausibility analysis. Observational studies and analysis of implementation data do not allow for analysis at the level of the individual, thereby making interpretation subject to ecological fallacy. Additionally, reproducibility of the HASTE system for MSM in LMIC may be difficult, as it involves retrieving implementation data that are not all indexed in the public domain. Significant effort would be required to retrieve these data. Finally, the HASTE system purposefully does not assign a numerical score to different components of the assessment, which may be perceived as being less objective than score-based systems such as GRADE. The development team decided that the assignment of abstract numerical values to evidence attributes also represents a subjective judgment assessment. However, we believe that despite these limitations, the system presented represents a tool that balances scientific integrity and practicality in assessing the quality of evidence of HIV/AIDS preventive interventions targeting MARPS for HIV.

CONCLUSIONS

In the context of a slowing HIV pandemic, HIV incidence and prevalence continue to increase among MARPS. HIV interventions for these populations have been inconsistently implemented, which is likely secondary to inadequate political motivation to address vulnerable populations, insufficient targeted funding, and the lack of a means to define an optimal package of services in resource-constrained settings. We have developed the HASTE system to define an appropriate package of HIV services including biomedical, behavioral, and structural approaches using the highest standard of attainable data. This system can be used to advocate for a package of HIV prevention, treatment, and care services for MARPS in low- and middle-income settings, allowing advocacy for the appropriate scale-up of these services in response to evidence-based need.

This study was a component of a project funded by the Global AIDS Monitoring and Evaluation Team of the World Bank. The World Bank did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the

article. This study relied on publicly available documents and was, therefore, exempt from Institutional Review Board determination.

REFERENCES

1. Rychetnik L, Frommer M, Hawe P, Shiell A. Criteria for evaluating evidence on public health interventions. *J Epidemiol Community Health* 2002;56:119-27.
2. Wolfe D, Carrieri MP, Shepard D. Treatment and care for injecting drug users with HIV infection: a review of barriers and ways forward. *Lancet* 2010;376:355-66.
3. Vandembroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med* 2008;5:e67.
4. Degenhardt L, Mathers B, Vickerman P, Rhodes T, Latkin C, Hickman M. Prevention of HIV infection for people who inject drugs: why individual, structural, and combination approaches are needed. *Lancet* 2010;376:285-301.
5. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. Going from evidence to recommendations. *BMJ* 2008;336:1049-51.
6. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6.
7. Moher D, Altman DG, Liberati A, Tetzlaff J. PRISMA statement. *Epidemiology* 2011;22:128; author reply 128.
8. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637-9.
9. Ansari MT, Tsertsvadze A, Moher D. Grading quality of evidence and strength of recommendations: a perspective. *PLoS Med* 2009;6:e1000151.
10. Briss PA, Zaza S, Pappaioanou M, Fielding J, Wright-De Agüero L, Truman BI, et al. Developing an evidence-based Guide to Community Preventive Services—methods. The Task Force on Community Preventive Services. *Am J Prev Med* 2000;18(1 Suppl 1):35-43.
11. Petticrew M, Roberts H. Evidence, hierarchies, and typologies: horses for courses. *J Epidemiol Community Health* 2003;57:527-9.
12. McQueen DV. The evidence debate. *J Epidemiol Community Health* 2002;56:83-4.
13. Tang KC, Choi BC, Beaglehole R. Grading of evidence of the effectiveness of health promotion interventions. *J Epidemiol Community Health* 2008;62:832-4.
14. Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295-300.
15. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;14:32-8.
16. John S. Why the prevention paradox is a paradox, and why we should solve it: a philosophical view. *Prev Med* 2011;53:250-2.
17. Strathdee SA, Hallett TB, Bobrova N, Rhodes T, Booth R, Abdool R, et al. HIV and risk environment for injecting drug users: the past, present, and future. *Lancet* 2010;376:268-84.
18. Beyrer C, Baral SD, Walker D, Wirtz AL, Johns B, Sifakis F. The expanding epidemics of HIV type 1 among men who have sex with men in low- and middle-income countries: diversity and consistency. *Epidemiol Rev* 2010;32:137-51.
19. Global HIV Prevention Working Group. Global HIV prevention: the access, funding, and leadership gaps. Washington: Global HIV Prevention Working Group; 2009.
20. United Nations. The Universal Declaration of Human Rights. Paris: UN General Assembly; December 10, 1948.
21. Office of the United Nations High Commissioner for Human Rights. International Covenant on Economic, Social and Cultural Rights. New York: UN General Assembly; January 1976.
22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005;8:19-32.
23. Johnson WD, Diaz RM, Flanders WD, Goodman M, Hill AN, Holtgrave D, et al. Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men. *Cochrane Database Syst Rev* 2008(3):CD001230.
24. Padian NS, Holmes CB, McCoy SI, Lyerla R, Bouey PD, Goosby EP. Implementation science for the US President's Emergency Plan for AIDS Relief (PEPFAR). *J Acquir Immune Defic Syndr* 2011;56:199-203.
25. Abu-Raddad LJ, Hilmi N, Mumtaz G, Benkirane M, Akala FA,

- Riedner G, et al. Epidemiology of HIV infection in the Middle East and North Africa. *AIDS* 2010;24 Suppl 2:S5-23.
26. Donabedian A. An assessment of the criteria approach to medical care evaluation. *Hosp Med Staff* 1976;5:1-9.
27. Allen K, Zoellner J, Motley M, Estabrooks PA. Understanding the internal and external validity of health literacy interventions: a systematic literature review using the RE-AIM framework. *J Health Commun* 2011;16 Suppl 3:55-72.
28. Glasgow RE, Dickinson P, Fisher L, Christiansen S, Toobert DJ, Bender BG, et al. Use of RE-AIM to develop a multi-media facilitation tool for the patient-centered medical home. *Implement Sci* 2011;6:118.
29. Shubert TE, Altpeter M, Busby-Whitehead J. Using the RE-AIM framework to translate a research-based falls prevention intervention into a community-based program: lessons learned. *J Safety Res* 2011;42:509-16.
30. Zaza S, Wright-De Agüero LK, Briss PA, Truman BI, Hopkins DP, Hennessy MH, et al. Data collection instrument and procedure for systematic reviews in the Guide to Community Preventive Services. Task Force on Community Preventive Services. *Am J Prev Med* 2000;18(1 Suppl):44-74.
31. Sanchez J, Sal Y Rosas VG, Hughes JP, Baeten JM, Fuchs J, Buchbinder SP, et al. Male circumcision and risk of HIV acquisition among MSM. *AIDS* 2011;25:519-23.
32. Beyrer C, Wirtz AL, Walker D, Johns B, Sifakis F, Baral SD. The global HIV epidemics among men who have sex with men. Washington: World Bank Publications; 2011.
33. World Bank. Gross national income per capita 2008, Atlas method and PPP. Geneva: World Bank; 2009.
34. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004;94:361-6.