

# HIV-Related Implementation Research for Key Populations: Designing for Individuals, Evaluating Across Populations, and Integrating Context

Sheree R. Schwartz, PhD, MPH, Amrita Rao, ScM, Katherine B. Rucinski, PhD, Carrie Lyons, MPH, Nikita Viswasam, MSPH, Carly A. Comins, MPH, Oluwasolape Olawore, ScM, and Stefan Baral, MD, MPH, MBA

**Introduction:** Key populations, including men who have sex with men, transgender people, sex workers, people who inject drugs, and incarcerated populations, experience high burdens of HIV and urgently need effective interventions. Yet the evidence base for implementation research (IR) with key populations remains weak and poses specific challenges to epidemiologic inference. We apply the Consolidated Framework for IR to consider specific challenges and recommendations for IR with key populations.

**Discussion:** Individuals within key populations exist within inner and outer settings—including organizational structures, legal (eg, criminalization), and funding environments—which influence the design, adoption and fidelity of interventions, and the potential sustainability of intervention scale-up. Underlying vulnerabilities and external stressors experienced at the individual level (eg, homelessness, violence) further impact participation and retention in IR. Thus, researchers should account for representation in the research process, beginning with community engagement in IR design and consideration of enumeration/sampling methods for key populations who lack probabilistic sampling frames. Interventions for key populations require substantial adaptation and complexity (eg, individually tailored, multicomponent) to ensure appropriateness; however, there is tension between the need for complexity and challenges to internal validity (fidelity) and external validity (generalizable scale-up). Finally, integrating contextual, sampling, and implementation elements into analytic approaches is critical for effectiveness evaluation.

**Conclusions:** Translation of efficacious findings at the individual level to effectiveness at the population level requires recognition of risk heterogeneity. Recognizing the nuances of working with key populations is essential to ensure that individuals are represented by design and therefore gains in population health can be achieved.

**Key Words:** implementation science, key populations, HIV, epidemiology, consolidated framework for implementation research

(*J Acquir Immune Defic Syndr* 2019;82:S206–S216)

## INTRODUCTION

Given intersecting individual-level, network-level, and structural risks, key populations, including sex workers, men who have sex with men (MSM), transgender persons, people who inject drugs (PWID), and incarcerated populations, bear high burdens of HIV independent of setting.<sup>1–4</sup> The epidemiology of HIV among key populations consistently demonstrates the disproportionate impact of HIV within these groups, which necessitates different needs in interventions. Evidence is required to guide effective implementation and scale programs and services for key populations to reduce HIV burden, particularly in low- and middle-income settings. Recent data highlight this need for addressing the heterogeneity of populations requiring effective HIV interventions, given the limited population-level effectiveness of interventions that are highly efficacious at the individual level among the general population.<sup>5–8</sup> Advancing the HIV response thus requires characterizing optimal implementation strategies for HIV prevention and treatment interventions that are responsive to individual needs while balancing effectiveness and sustainability.

HIV-related implementation research (IR) with key populations poses specific challenges to epidemiologic inference emerging from multiple intersecting (co-occurring) stigmas that result in marginalization, such as stigma toward same-sex behavior and HIV and stigmas faced by women engaged in sex work who use drugs and are living with HIV.<sup>9,10</sup> This can lead to selection biases, information biases, fidelity of implementation execution, and limited transportability of results across populations and settings. Identification of key populations is often challenging because group membership is defined by behavior that is often difficult to verify.<sup>11–13</sup> Furthermore, the absence of probabilistic sampling frames for key populations is a barrier to understanding denominators for key populations and to measuring the extent to which individuals within key populations are included or missing from research or programs.<sup>11–13</sup> Yet, recognition of who is and who is not represented is critical to understanding

From the Key Populations Program, Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, MD.

Supported by the Johns Hopkins University Center for AIDS Research (P30AI094189).

The authors have no conflicts of interest to disclose.

Correspondence to: Sheree R. Schwartz, PhD, MPH, 615 N. Wolfe St, E7644, Baltimore, MD 21205 (e-mail: sschwartz@jhu.edu).

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

the generalizability of the findings and the potential for population-level impact, particularly given the social, political, and economic marginalization of key populations. Comprehensive consideration of representation in research and programs across the IR process includes attention to population engagement, epidemiologic design, sampling, recruitment, retention, implementation outcomes, individual outcomes, dissemination, and the interplay of marginalization with each of these elements.

We propose to use the Consolidated Framework for Implementation Research (CFIR) to assess specific challenges for IR with key populations. Using the IR literature and our experiences conducting epidemiologic and IR with key populations, our objectives were to fill a gap by documenting IR challenges with key populations, particularly in resource-constrained settings, assess their potential epidemiologic impact, and recommend strategies for conducting IR moving forward.

### CONSOLIDATED FRAMEWORK FOR IR

Figure 1 illustrates the proposed, applied CFIR framework highlighting constructs that inform IR for key populations and their interdependencies.<sup>14,15</sup> Broadly, the 5 CFIR domains that facilitate studying the potential for generaliz-

ability and scalability of research findings include *characteristics related to individuals*, the *inner* and *outer settings* that they are embedded, *intervention characteristics* and *implementation and research processes*.<sup>14</sup> The interplay of the CFIR domains is a strength of the framework and a critical reason it was selected for this assessment. Although the “research process” and “intervention characteristics” are the domains most commonly considered for IR challenges, threats to epidemiologic inference may emerge from each of these domains because invariably the outer setting, inner setting, and individual characteristics will influence the research process and consideration of intervention characteristics (Table 1). Within the 5 broad CFIR domains, we have highlighted subconstructs of particular relevance to key populations. Other subconstructs may have relevance but were determined, based on the literature and our experience, to have less prominence and not specifically highlighted. Although we recognize that key populations represent very different groups of individuals, themes of marginalization are shared across groups and result in varied, but similar, IR challenges. Thus, although we apply this construct overall to key populations across domains and subconstructs, we highlight issues of relevance within specific key population groups and provide illustrative examples of challenges, solutions, and threats to epidemiologic inference within

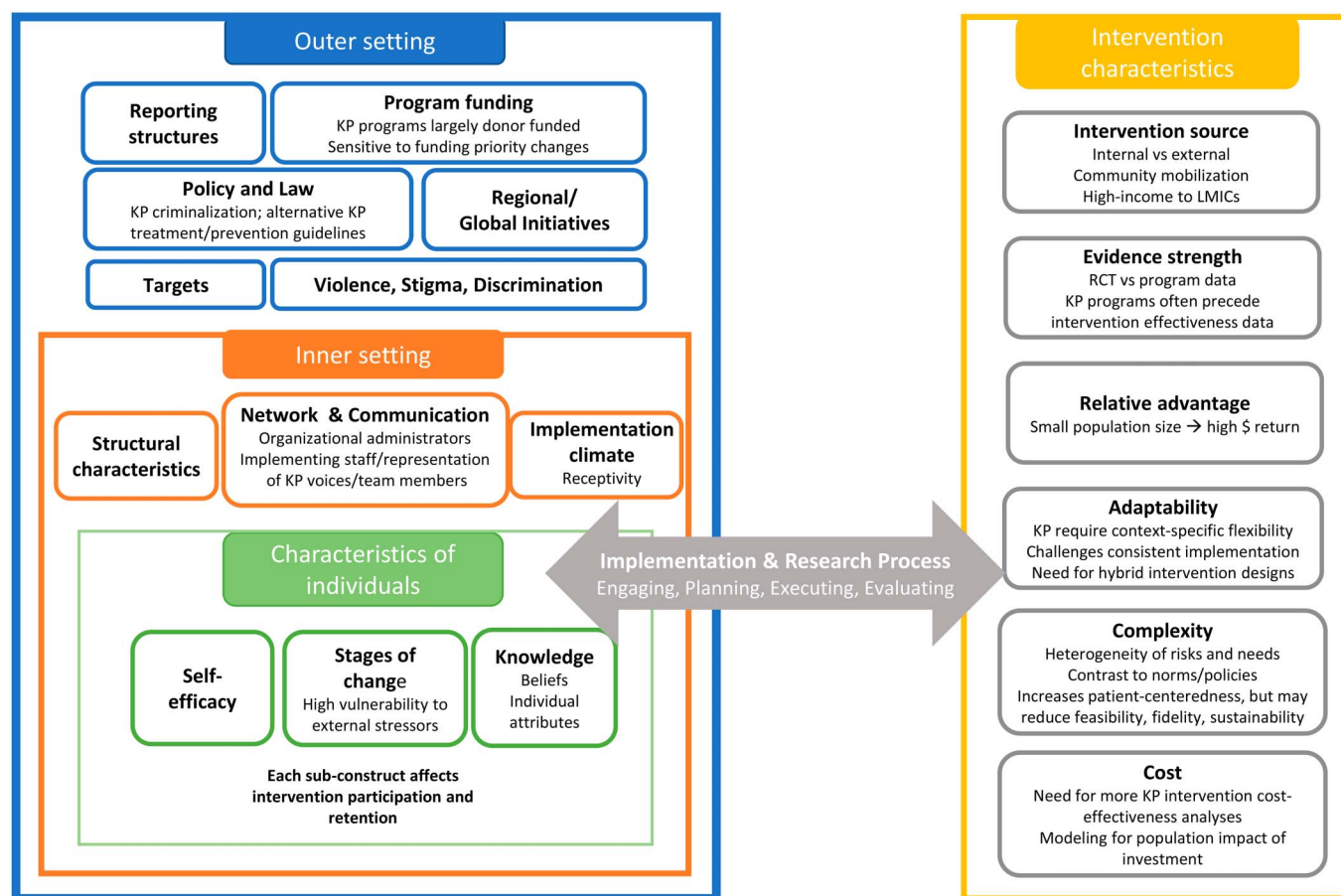


FIGURE 1. CFIR framework applied for key populations.

**TABLE 1.** Illustrative Examples of Key Population-Specific Threats to Scientific Inference Emanating Across Domains of the CFIR

Exemplary Challenges and Recommendations by Domain and Subconstruct	Implications for Inference
<p><b>Outer setting</b></p> <p><b>External policies:</b> Criminalization and limited legal protections for those who engage in sex work, practice same-sex behaviors, and use drugs may limit uptake of HIV prevention and treatment services and study participation. Evaluating the impact of laws can be difficult if working in a single setting. Mathematical modeling of structural drivers of HIV transmission among FSW in Canada and pre-post evaluations using legislative changes as a natural experiment with MSM in Nigeria are examples of approaches evaluating impact of external forces on health outcomes. As demonstrated by an increase in loss to follow-up in the Nigerian study with MSM, populations may avoid study participation in response to laws, policy initiatives, or external environments of stigma and discrimination.<sup>20,25</sup></p> <p><b>Peer pressure:</b> To facilitate participant trust and intervention engagement, key populations research is often conducted with local nongovernmental or community-based organizations already engaged in service delivery with the key population. However, given lags in research funding from time of concept to award, ethical approvals, and implementation, often donors funding programs decide to implement interventions before evidence for key populations is available. These decisions aim to accelerate epidemic response, and program grantee recipients are expected to swiftly implement changes. This creates challenges when similar interventions are being tested through randomized studies with key populations because researchers cannot ethically withhold the standard of care from participants. Solutions to this challenge include creating shorter duration trials, step-wedged approaches to program scale-up, researching implementation outcomes of existing programs, and building adaptive elements into the study design that allow tested interventions to evolve with the program.</p>	<p>State-supported marginalization reinforces the utility of evaluating implementation and effectiveness outcomes within contextual settings; results from one setting may not generalize to areas in which behaviors are criminalized, highly stigmatized, and/or discriminated against. Mathematical modeling can measure the potential impact of policies and structural interventions but does not always account for variations in implementation and enforcement. Similarly, pre-post policy evaluation often lacks controls or allows individuals to serve as their own controls—which, in light of unanticipated policy shifts, may be impacted by informative censoring or poor measurement specification if the study was not initially designed to capture legal or policy impacts.</p> <p>Changes to the program and standard of care will necessitate changes within research programs, potentially affecting the internal validity of findings and diluting the intervention effect of the intention-to-treat analyses, if standard of care arms begin receiving the intervention. Furthermore, this may result in the absence of evidence around intervention effectiveness within key populations, which may differ from results seen more broadly.</p>
<p><b>Inner setting</b></p> <p><b>Network and communications:</b> A methadone-assisted treatment program in Tanzania for PWID was supported by 4 community based organizations working in different capacities to promote HIV prevention efforts among PWID. Yet despite recruitment efforts from trusted community based organizations, mistrust remained common, particularly among women who inject drugs. Utilization of female peers, including women who had used drugs and engaged in sex work, was identified as a critical strategy to more effectively engage women.<sup>49,50</sup></p> <p><b>Structural characteristics; network and communications:</b> Utilization of peer outreach workers in intervention delivery creates opportunities but can also lead to unanticipated challenges stemming from the structural characteristics and communication within networks. In a study run with FSW in South Africa by our team, peer FSW were strongly recommended by the community advisory group to fill the case manager positions providing ART adherence support to FSW living with HIV. Peer case managers received extensive case management, HIV and research training. However, organizational salary structures that use common guidelines of basing salary ranges on education create income disparities and higher rates of turnover after case managers have been upskilled. Hiring of complementary case managers outside of the sex work community also resulted in internal conflict. Solutions include budgeting for high turnover, investment in capacity building for peers, and organizational review of policies that undervalue real-world or street experience vs. education, when this experience is the most important qualification for the position.<sup>51</sup></p>	<p>Miscommunication and mistrust in peer networks can lead to selection bias/underrepresentation of women who inject drugs in program and effectiveness data.</p> <p>High turnover creates threats to intervention fidelity, particularly interventions which rely on continuity of care.</p>

**TABLE 1. (Continued)** Illustrative Examples of Key Population–Specific Threats to Scientific Inference Emanating Across Domains of the CFIR

Exemplary Challenges and Recommendations by Domain and Subconstruct	Implications for Inference
<p>Characteristics of individuals</p> <p>Knowledge and beliefs about the intervention: Inaccurate risk perception and lack of awareness about available prevention interventions can limit the ability to recruit and retain MSM into interventions or research studies. Across high- and low-income settings, MSM, particularly younger MSM, often attribute low PrEP uptake to low-risk perception for HIV infection. Because risk perceptions can operate as either a motivator or a barrier, MSM who readily enroll and are retained in studies may be fundamentally different from those who do not, impacting findings. Encouraging peers to recruit peers may allow for research and programs to reach those not otherwise engaged in care to limit these effects.<sup>43,52–54</sup></p> <p>Self-efficacy: Factors in the outer settings such as criminalization or discrimination of same-sex practices, lack of recognition of different gender expressions and associated stigmas may impact personal agency of MSM and TGW, affecting their ability to engage in or adhere to interventions and compromising intervention effectiveness. Approaches including motivational interviewing that empower participants and bolster self-efficacy can mitigate these effects.</p>	<p>Level of knowledge, risk perception, and other characteristics of individuals can lead to inadequate recruitment, sample representation, and generalizability of findings.</p> <p>Decreased self-efficacy can influence who chooses to enroll and continue in studies or programs (representativeness) and may further modify the effectiveness of implementation strategies.</p>
<p>Intervention characteristics</p> <p>Intervention source, adaptability: In sex work–focused implementation science research, engaging or enabling community mobilization is critical to establish stakeholder ownership and the capacity for sustained implementation. It is ideal for sex workers to assume an active leadership role in the development, implementation, and evolution of research. Such was the case with India’s Sonagachi project, which operated for 16 years.<sup>1</sup> In this program, initially started as a clinic program to improve uptake of sexually transmitted infection treatment and condom use among sex workers through peer outreach, early stakeholder feedback indicated that sex workers’ environment, literacy rates, and limited self-efficacy compromised implementation effectiveness. To mitigate these challenges, the program expanded its training of peer outreach workers to include HIV prevention education through pocket visual charts, and “helper” skills to assist in facilitating condom negotiations between sex workers and clients.<sup>55,56</sup></p> <p>Complexity: Complex interventions tailored to individual needs and/or which require content delivery or multiple engagements may experience implementation challenges that can undermine the effectiveness of an intervention, if not circumvented. For example, individual or group-based interventions with incarcerated populations and ex-prisoners upon release may be complex to implement—from the perspective of ensuring repeat attendance of participants during a time of potential transition and mobility, full delivery of session content, high quality of content delivery, and appropriate tailoring of counseling and support to the individual client needs. Measurement of intervention feasibility and fidelity are essential to understand intervention effectiveness results and potential scalability—although often omitted or not reported. Measurement of implementation from both participant and implementer perspectives are valuable to triangulate the evidence, and potential chart review and mixed methods approaches.</p>	<p>Measurement of community engagement and ownership in interventions should be considered as an outcome and potential effect measure modifier of intervention effectiveness, but measurement may be complex. Furthermore, planned and real-time adaptations to existing interventions to better guide key populations-specific implementation should strive to maintain fidelity to intervention core components; tension between achieving “full and consistent implementation” while making rapid adaptations to meet setting-specific challenges may undermine internal validity or reproducibility.</p> <p>Inconsistencies in implementation documentation will result in information bias and reduce the ability to evaluate constraints or facilitators of effectiveness, including dose, quality, and appropriateness. This will be particularly problematic if there is differential documentation, such that those executing with lower intervention fidelity also have poorer documentation.</p>

(continued on next page)

**TABLE 1. (Continued)** Illustrative Examples of Key Population–Specific Threats to Scientific Inference Emanating Across Domains of the CFIR

Exemplary Challenges and Recommendations by Domain and Subconstruct	Implications for Inference
Implementation and research process	
Engagement, planning: Recruitment of key populations often requires nonprobability sampling approaches. These can result in nonrepresentative samples, depending on the methodology, degree, and quality of community engagement and partner selection before recruitment. Working with key decision-making stakeholders and key informants is critical to ensure buy-in, smooth study operation, and insight into when and where key populations groups commonly meet. The final methodology used may impact results. For example, a comparison of results from RDS and venue-based snowball sampling studies conducted in the same area with the same populations (FSW and MSM in both eSwatini and Cameroon) found that the RDS methodology sampled participants who were younger, less educated, and less engaged in care as compared with venue-based snowball sampling. Venue-based methods may be more reflective of programmatic estimates (as programs often target venues for implementation), whereas RDS methods may be more reflective of underlying population prevalence. Regardless, programmatic data should be interpreted in the context of existing integrated biobehavioral data (where available) and can be recalibrated when surveillance estimates emanate from program data. <sup>12</sup>	Selection biases may prevent participants from representing the larger source population; additionally, differences in the source and target population will impact generalizability.
Planning, execution: The Sisters Programme in Zimbabwe worked with FSW in 3 sites to intensify community-based activities and utilization of HIV prevention modalities and ART clinical services. Beyond program evaluation, researchers implemented a serial cross-sectional RDS study to compare results at baseline and at 5 years. Although this was a proactive attempt to get a representative sample and observe changes over time, one challenge in execution was that no control sites were included and rapid national program scale-up of the ART program in Zimbabwe co-occurred with the intervention, limiting the interpretation of positive study findings. <sup>57</sup>	Secular trends, if not accounted for, may result in overestimation or underestimation of intervention effect.

PrEP, pre-exposure prophylaxis; RDS, respondent-driven sampling; TGW, transgender women.

populations (Table 1). Heterogeneity also exists within populations, such as by age, venue type among sex workers, geographic location, outness among MSM and transgender women, prison release status among incarcerated or previously incarcerated populations, and type of drug use among PWID, the implications of which are considered.

OUTER SETTING

The outer setting comprises the economic, political, and social contexts that inform patient needs, external policies, and resources.<sup>14</sup> Specific to key populations, the legal and policy contexts, funding, and intersecting social stigmas, and marginalization will influence the design and fidelity of HIV-focused IR, and the potential feasibility of intervention scale-up and sustainability.

National or subnational policies will influence the content, design, and success of a study and the potential scale-up of an intervention. For implementation of studies among key populations, criminalization and limited legal protections for those who engage in sex work, practice same-sex behaviors, and use drugs are frequent challenges.<sup>16</sup> Criminalization or legal restrictions may drive populations to be more hidden and therefore less likely to engage in

services and studies. Service delivery through the government healthcare system or nongovernmental organizations may show discrepancies between the established standard-of-care and the practical implementation of those guidelines. Therefore, process evaluation of implementation, and communication and coordination with higher-level officials, clinic staff, and key population representatives, is critical to assess the adoption and fidelity of guideline implementation. For example, clinic staff may hold stigmatizing beliefs against MSM, female sex workers (FSW), transgender populations, or PWID, which may prevent full or high-quality implementation of the guidelines; sensitization efforts to address these stigmas may need to be included as intervention components and/or elements of process evaluation.<sup>17,18</sup> Furthermore, ensuring safety for both study participants and study staff is essential, especially in the context of criminalization. Therefore, engagement and partnership with key population representatives, and government, will support navigation of IR in these contexts.

Studies aiming to evaluate components of the outer setting are complex but important to assess the impact of a law or policy on the health of key populations.<sup>19</sup> In the context of national-level policies affecting key populations, rigorous evaluation of policy or structural

changes is difficult with the limited ability to use control groups. The impact of policy changes may be evaluated using preassessment and postassessment through natural experiments<sup>20</sup>; however, this requires data availability before a policy change. Interrupted time series is possible, but assumes that the preintervention context remains otherwise unchanged into the postintervention period. Ecological studies may be used to provide insight of policy effects across countries, but they do not account for individual-level characteristics and are prone to ecological fallacy.<sup>21–24</sup> Mathematical modeling may be used to estimate the potential effect of a policy or structural-level intervention that has not yet occurred. For example, mathematical modeling has been used to estimate the impact of sex work criminalization on HIV acquisition.<sup>25</sup>

Reporting structures and external incentives, such as funding, donor-driven targets, and regional or global initiatives, have implications for IR and program implementation. Modeling may demonstrate the population impact of implementing country-level strategies and scaling-up services for key populations, especially in response to meeting external mandates or adopting recommendations or guidelines.<sup>26</sup> However, scale-up of intervention strategies that have demonstrated efficacy, such as pre-exposure prophylaxis and universal treatment, may face real-world implementation challenges. Thus, scaling interventions shown efficacious with non-key populations has presented challenges in the absence of evidence of “how” to effectively implement services at a small scale, let alone nationally or with key populations. Representative research to appropriately adapt or design interventions before scale-up with key population groups will support more effective implementation.

### INNER SETTING

In the inner setting, considering the structures and experiences of particular organizations working with key populations plays a key role in influencing both the scope and the quality of IR.<sup>27</sup> Working with key populations is fundamentally contingent upon the establishment of relationships between the organization and the community, and thus service delivery is often led by key population-focused nongovernmental organizations. Three key constructs in the inner setting will impact key population intervention penetration and sustainability: (1) the structural characteristics of the organization; (2) the networks, communications, and overlap between organizational administrators, implementing staff, and the population served; and (3) the implementation climate and receptivity of the organization. Specifically, structural characteristics, inclusive of the social architecture and maturity of the organization, facilitate or hamper the organization's ability to reach and sustain contact with key populations through programming and research.<sup>28</sup> The degree to which organizations work in a top-down vs. a bottom-up approach to determining programming and research priorities will impact engagement and confidence of key populations in IR and in the services delivered. These elements, however, are not determined in organizational isolation, but are themselves embedded within the outer setting, including donor

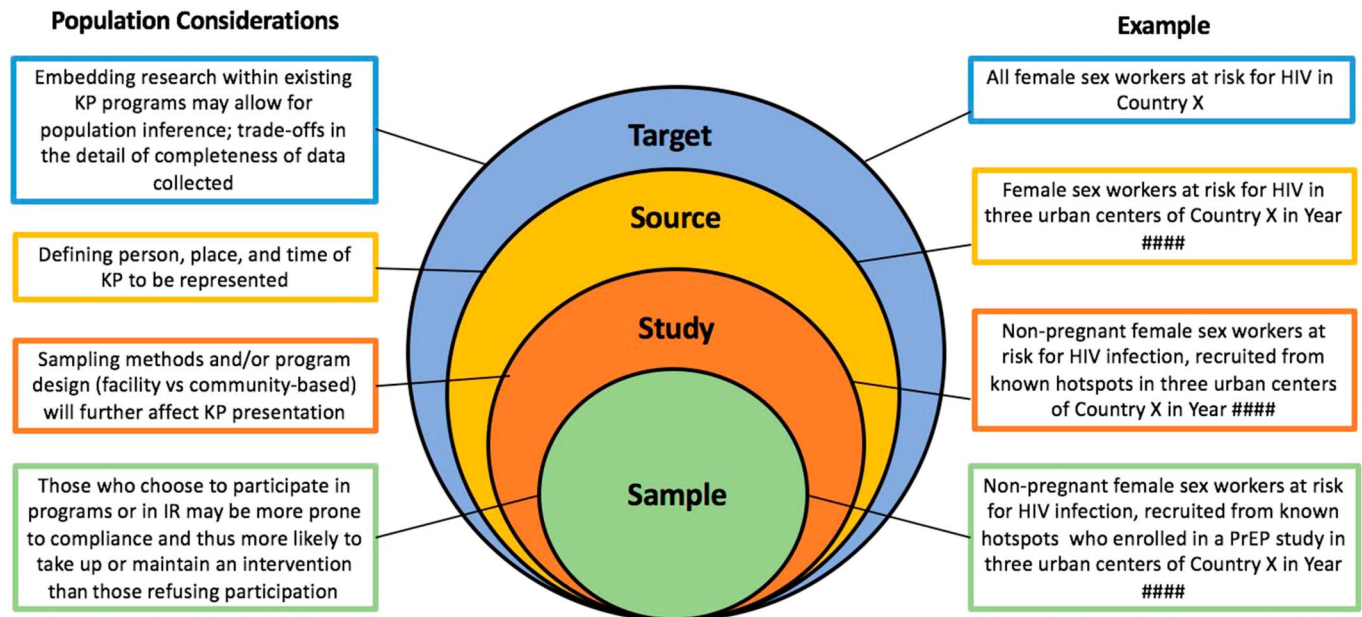
priorities. Because recruitment and intervention delivery with key populations are frequently decentralized and conducted via peer outreach workers, particularly with FSW, MSM, PWID, and transgender populations, the capacity and autonomy of research and implementation teams will impact intervention success, requiring organizational investment and growth opportunities.<sup>29</sup> The close networking of peers within the communities may also create real or artificial tiers within the organization, and peers may face stigma and discrimination within their role or, conversely, work together to undermine team members not from the community. The organizational capacity to change and the extent to which the research is championed within the organization plays a mediating role between implementation effectiveness and the overall management support needed within IR.<sup>30</sup> Yet, although the scope of IR necessitates consideration of an organization's inner setting, limited flexibility may exist in the selection of a key population implementing organization, and once selected, there is a need to implement within the existing structures of the organization. Furthermore, considering the generalizability of the results to areas in which key population-focused organizations are not present is critical to scale-up.

### CHARACTERISTICS OF INDIVIDUALS

The characteristics of key populations and service delivery providers are embedded within and consequently are often shaped by the outer and inner settings (Fig. 1). Although program and study design will affect the source and study populations, participation will be impacted by characteristics of individuals, affecting at the most central level who is represented in IR (Fig. 2). Specifically, knowledge, beliefs, and attributes of key populations will affect intervention appropriateness and uptake, whereas self-efficacy to adhere to interventions and move through the stages of change will impact both uptake and maintenance. Together, these elements will affect how representative a sample of the study population is, and the internal validity of the results based on intervention dropout.

Knowledge and beliefs about interventions will directly impact service and research participation. In the context of pervasive stigma, discrimination, and criminalization, prospective participants weigh the risks of disclosing their key population status—often a criterion for study participation—with potential benefits of engaging with research. This may be particularly true for HIV-negative individuals at high risk of acquisition who could benefit from prevention interventions but may not believe the disclosure risks are justified given that they are well.<sup>1,2</sup>

Similarly, self-efficacy and stage of change will impact the successful adoption, maintenance, and ultimately impact of interventions and will be heavily impacted by the outer and inner settings within which individuals reside. For example, stigma has been shown to negatively impact self-efficacy among transgender women,<sup>2</sup> and thus, intervention engagement may be limited among transgender women with reduced agency. Limited self-efficacy can diminish individuals' beliefs that the intervention will work for them, potentially



**FIGURE 2.** Key populations representation and data considerations from the individual participants to the broader population.

impacting recruitment, retention, and ultimately generalizability of findings. Furthermore, because key populations in many settings are more vulnerable to major life stressors and threats to survival, including arrests, violence, comorbidities, and mortality, high dropout can be anticipated because stability supports completion of the stages of change necessary for effective intervention. It is recommended that approaches both build self-efficacy components into interventions and measure self-efficacy and stigma to improve and interpret intervention results.

Successful implementation also involves consideration of other individual-level attributes that may serve as barriers to uptake of interventions, including demographics (age, literacy, and income), physical and mental health status, mobility, and relationships, which again will each be impacted by the outer setting.<sup>1</sup> Use of peers, monetary reimbursement, conveniently offered services, and long windows of follow-up visits may help mitigate these barriers. However, these improvements in accessibility may increase the cost of implementation, reduce scale-up feasibility, and introduce additional risks for key populations. For example, reimbursements should be carefully considered to reflect appropriate compensation for time and not as a form of coercion to participate in research activities.

### INTERVENTION CHARACTERISTICS

Characteristics of interventions—including what, for whom, by whom, when, for how long, with what intensity, and where—have particular relevance to conducting IR with key populations. Recognizing that heterogeneity between and within populations requires different intervention approaches is essential, but does not negate implementation challenges. Differentiated service delivery for key populations requires attention to intervention characteristics, including interven-

tion source, evidence strength, relative advantage, adaptability, complexity (eg, duration, scope, number of steps, and intricacy), and cost (Fig. 1).<sup>31</sup>

Key stakeholders' perception of an intervention as internally or externally developed can significantly affect implementation among key populations. Internal development approaches, such as community empowerment and mobilization led by key population members, can play a key role in developing tailored, effective interventions that encourage participation, retention, and sustainability.<sup>32</sup> This is particularly relevant for HIV interventions evaluating structural determinants of HIV infection among key populations, such as stigma, violence, and policing, which community mobilization approaches have sought to address in combination with individual-level clinical services.<sup>33,34</sup> Researchers can support consensual dialogues and development of shared language between authorities and key populations, and both attempt to intervene on these structural levels, and monitor engagement of key populations in structural interventions.<sup>35</sup>

High-quality evidence of interventions is limited with key populations, particularly in resource-constrained settings. In general, few randomized controlled trials assessing the effectiveness of HIV interventions have been conducted among key populations, presenting a persistent challenge to the translation of efficacy findings from broader population-based studies to key populations. Noting this, the Highest Attainable Standard of Evidence was developed to evaluate HIV interventions among MSM, focusing on the triangulation of efficacy data, implementation data, and biological and public health plausibility.<sup>36</sup> However, applications of these systems have noted barriers to generating high levels of evidence for specific implementation among key populations. Therefore, expanding IR for key populations, including adaptation of existing interventions, will likely require the



use of hybrid approaches to supplement existing programmatic data with targeted data collection. Furthermore, small-scale iterative testing of implementation strategies using tools that leverage existing resources, such as the multiphase optimization strategy,<sup>37</sup> can maximize efficiency and ultimately improve the strength of evidence for key populations.

Consideration of the adaptability of interventions, or the degree to which the peripheral elements of an intervention may be modified for context while maintaining fidelity to the tested core components, is particularly salient for key populations research. For key populations, structural issues including socially sanctioned discrimination and violence coupled with criminalization may further compromise effective implementation.<sup>38</sup> Thus, modifications to both the content (eg, changes to intervention materials, procedures, or delivery) and the context (eg, changes to format, location, or personnel)<sup>39</sup> of existing interventions may be most effective when accounting for the equity status of key populations and the intersection of intervention components with the outer, inner, and individual domains.<sup>33</sup>

Complexity is another critical construct for IR consideration. Complex interventions are likely needed to overcome the circumstances by which key populations experience an increase in HIV risk and transmission, yet complexity also creates implementation barriers. Although there is increasing recognition of the need for more human-centered design to meet heterogeneity of needs, there is likely to be a concomitant tension between individualized, patient-centered care and feasibility of implementation by overburdened health systems. Thus, assessment of feasibility and appropriateness before fully scaled implementation, followed by a comprehensive assessment of intervention fidelity, is essential to generate clear evidence of why interventions succeed or fail.<sup>40</sup>

Investments in comprehensive HIV programs for key populations are necessary to accelerate the HIV response, yet programs that address the specific needs of key populations remain severely underfunded.<sup>41</sup> Although intervention and implementation costs may be perceived as high per individual reached, programs that address the needs of key populations are generally considered cost-effective because of the potential for achieving large-scale epidemic impact through disruption of ongoing chains of transmission.<sup>42</sup> Therefore, increasing overall investment in IR for key populations, and including cost-effectiveness and mathematical modeling assessments of different implementation strategies across populations, is essential to evaluate and optimize intervention delivery and impact.

## PROCESS

The research and implementation process with key populations comes with a unique set of challenges impacting participant representation, information precision and accuracy, and ultimately epidemiologic inference (Table 2). Broadly, these process issues can be categorized into issues of engagement, planning, execution, and evaluation.

Key population members are not only recipients of the intervention but also often play a major role in the

implementation process and the uptake of results. Therefore, engaging with key population communities before study start, during implementation, and throughout interpretation of the results is crucial. This engagement can take the form of community advisory groups or dissemination of information through key population networks. Because misinformation can spread quickly in highly networked groups, clear and concise information is needed around the objectives and implications of the research. Misinformation can lead to limited or differential participation, with potential participants becoming distrustful of study activities.

IR planning should account for real-world implementation structures, context-specific challenges, and community feedback. The selected sampling method will impact the subset of the population reached and the potential for subsequent inference and generalizability (Fig. 2). Given the absence of a sampling frame and the challenges recruiting representative samples of key populations, sampling methods that allow for bias assessment and correction are frequently recommended, such as respondent-driven sampling or time-location sampling.<sup>11–13</sup> These methods may be effective at recruiting individuals into programs or interventions who are less readily reached,<sup>43</sup> and a limited but growing evidence base is emerging to compare the heterogeneity and reach of these approaches.<sup>11–13</sup> Even using these approaches, however, there will be individuals within the population who are less visible and potentially at greater risk if prevention efforts are not penetrating. These populations, such as FSW controlled by venue managers or MSM married to women, may have differing intervention needs than those “easily” accessible. Additionally, using program data and embedding IR within existing programming structures may foster broader generalizability of intervention effects but may be misinterpreted if used for health surveillance purposes. For example, HIV prevention programming data may underestimate HIV prevalence because those knowingly living with HIV are appropriately less likely to engage in HIV testing and HIV acquisition prevention activities. Key populations IR planning should also consider identification and double-counting issues because collection of identifiers from stigmatized and criminalized groups may be discouraged in certain contexts or members may provide false identities to protect themselves from criminalization or discrimination or to seek remuneration for research participation given underlying economic vulnerabilities.<sup>44,45</sup> Generation of unique IDs through replicable questions, or more recently through biometric data capture (eg, fingerprints, iris scanner), can help to address this, but has its own challenges.<sup>44</sup>

Executing any study requires adherence to clearly outlined standard operating procedures, but evaluation designs for key populations will need to be particularly flexible to rapid changes in legal, social, or programmatic context. Using iterative designs with smaller trials that work quickly to develop incremental evidence may offer solutions to leverage program infrastructure and ongoing key population engagement for IR while remaining nimble to frequent shifts in program and donor priorities.<sup>37</sup> Program data may also have increased rates of missing data, and trade-offs in data integrity and large-scale data should be evaluated, with additional



**TABLE 2.** Research Process, Implementation, and Implications for Inference

CFIR Construct	Research Process and Implementation	Implications for Inference
Engaging	Coordinating with implementing partner CAGs	Competing interests between program and research (external vs. internal validity) Perception of study within key population communities affects study participation Information and misinformation about study objectives or procedures can spread quickly through key population networks if CAGs are nonexistent, inactive, not representative of the larger population to be reached, or not fully aware of details
Planning	Study design Protocol development	Selection of sampling method will influence who is recruited Protocol needs to take into account real-world implementation structure Reimbursements can play a larger role among key populations in decisions to participate or attempts to doubly enroll
Executing	Study implementation	Training of study staff in <i>research methods</i> and how these differ from program implementation Want the research to be rigorous but do not want to interrupt the program itself Adaptability of objectives and methods based on changing program dynamics Assessment of fidelity of implementation perspective should capture lack of fidelity because of program/research teams vs. inability to find key populations for intervention delivery. This latter impact may have a larger impact on intervention effectiveness among key populations and also speaks to intervention appropriateness for all or some key population groups.
Reflecting and evaluating	Analysis Dissemination of results back to key stakeholders	Weighting/accounting for sampling biases and missing data (especially important if using program data) Measurement and reporting of implementation outcomes alongside effectiveness data to better inform reasons for why interventions succeed or fail Lack of research dissemination may affect future participation/representation. Furthermore, less top-down approaches to dissemination that include active community interpretation of results often provide invaluable insights to the final interpretation of results and potential drivers of intervention failure or success.

CAGs, community advisory groups.

efforts to supplement data gaps. However, program data may be less prone to social desirability bias, when community trust in programs and staff members is high. Accounting for loss to follow-up, missing data, or suspected overreporting/underreporting can be addressed through recalibration of estimates through bias validation samples (eg, pre-exposure prophylaxis adherence), inverse probability weights, and multiple imputation, and utilization of these approaches in HIV IR with key populations' programmatic data are urgently needed. Finally, evaluation should incorporate measurement of implementation outcomes beyond acceptability and adoption,<sup>46,47</sup> including fidelity (both attempted intervention delivery and participant receipt), feasibility (from an organizational and participant retention perspective), cost, maintenance, and penetration, alongside measurement of service and client outcomes (eg, efficiency, timeliness, effectiveness, equity, and satisfaction).

## CONCLUSIONS

The disconnect between the efficacy of HIV prevention and treatment interventions and the effectiveness of these interventions at the population level brings to focus the need for improved and focused implementation strategies.<sup>48</sup> Moving forward necessitates better addressing the needs of key populations who are often both at high risk of HIV acquisition and transmission and marginalized from and underserved by existing efficacious interventions. Although research and programs focused on serving key populations may be more complex and varies widely across locations and key populations, this work is central to the HIV response, independent of epidemic setting. Research that addresses or measures the embedded nature of individuals within their settings, prioritizes flexibility, and evaluates implementation barriers and strategies to overcome them for key populations, combined with programs that integrate consideration of these

elements may facilitate a more efficient HIV pandemic response. This nuanced HIV response is essential to close the gaps between efficacy, effectiveness, and ultimately population-level impact.

## ACKNOWLEDGMENTS

*The authors are grateful to the many key population members, community advisory groups, and local implementing partners who the authors have had the privilege to work with and learn from over the years. May your voices and efforts guide the next generation of implementation research.*

## REFERENCES

- Baral S, Beyrer C, Muessig K, et al. Burden of HIV among female sex workers in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Infect Dis*. 2012;12:538–549.
- Baral SD, Poteat T, Strömdahl S, et al. Worldwide burden of HIV in transgender women: a systematic review and meta-analysis. *Lancet Infect Dis*. 2013;13:214–222.
- Beyrer C, Baral SD, van Griensven F, et al. Global epidemiology of HIV infection in men who have sex with men. *Lancet*. 2012;380:367–377.
- Mathers BM, Degenhardt L, Phillips B, et al; for the 2007 Reference Group to the UN on HIV and Injecting Drug Use. Global epidemiology of injecting drug use and HIV among people who inject drugs. *Lancet*. 2008;371:1733–1745.
- Mountain E, Mishra S, Vickerman P, et al. Antiretroviral therapy uptake, attrition, adherence and outcomes among HIV-infected female sex workers: a systematic review and meta-analysis. *PLoS One*. 2014;9:e105645.
- Hayes RJ, Donnel DJ, Floyd S, et al. Impact of universal testing and treatment in Zambia and South Africa: HPTN071 (PopART). Conference on Retroviruses and Opportunistic Infections (CROI); Seattle, Washington; 2019.
- Wylie K, Knudson G, Khan SI, et al. Serving transgender people: clinical care considerations and service delivery models in transgender health. *Lancet*. 2016;388:401–411.
- Schwartz SR, Kavanagh MM, Sugarman J, et al. HIV viral load monitoring among key populations in low- and middle-income countries: challenges and opportunities. *J Int AIDS Soc*. 2017;20(suppl 7).
- Friedland BA, Sprague L, Nyblade L, et al. Measuring intersecting stigma among key populations living with HIV: implementing the people living with HIV Stigma Index 2.0. *J Int AIDS Soc*. 2018;21(suppl 5):e25131.
- Kennedy CE, Baral SD, Fielding-Miller R, et al. “They are human beings, they are Swazi”: intersecting stigmas and the positive health, dignity and prevention needs of HIV-positive men who have sex with men in Swaziland. *J Int AIDS Soc*. 2013;16(suppl 3):18749.
- Hakim AJ, Johnston LG, Dittrich S, et al. Defining and surveying key populations at risk of HIV infection: towards a unified approach to eligibility criteria for respondent-driven sampling HIV biobehavioral surveys. *Int J STD AIDS*. 2018;29:895–903.
- Rao A, Stahlman S, Hargreaves J, et al. Sampling key populations for HIV surveillance: results from eight cross-sectional studies using respondent-driven sampling and venue-based snowball sampling. *JMIR Public Health Surveill*. 2017;3:e72.
- Sabin KM, Johnston LG. Epidemiological challenges to the assessment of HIV burdens among key populations: respondent-driven sampling, time-location sampling and demographic and health surveys. *Curr Opin HIV AIDS*. 2014;9:101–106.
- Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. 2009;4:50.
- Pinto C, Bristowe K, Witt J, et al. Perspectives of patients, family caregivers and health professionals on the use of outcome measures in palliative care and lessons for implementation: a multi-method qualitative study. *Ann Palliat Med*. 2018;7(suppl 3):S137–S50.
- Davis SL, Goedel WC, Emerson J, et al. Punitive laws, key population size estimates, and Global AIDS Response Progress Reports: an ecological study of 154 countries. *J Int AIDS Soc*. 2017;20:21386.
- van der Elst EM, Gichuru E, Muraguri N, et al. Strengthening healthcare providers’ skills to improve HIV services for MSM in Kenya. *AIDS*. 2015;29(suppl 3):S237–S240.
- Duby Z, Fong-Jaen F, Nkosi B, et al. “We must treat them like all the other people”: evaluating the integrated key populations sensitivity training programme for healthcare workers in South Africa. *South Afr J HIV Med*. 2019;20:909.
- Shannon K, Strathdee SA, Shoveller J, et al. Structural and environmental barriers to condom use negotiation with clients among female sex workers: implications for HIV-prevention strategies and policy. *Am J Public Health*. 2009;99:659–665.
- Schwartz SR, Nowak RG, Orazulike I, et al. The immediate effect of the Same-Sex Marriage Prohibition Act on stigma, discrimination, and engagement on HIV prevention and treatment services in men who have sex with men in Nigeria: analysis of prospective data from the TRUST cohort. *Lancet HIV*. 2015;2:e299–306.
- Milcarek BI, Link BG. Handling problems of ecological fallacy in program planning and evaluation. *Eval Program Plann*. 1981;4:23–28.
- Reeves A, Steele S, Stuckler D, et al. National sex work policy and HIV prevalence among sex workers: an ecological regression analysis of 27 European countries. *Lancet HIV*. 2017;4:e134–e40.
- Cunningham S, Shah M. Decriminalizing indoor prostitution: implications for sexual violence and public health. *Rev Econ Stud*. 2018;85:1683–1715.
- Boily MC, Shannon K. Criminal law, sex work, HIV: need for multi-level research. *Lancet HIV*. 2017;4:e98–e99.
- Shannon K, Strathdee SA, Goldenberg SM, et al. Global epidemiology of HIV among female sex workers: influence of structural determinants. *Lancet*. 2015;385:55–71.
- Heffernan A, Cooke GS, Nayagam S, et al. Scaling up prevention and treatment towards the elimination of hepatitis C: a global mathematical model. *Lancet*. 2019;393:1319–1329.
- Dopson S, Fitzgerald L. The active role of context. In: *Knowledge to Action: Evidence-Based Health Care in Context*. Oxford: Oxford University Press; 2005.
- Damanpour F. Organizational innovation: a meta-analysis of effects of determinants and moderators. *Acad Manag J*. 1991;34:555–590.
- Dewar RD, Dutton JE. The adoption of radical and incremental innovations: an empirical analysis. *Management Sci*. 1986;32:1422–1433.
- Klein KJ, Conn AB, Sorra JS. Implementing computerized technology: an organizational analysis. *J Appl Psychol*. 2001;86:811–824.
- Macdonald V, Verster A, Baggeley R. A call for differentiated approaches to delivering HIV services to key populations. *J Int AIDS Soc*. 2017;20(suppl 4):21658.
- Kerrigan D, Kennedy CE, Morgan-Thomas R, et al. A community empowerment approach to the HIV response among sex workers: effectiveness, challenges, and considerations for implementation and scale-up. *Lancet*. 2015;385:172–185.
- Lyons CE, Ketende S, Diouf D, et al. Potential impact of integrated stigma mitigation interventions in improving HIV/AIDS service delivery and uptake for key populations in Senegal. *J Acquir Immune Defic Syndr*. 2017;74(suppl 1):S52–S9.
- Wheeler T, Kiran U, Dallabetta G, et al. Learning about scale, measurement and community mobilisation: reflections on the implementation of the Avahan HIV/AIDS initiative in India. *J Epidemiol Community Health*. 2012;66(suppl 2):ii16–25.
- Scheibe A, Howell S, Muller A, et al. Finding solid ground: law enforcement, key populations and their health and rights in South Africa. *J Int AIDS Soc*. 2016;19(suppl 3):20872.
- Baral SD, Wirtz A, Sifakis F, et al. The highest attainable standard of evidence (HASTE) for HIV/AIDS interventions: toward a public health approach to defining evidence. *Public Health Rep*. 2012;127:572–584.
- Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *Am J Prev Med*. 2007;32(5 suppl):S112–S118.
- Wheeler T, Wolf RC, Kapesa L, et al. Scaling-up HIV responses with key populations in West Africa. *J Acquir Immune Defic Syndr*. 2015;68(suppl 2):S69–S73.
- Escoffery C, Lebow-Skelley E, Haardoefer R, et al. A systematic review of adaptations of evidence-based public health interventions globally. *Implement Sci*. 2018;13:125.

40. Lambdin BH, Cheng B, Peter T, et al. Implementing implementation science: an approach for HIV prevention, care and treatment programs. *Curr HIV Res*. 2015;13:244–249.
41. UNAIDS. *Miles to Go—Closing Gaps, Breaking Barriers, Righting Injustices*. Geneva, Switzerland: UNAIDS; 2018.
42. Maheu-Giroux M, Diabate S, Boily MC, et al. Cost-effectiveness of accelerated HIV response scenarios in Cote d'Ivoire. *J Acquir Immune Defic Syndr*. 2019;80:503–512.
43. Baral SD, Ketende S, Schwartz S, et al. Evaluating respondent-driven sampling as an implementation tool for universal coverage of antiretroviral studies among men who have sex with men living with HIV. *J Acquir Immune Defic Syndr*. 2015;68(suppl 2):S107–S113.
44. Kavanagh MM, Baral SD, Milanga M, et al. Biometrics and public health surveillance in criminalised and key populations: policy, ethics, and human rights considerations. *Lancet HIV*. 2018. doi: 10.1016/S2352-3018(18)30243-1.
45. UNAIDS. *Developing and Using Individual Identifiers for the Provision of Health Services Including HIV*. Geneva, Switzerland: UNAIDS; 2009.
46. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*. 2011;38:65–76.
47. Smith JD, Benbow N, Brown CH, et al. A scoping review of recently-funded implementation research in HIV: examining characteristics of interventions and implementation research methods. In: Lyles C, ed. Johns Hopkins University Center for AIDS Research Implementation Science Symposium, Baltimore, MD, September 6, 2018; 2019.
48. Baral S, Rao A, Sullivan P, et al. The disconnect between individual-level and population-level HIV prevention benefits of antiretroviral treatment. *Lancet HIV*. 2019;6:e632–e638.
49. Zamudio-Haas S, Mahenge B, Saleem H, et al. Generating trust: programmatic strategies to reach women who inject drugs with harm reduction services in Dar es Salaam, Tanzania. *Int J Drug Policy*. 2016; 30:43–51.
50. Lambdin BH, Bruce RD, Chang O, et al. Identifying programmatic gaps: inequities in harm reduction service utilization among male and female drug users in Dar es Salaam, Tanzania. *PLoS One*. 2013;8: e67062.
51. Comins CA, Schwartz SR, Phetlhu DR, et al. Siyaphambili protocol: an evaluation of randomized, nurse-led adaptive HIV treatment interventions for cisgender female sex workers living with HIV in Durban, South Africa. *Res Nurs Health*. 2019;42:107–118.
52. Blumenthal J, Jain S, Mulvihill E, et al. Perceived versus calculated HIV risk: implications for pre-exposure prophylaxis uptake in a randomized trial of men who have sex with men. *J Acquir Immune Defic Syndr*. 2019; 80:e23–e29.
53. Young I, Flowers P, McDaid LM. Barriers to uptake and use of pre-exposure prophylaxis (PrEP) among communities most affected by HIV in the UK: findings from a qualitative study in Scotland. *BMJ Open*. 2014;4:e005717.
54. Han J, Bouey JZ, Wang L, et al. PrEP uptake preferences among men who have sex with men in China: results from a National Internet Survey. *J Int AIDS Soc*. 2019;22:e25242.
55. Cornish F, Campbell C. The social conditions for successful peer education: a comparison of two HIV prevention programs run by sex workers in India and South Africa. *Am J Community Psychol*. 2009;44: 123–135.
56. Jana S, Basu I, Rotheram-Borus MJ, et al. The Sonagachi Project: a sustainable community intervention program. *AIDS Educ Prev*. 2004; 16:405–414.
57. Wolf RC, Bingham T, Millett G, et al. Building the evidence base to optimize the impact of key population programming across the HIV cascade. *J Int AIDS Soc*. 2018;21(suppl 5):e25146.