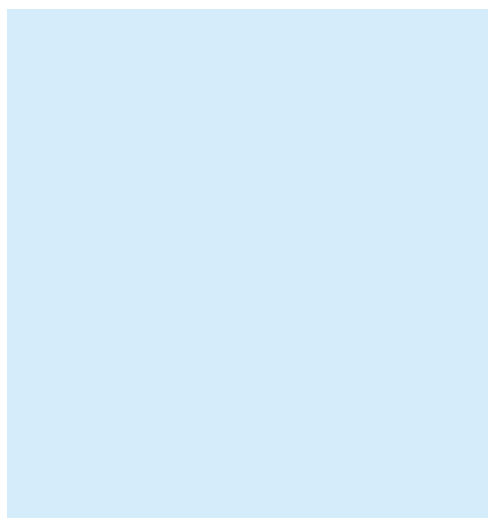
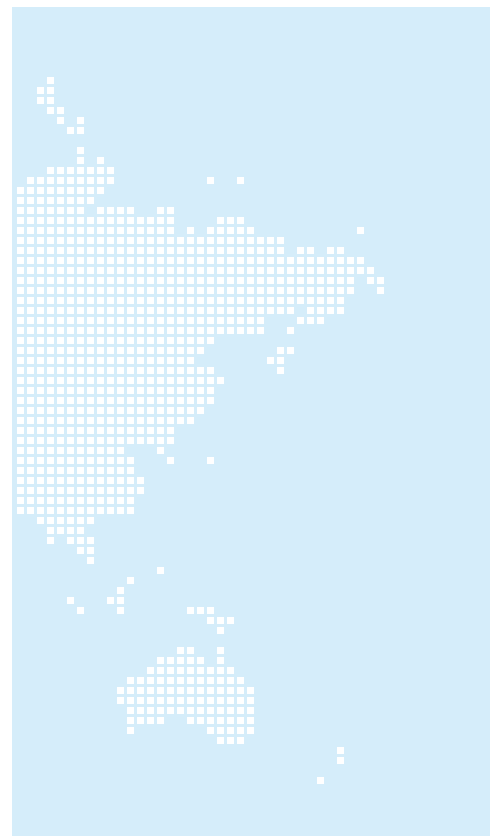


June 10

Strategic Guidance for Evaluating HIV Prevention Programmes



UNAIDS
JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS

UNHCR
UNICEF
WFP
UNDP
UNFPA

UNODC
ILO
UNESCO
WHO
WORLD BANK

MERG
MONITORING & EVALUATION
REFERENCE GROUP

Strategic Guidance for Evaluating HIV Prevention Programmes





Table of Contents

01 Acknowledgements

- 02 Purpose of the strategic guidance
- 02 Intended users of the strategic guidance
- 03 What the strategic guidance does **not** address

04 Acronyms

05 Executive Summary

- 05 Why focus on evaluation of HIV prevention now?
- 06 Recommendations for improving evaluations of HIV prevention programmes
 - 1. *Describe the programme impact pathway*
 - 2. *Determine what decisions need to be made and if an evaluation is needed and feasible*
 - 3. *Select appropriate measures to assess programme effects*
 - 4. *Assess programme implementation as well as programme effects, using mixed methods*
 - 5. *Focus on actionable results: a public health questions approach to HIV monitoring and evaluation*

10 Introduction

- 10 Why focus on evaluation of HIV prevention now?
- 14 Determining HIV prevention effectiveness: from research to practice
- 15 Professional standards in evaluation

18 Strategic Guidance for Evaluating HIV Prevention Programmes

- 18 Setting realistic expectations for monitoring and evaluation of HIV prevention programmes
- 21 Recommendations for improving evaluations of HIV prevention programmes
 - 1. *Describe the programme impact pathway*
 - 2. *Determine what decisions need to be made and if an evaluation is warranted and feasible*
 - 3. *Select appropriate measures to assess programme effects*
 - 4. *Assess programme implementation as well as programme effects, using mixed methods*
 - 5. *Focus on actionable results: a public health questions approach to HIV monitoring and evaluation*

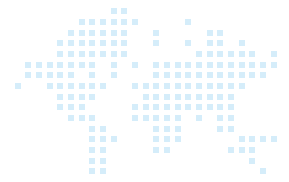
49 Conclusion

52 References

57 Appendix 1. M&E Glossary

60 Appendix 2. Programme Evaluation Standards

63 List of MERG Documents 2007–2009



Acknowledgements

This guidance was developed under the auspices of the UNAIDS-led Monitoring and Evaluation Reference Group (MERG), the global standard-setting body for monitoring and evaluation (M&E) in HIV and AIDS. During its 12th annual meeting (17–20 November 2008), the MERG called for guidance for evaluating HIV prevention programmes, taking into account the already substantial knowledge base and considering which methods would be most appropriate to the nature of the programmes and the specific evaluation questions asked.

A preparatory workshop was held in Geneva, Switzerland (26–28 May 2009) to learn from evaluation experience and to agree on specific evaluation issues to address¹. This was followed by a Think Tank held in Wilton Park, UK (1–4 September 2009) to develop consensus recommendations on evaluating HIV prevention programmes². More than 60 programme managers and evaluation experts from service implementing organizations, evaluation/research institutions, and international and donor agencies participated in these events.

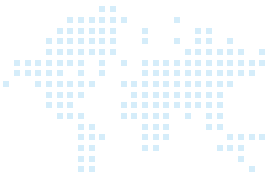
The strategic guidance was written, based on recommendations from the Think Tank event, by: Greet Peersman (Tulane University, New Orleans, USA); Marie Laga (Institute of Tropical Medicine, Antwerp, Belgium), and Deborah Rugg, Barbara deZalduondo, Eva Kiwango, Michael Bartos (UNAIDS, Geneva, Switzerland).

With thanks to the following individuals who provided input and/or comments [listed in alphabetical order]: Michel Alary (Centre Hospitalier affilié Universitaire de Quebec/Canada), Jane

Bertrand (Tulane University/USA), Patchara Benjarattanaporn (UNAIDS/Thailand), Gillian Callaghan (Durham University/UK), Sean Curtis (MEASURE Evaluation/USA), Roger Drew (DFID Consultant/UK), Peter Ghys (UNAIDS/Switzerland), Jean Pierre Habicht (Cornell University/USA), Susan Kippax (University of New South Wales/Australia), Douglas Kirby (ETR Associates/USA), Andrew Koleros (MEASURE Evaluation/Rwanda), Daniel Low-Ber (Global Fund/Switzerland), Sandy Oliver (University of London/UK), Nancy Padian (RTI International/USA), Rosalia Rodriguez-Garcia (World Bank/USA), Caroline Ryan (PEPFAR/USA), Naline Sangrujee (CDC/USA), Daniel Tarantola (University of New South Wales/Australia), Robin Vincent (PANOS/UK), Alice Welbourn (Salamander Trust/UK), and members of the HIV Prevention and the M&E Reference Groups.

¹ See: *Proceedings. Pre-Think Tank workshop on "Lessons Learned from Evaluation Experiences"*, Geneva (Switzerland), 26–28 May 2009. Geneva: UNAIDS, 2009.

² See: *Proceedings. Think Tank on "Evaluation Methods for HIV Prevention"*, Wilton Park (UK), 1–4 September 2009. Geneva: UNAIDS, 2009.



© N. Sorokin/Dreamstime



Purpose of the strategic guidance

Evaluation informs decision-making about sustaining, improving or discontinuing a programme and contributes to the local and international knowledge base on HIV prevention effectiveness. This document provides strategic guidance for evaluating the effectiveness of HIV prevention programmes considering issues of programme relevance and appropriateness, reach and coverage, quality, outcome/impact and cost-effectiveness. Many guidelines and tools about evaluation already exist but this guidance specifically addresses current challenges in evaluating HIV prevention programmes which aim to address HIV transmission through sexual intercourse and injecting drug use. The guidance is relevant to both generalized and concentrated/low HIV epidemics.

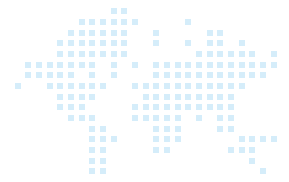
A glossary of key monitoring and evaluation terms used throughout the guidance is provided in **Appendix 1**.

Intended users of the strategic guidance

This guidance addresses what programme planners, managers and implementers (labelled “programme managers” for ease of reference) need to know about evaluation, and how evaluation links to programme monitoring. The extent to which large-scale national or subnational HIV prevention portfolios (i.e. the mix of projects, interventions and services) positively affect the HIV epidemic depends on how well their component parts are functioning. Thus, programme managers at all levels have an important role to play in maximizing programme effectiveness, even though they are often not conducting the evaluations themselves.

All managers of HIV prevention programmes need to:

- ▶ be able to identify what monitoring and evaluation activities are needed to guide programme management and improvement;



- ▶ understand how to implement and use routine input-output monitoring and process evaluations to ensure the programme is implemented as planned; and
- ▶ be able to identify what outcomes are appropriate for a particular programme and what evaluation design is best used to assess its effects based on what is already known about the programme and the decisions that need to be made.

Managers of specific projects, interventions or services also need to:

- ▶ understand the contribution of these programmes to the effectiveness of the subnational and national HIV prevention portfolio and the cumulative evidence base on HIV prevention.

Managers of national and subnational HIV prevention portfolios also need to:

- ▶ be able to interpret and use data to understand the HIV epidemic and to determine an appropriate and evidence-based HIV prevention portfolio to impact the epidemic; and
- ▶ be able to coordinate a national evaluation agenda focused on actionable results for improving HIV prevention programmes.

The guidance is also relevant to evaluators/researchers and to international and donor agencies to encourage a more unified and dynamic approach to evaluation in HIV prevention, grounded in country realities and focused on improving decision-making and practice.

What the strategic guidance does not address

The guidance does *not* address the strengths and weaknesses of specific HIV prevention interventions nor how to prioritize and adapt HIV prevention interventions to a specific HIV context. It does *not* address how to manage or conduct an evaluation

study nor how to document evaluation findings and practically apply them for programme improvement. These important topics are the subject of ongoing initiatives and/or other guidelines or tools developed by UNAIDS and its partners, including:

- ▶ Practical guidelines for intensifying HIV prevention.
- ▶ Taxonomy of HIV prevention activities.
- ▶ Developing Minimum Quality Standards for HIV Prevention Interventions.
- ▶ Organizing Framework for a Functional National HIV Monitoring and Evaluation System.
- ▶ 12 Components Monitoring and Evaluation System Assessment. Guidelines to support preparation, implementation and follow-up activities.
- ▶ 12 Components Monitoring and Evaluation System Strengthening Tool.
- ▶ Guidance on capacity building for HIV monitoring and evaluation.
- ▶ A Framework for Monitoring and Evaluating HIV Prevention Programmes for Most-At-Risk Populations.
- ▶ Monitoring and Evaluation Guidelines for HIV Prevention for People who Inject Drugs.
- ▶ Monitoring and Evaluation Guidelines for HIV Prevention for Men Who Have Sex with Men.
- ▶ Indicator standards operational guidelines.
- ▶ UNGASS. Monitoring the Declaration of Commitment on HIV/AIDS. Guidelines on Construction of Core Indicators. 2010 Reporting.
- ▶ Core Indicators for National AIDS Programmes. Guidance and Specifications for Additional Recommended Indicators.
- ▶ Guidance on developing terms of reference for HIV prevention evaluation.
- ▶ HIV triangulation resource guide. Synthesis of results from multiple data sources for evaluation and decision making.

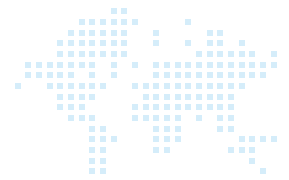
These documents are freely available at: <http://www.unaids.org> and/or <http://www.globalhivmeinfo.org/AgencySites/Pages/MERG%20UNAIDS%20ME%20Reference%20Group.aspx>



Acronyms

AIDS	Acquired immunodeficiency syndrome
ANC	Antenatal clinic
ART	Antiretroviral therapy
BCC	Behavior change communication
Global Fund	Global Fund to Fight AIDS, Tuberculosis and Malaria
HIV	Human immunodeficiency virus
IDU	Injecting drug use
KAB	Knowledge, attitude and behaviour
M&E	Monitoring and evaluation
MERG	Monitoring and Evaluation Reference Group
NASA	National AIDS spending survey
NHA	National health accounts
PEPFAR	United States President's Emergency Plan for AIDS Relief
PIP	Programme impact pathway
PMTCT	Prevention of mother-to-child transmission of HIV
RCT	Randomized controlled trial
STD	Sexually transmitted disease
STI	Sexually transmitted infection
SW	Sex worker
STARHS	Serologic testing algorithm for recent HIV seroconversion
UNAIDS	Joint United Nations Programme on HIV/AIDS
WHO	World Health Organization

Executive Summary



© UNAIDS/AVECC/H. Vincent



Why focus on evaluation of HIV prevention now?

HIV prevention remains one of the world's top public health and development priorities. Global efforts to control the AIDS epidemic cannot succeed without effective HIV prevention. There is no “magic bullet” solution, but combination prevention³ offers the best hope for successful HIV prevention and thus for sustainable AIDS treatment.

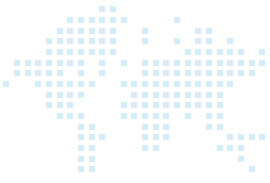
Though data from more and more countries show that HIV prevention has measurable population benefits, the evidence base for specific programmes is varied and incomplete. Thus, there is an urgent need to continue to accumulate credible evidence about what works and does not work to avert HIV infections in particular populations and settings, and to apply the lessons learned in programme practice.

In addition to scaling up HIV prevention interventions with known effectiveness, programme planners have to take the risk of implementing HIV prevention strategies of uncertain effectiveness. Evaluation is the only way to understand the programme's effects within the specific social and structural context and know how to improve on them.

Even if there are methodological challenges, programme managers can and must do better in maximizing the effectiveness of the HIV response by supporting appropriate evaluation and using the results for improving programmes at all levels.

A central dilemma in HIV prevention is whether to hold all programmes accountable for reducing

³ *Combination HIV prevention* is a dynamic, rights-based approach to providing the right mix of biomedical, behavioural and structural interventions aiming to have the greatest, sustained effort on reducing new HIV infections. The interventions are prioritized and informed by evidence and the wisdom and ownership of communities and tailored to meet local needs (UNAIDS Prevention Reference Group, 2009).



HIV incidence, the indicator of impact, when many programmes/programme components reduce risk factors and/or vulnerabilities rather than averting HIV infections directly, and when measuring HIV incidence is particularly challenging.

Recommendations for improving evaluations of HIV prevention programmes

1. Describe the programme impact pathway

Every programme manager should construct and regularly review the programme logic model or programme impact pathway drawing on existing evidence and theory, supplemented if necessary by new primary research. The programme impact pathway should describe the main components of the programme and how they are intended to work together to reach measurable objectives. Programme components that are not intended to avert HIV infections directly should be planned in concert with others to ensure that, together, they are accountable for significantly reducing new HIV infections.

2. Determine what decisions need to be made and if an evaluation is needed and feasible

Not all monitoring and evaluation activities are appropriate for all programmes or for the stage of development at which a programme happens to be at a given time. Expectations to conduct evaluation and choosing an evaluation design depends on what decisions need to be taken, the nature of the programme, and what is already known about the programme.

Every programme manager should use the programme impact pathway to determine, in collaboration with the programme's stakeholders, what decisions need to be made about the programme and what data and methods are most appropriate to support these.

3. Select appropriate measures to assess programme effects

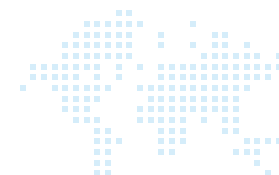
Each programme or programme component should be judged for its effectiveness in delivering the outcomes appropriate to its place in the causal chain towards averting HIV infections.

However, to justify any programme or programme component as essential parts of HIV prevention, it is critical for national programme managers to determine how the many component parts of the national programme "add up" to averting HIV infections.

4. Assess programme implementation as well as programme effects, using mixed methods

Programmes that are subjected to an outcome or impact evaluation should have implemented some level of process evaluation to identify any implementation problems which may negatively affect the programme's effectiveness and document important information for programme scale-up or replication elsewhere should this be warranted.

A combination of qualitative and quantitative mixed methods with nested designs and triangulation of different data sources (and if possible modelling) will most likely provide more complete information of HIV prevention



effectiveness than applying *one* method as a definitive gold standard.

5. Focus on actionable results: a public health questions approach to HIV monitoring and evaluation

Monitoring and evaluation activities differ in purpose and design but complement one another.

The utilization-focused approach focuses on eight basic questions which address key issues in the programme design and management cycle and reflect on the role of smaller-scale projects, interventions and services as components of a large-scale national or subnational HIV prevention portfolio.

Q1: What is the problem?

Purpose: To identify the nature, magnitude and course of the overall HIV epidemic and what population subgroups are most affected.

National programme managers need to ensure the following actions are undertaken:

- ▶ Continue to measure HIV prevalence in the general population using sentinel surveillance and 3–5 yearly nationally representative population-based surveys.
- ▶ Use the best available methods for estimating the size of most-at-risk populations and conduct regular surveys addressing HIV prevalence and behavioural trends.
- ▶ Apply appropriate modelling to estimate HIV incidence trends as well as HIV incidence by modes of transmission.

Q2: What are the contributing factors?

Purpose: To identify the contributing factors to the HIV epidemic and the determinants for vulnerability and risk for HIV infection.

Key steps for *national* programme managers are:

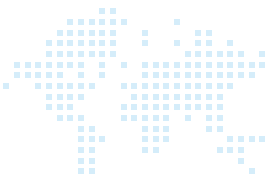
- ▶ Commission a multi-disciplinary group to engage with affected communities and to conduct an in-depth situational analysis of the HIV epidemic context, including the social factors that increase and protect against HIV risk and vulnerability.
- ▶ Commission additional determinants research where needed, using quantitative and/or qualitative approaches to seek a deeper understanding of the identified gaps and to identify strategies for change.
- ▶ Develop a working hypothesis of the underlying causes or drivers of risk and vulnerability and the likely pathways and social networks for influencing these. Identify stakeholders and “gate keepers” who influence the pathways and possible points of intervention at multiple levels.

Q3: What interventions can work (efficacy and effectiveness)?

Purpose: To determine what interventions might work under ideal circumstances and under specific field conditions.

Though programme managers are not conducting these evaluations themselves, they need to understand the following evaluation recommendations to be able to identify evaluation needs and oversee evaluation implementation and use:

- ▶ To evaluate a *mature, existing* programme (i.e. already part of routine practice) which has not been previously evaluated and use a mixed method approach. Include an experimental or quasi-experimental design if the programme has unknown effectiveness or is costly to imple-



ment in the population. Before embarking on this rigorous evaluation, prerequisites need to be fulfilled.

- ▶ To evaluate a *new* programme which has not been previously evaluated, use an experimental or quasi-experimental design if *any* of the following conditions apply: the programme has unknown effectiveness, or is risky politically or otherwise, or there is potential for negative effects. Before embarking on a full-scale evaluation, conduct a pilot study.
- ▶ To evaluate a *structural* programme: designing the programme should not be approached as a one-off event but closely integrated with continuing evaluation and reflecting an appropriate time scale. Include members of the audience in both the programme and evaluation teams and consider participatory evaluation methods to get maximum insight into internal and external influences. Adjust the programme according to the evaluation findings and continue to evaluate and adjust the programme throughout its life cycle.

Q4: What specific interventions and resources are needed?

Purpose: To determine what specific interventions are needed to address the local needs and what resources need to be available to implement them.

Programme managers need to:

- ▶ Plan structural interventions in an integrated way with needed biomedical and behavioural interventions at different levels to address the identified social drivers.
- ▶ Describe the *initial* programme impact pathway (PIP) including the selection of appropriate measures to monitor and/or evaluate the effects of the programme.
- ▶ The resources required for specific programmes

will be better informed the more all programmes and evaluations track their costs. Better collection of programme costs will allow cost-effectiveness analysis to be performed in the future.

Q5: What are we doing? Are we doing it right?

Purpose: To determine who is doing what, where, and with what intended outputs/outcomes.

Programme managers can improve routine monitoring by taking the following actions:

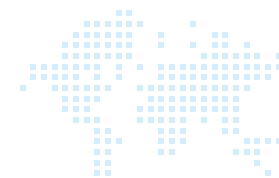
- ▶ Develop standards for output monitoring including data quality procedures to ensure sound routine monitoring, feedback and use of data at the point of collection.
- ▶ Supplement routine programme monitoring with data from surveys which include questions on programme exposure and with in-depth assessments of the quality of services provided.

National programme managers should also:

- ▶ Fund and support a responsible unit/person at the national level responsible for collating and analysing routine monitoring data for *all* HIV prevention interventions. They should engage with donors to ensure that pertinent monitoring data collected from donor-funded programmes are reported to the government.

Q6: Are we implementing the programme as planned?

Purpose: To determine whether the programme is implemented as intended and identify any problems in programme implementation for timely correction.



Every programme manager should ensure that:

- ▶ The programme (and key projects within it) is periodically assessed using process evaluation that examines fidelity to the programme design, the quality of services provided, client recruitment and retention, reach, intensity of programme delivered and received, client reaction/satisfaction, contextual changes, etc., to identify problems in implementation and design and take corrective action in a timely manner.

Q7: Are interventions working/making a difference?

Purpose: To determine if, and by how much, programmes achieved their intended outcomes.

National/subnational programme managers who need to make a decision about scaling up a specific programme:

- ▶ Should commission a retrospective study to establish whether the programme achieved its intended results in a similar context elsewhere (if available, systematic reviews may be useful here). External factors (such as social, cultural, economic or political factors, continuity of funding, etc.) that may affect the decision for scale-up or how to scale up, should be carefully considered before going ahead.
- ▶ If the efficacy and/or effectiveness of the programme is well established and there are no major external factors expected to affect the scale-up, then it may suffice to monitor if the programme is being scaled up according to plan and if it is still achieving its intended results.
- ▶ If there are important uncertainties about the programme's effectiveness in the context in which it is being scaled up, then a strong prospective evaluation should be commissioned. This should include rigorous qualitative and quantitative methods to collect context data, descriptive

programmatic data, data on the implementation of the programme, and behavioural and HIV prevalence data. To complement these data sources, special studies may need to be included to address any data gaps (e.g. assessing the effectiveness of particular programme components; testing specific assumptions in the programme impact pathway). Data from these different data sources should be analysed using triangulation methods. In most circumstances, a convergence of evidence provides sufficient evidence for a plausible link between the programme's operations and the observed results.

Q8: Are collective efforts being implemented on a large enough scale to impact the HIV epidemic (coverage; impact)?

Purpose: To understand national/subnational HIV trends and plausible association of results to the national/subnational HIV prevention programme.

National programme managers need to:

- ▶ Focus on national (and subnational where available) HIV trends, behaviours, determinants and the mix of HIV prevention programmes.
- ▶ Focus on data that are of critical importance for strategic planning and programme improvement.
- ▶ Focus on collecting a consistent, comparable data package and on analysis of determinants to understand both the programme and the context.

National programme managers need to ensure the following actions are undertaken:

- ▶ Collect a minimum package of repeat surveillance and survey data addressing epidemiology, behaviour and social and structural factors; apply strong analytic methods including data triangulation; and appropriately use modelling techniques.

Introduction



© UNAIDS/P. Virost

Why focus on evaluation of HIV prevention now?

HIV prevention remains one of the world's top public health and development priorities. In 2007, for every two patients entering antiretroviral therapy, five new HIV infections occurred (Merson et al., 2008). Hence, global efforts to control the AIDS epidemic cannot succeed without more intense and effective prevention programmes at country level. There is no "magic bullet" solution, but *combination HIV prevention* offers the best hope for successful HIV prevention, and thus, for sustainable AIDS treatment.

Combination HIV prevention is a dynamic, human rights-based approach to providing the right mix of biomedical, behavioural and structural interventions tailored to meet local needs in order to have the greatest, sustained effort on reducing new HIV infections (UNAIDS Prevention Reference Group, 2009). A combination HIV prevention approach does not mean doing everything for everyone; it means selecting appropriate interventions for the epidemiological and social context and the needs of those most at risk, prioritized and informed by evidence and the wisdom and ownership of communities. Combination HIV prevention is a human rights-based approach which aims to ensure that:

- ▶ the needs of the most affected, vulnerable and marginalized populations are addressed;
- ▶ these populations are empowered to have informed, active, free and meaningful participation in HIV-related decision-making processes;
- ▶ programmes are designed to achieve specific human rights-related objectives such as protection from sexual violence, gender equality, education, health, employment and access to scientific progress;
- ▶ there is equality and non-discrimination in programmatic expenditure and implementation which is carefully monitored by gathering information on service access by sex, age,

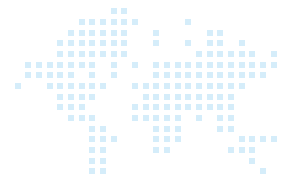
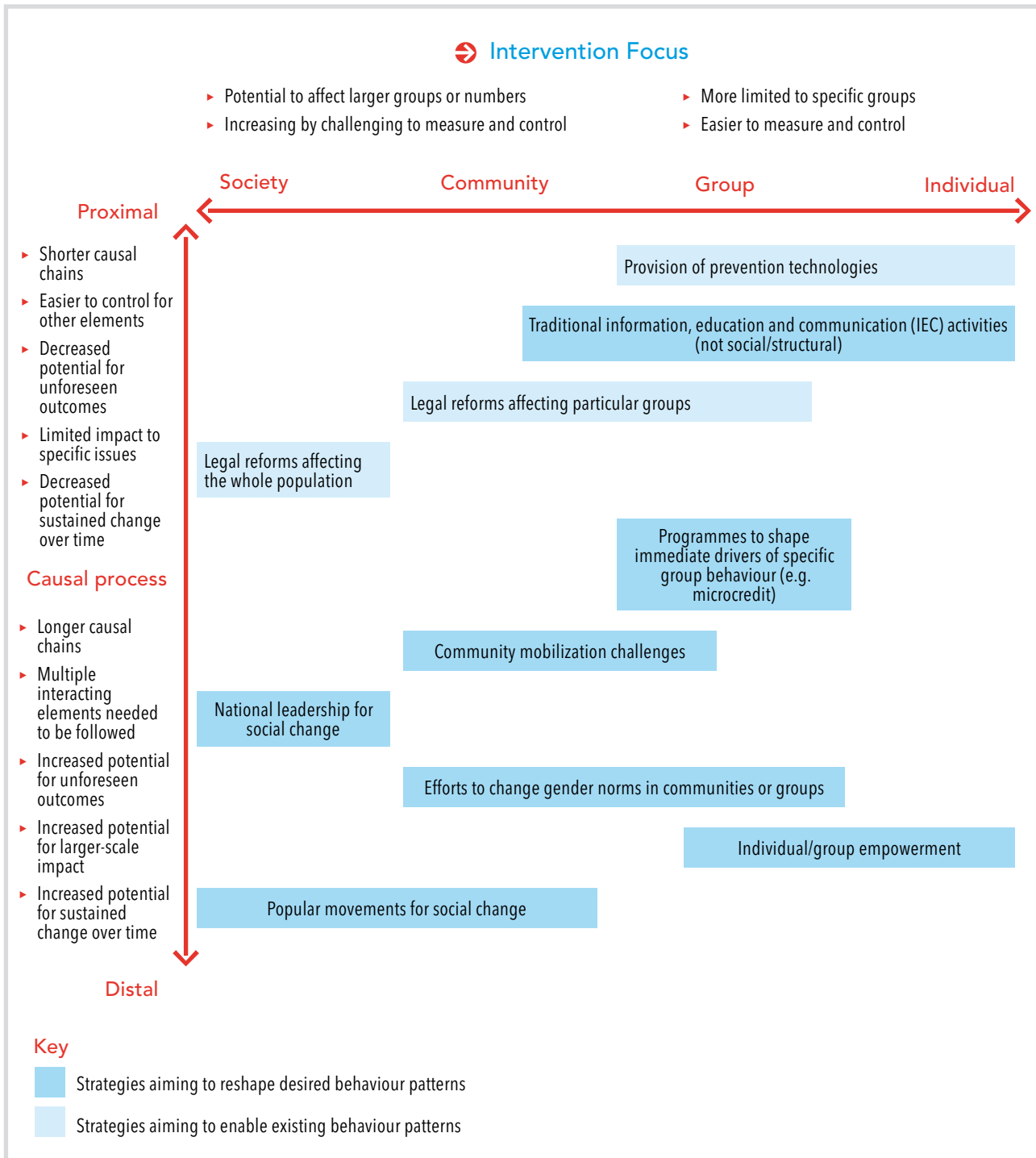
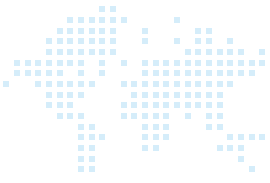


FIGURE 1.

Conceptual mapping of structural interventions based on the scale of the intervention (horizontal axis) and the level of intervention (vertical axis)

[Source: Auerbach J, Parkhurst J, Cáceres C, Keller K (2009). Addressing social drivers of HIV/AIDS. Some conceptual, methodological, and evidentiary considerations. Aids2031, Working Paper No 24.]





marital status, rural/urban status, economic status, language, and ethnicity; and

- ▶ accountability mechanisms are implemented in governments, intergovernmental organizations, donor agencies and the private sector.

(Adapted from the UNAIDS 2008 Report on the Global AIDS Epidemic)

The *combination HIV prevention* approach recognizes 25 years of science and programme experience which show that the behaviours and conditions that promote HIV transmission are socially embedded, and individual capacity to access and use prevention strategies is influenced by factors ranging from community norms to national laws and policies. Thus, averting new HIV infections depends on reducing both the immediate risks and the underlying drivers or causes of risk. For example, promoting knowledge and a desire to avoid HIV risks while acting to shift social norms and the broader social environment so that the behaviour change is accepted and supported. The *causal chains* – the hypothesized cascade of cause-effect relationships that explain final exposure to and infection with HIV will be different for different groups and settings and can include multiple linkages [Figure 1]. Since there are a range of strategies and actors involved in providing *combination HIV prevention*, there are additional benefits when prevention, treatment, care and support programmes are coordinated and reinforce one another.

The imperative to avert new HIV infections is clear, but support for HIV prevention is under threat due to competition for limited resources and broad claims that HIV prevention programmes are not working. Most of these claims are argued on the results from a handful of community randomized trials which have failed to show an intervention effect on HIV incidence (Kamali et al., 2003; Pronyk et al., 2006; Gregson et al., 2007; Ross et al., 2007; Cowan et al., 2008; Jewkes et al., 2008). While it is acknowledged that the evidence base for the

range of specific HIV prevention interventions is varied and incomplete, data from more and more countries show that HIV prevention has measurable population benefits [Box 1]. Nevertheless, there is a clear and urgent need to continue to accumulate credible evidence about what works and does not work in HIV prevention in specific settings. Evaluation is the only way to determine and understand a programme's effects and to know how to improve it.

When considering their investments, funders and managers of HIV prevention programmes need information to determine: whether their programme is implementing the right activities to overcome specific local barriers to HIV prevention (*Are we doing the right things?*); whether the activities are being delivered correctly (*Are we doing them right?*); and, whether the activities, collectively, are doing enough to reduce HIV incidence (*Are we doing them on a large enough scale?*). These are the basic questions programme evaluation should answer to improve programmes.

Many guidelines and tools about evaluation already exist, but this guidance specifically addresses current challenges in evaluating HIV prevention programmes which aim to address HIV transmission through sexual intercourse and injecting drug use:

- ▶ HIV prevention programmes are increasingly complex, multi-component and context-specific and the appropriate use of different evaluation methods needs to be clarified;
- ▶ The scientific evidence base to support the causal relationship between input/output and outcome/impact is incomplete and fragmented. The underlying behavioural theories leading to multiple behaviour changes and ultimately impact (i.e. reduction in HIV incidence) are difficult to assess;
- ▶ Many projects/interventions/services aim to affect HIV risk factors and/or vulnerabilities rather than averting HIV infections directly. The usefulness of different outcome measures

HIV prevention success in countries with high HIV prevalence

HIV prevalence and behavioural trends among 15 to 24 year-olds in countries with high HIV prevalence show positive effects of HIV prevention*.

[Source: UNAIDS. *Report on the global AIDS epidemic*. Geneva: UNAIDS, 2008b, p.34]

COUNTRY	Time period for which prevalence data were available	Prevalence trend ^a		Percent of young people (15–19 years) having had sex before age 15 ^b		Proportion having sex with more than one partner in the last 12 months ^c		Condom use during last sex among those with more than one partner in the last 12 months ^d	
		Urban	Rural	Females	Males	Females	Males	Females	Males
Angola ^a		ID	ID						
Bahamas		↓							
Benin ^b	2000–2006	↓*	→	→	↓*	↓*	↑	↑	↑
Botswana	2001–2006	↓*	↓*						
Burkina Faso	2000–2006	↓*	↓	↓*	↓*	↓*	↓*	→	↑*
Burundi	1999–2004	↓	→	↑*					
Cameroon ^a		ID	ID	↓*	↓*	↓*	↓*	↑*	↑*
Central African Republic ^b					↓				
Chad ^a		ID	ID	↓*	↑		↓*	↓	↑
Congo ^a		ID	ID						
Côte d'Ivoire	2000–2004	↓*	ID	↓*	↑	↓	↓*	↑	→
Democratic Republic of the Congo ^a		ID	ID						
Djibouti ^b									
Ethiopia ^b				↓*	↓*	↓*	↓*		↓
Gabon ^b									
Gambia ^b									
Ghana ^b				↓*	↓*	→	→	↑*	↑
Haiti ^b				↑	↑	↑	↓	↓	↑*
Kenya	2000–2005	↓*	↓*	→	→	↓*	↓*	↓	↑*
Lesotho ^a	2003–2007	→	↓						
Liberia ^b									
Malawi ^e	1999–2005	↓*	→	↓*	↓*	↑	↓*	→	↑
Mozambique ^f	2000–2007		→	→	↑*				
Namibia	2002–2006	↓	↓	→	↓*	→	↓*	↑*	↑
Nigeria ^a									
Rwanda	1998–2003	↓	ND	↑*	↑*	↓*	↓		
Sierra Leone ^a		ID	ID						
South Africa ^g	2000–2006		→			↓*	→		
Sudan ^a									
Swaziland	2002–2006	↓	↓						
Togo ^a									
Uganda ^b				↓*	↓	→	→	↑*	↑*
United Republic of Tanzania	2000–2006	→	↓	→	↓*	↓*	↓*	↑*	↑*
Zambia ^h	1998–2004		→	↓*	↓*	↓*	↓*	↑	↑
Zimbabwe	2000–2004	↓*	↓	→	↓*	↓*	↓*	→	→

Notes:

[1] Highlighted cells indicate positive trends in prevalence or behaviour.

[2] * Consistent sites only were used in the analysis of change in HIV prevalence over time, for a minimum of three years. Significance test based on $H_0: \text{slope} = 0$

Legend:

a Prevalence obtained from pregnant women attending antenatal clinics in selected countries.

b Among 15–19-year-olds, proportion reported having had sex by age 15. Analyses based on DHS, MICS or national surveys conducted between 1990 and 2007.

c Among 15–24-year-olds, proportion reported having had sex with more than one partner in the last 12 months. Analyses based on data from repeat DHS or national surveys conducted between 1990 and 2007.

d Among 15–24-year-olds, proportion of those with more than one partner reporting having used a condom the last time they had sex. Analyses based on data from repeat DHS or national surveys conducted between 1990 and 2007.

e Semi-urban and urban areas were combined in analysis of urban data.

f Analysis in Mozambique combined for South, North and Central.

g No data received in response to working group process; analyses based on data in South Africa surveillance report.

h No data received in response to working group process; analyses based on data reported in Zambia 2005 surveillance report. Analysis based on urban and rural data combined.

↑ Observed increase in HIV prevalence or behaviour.

↓ Observed decrease in HIV prevalence or behaviour.

↓* Statistically significant decrease in HIV prevalence of more than 25% or significant decrease in measured behavioural indicator. Analysis of prevalence based on regression analysis; analysis of behaviour based on Chi-square or Chi-square test for trend.

↑* Statistically significant increase in measured behavioural indicator. Analysis of behaviour based on Chi-square or Chi-square test for trend.

→ No evidence of change.

*ID Insufficient data, i.e. less than three years of data received.

**ND Data not received.

*HIV prevalence among pregnant women (2000–2007) in sentinel surveillance systems and selected sexual behaviours among women and men (1990–2007) from national surveys in all countries with a national HIV prevalence that exceeded 3% and four additional countries in Africa with notable prevalence levels.)

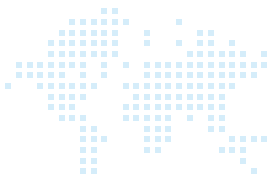
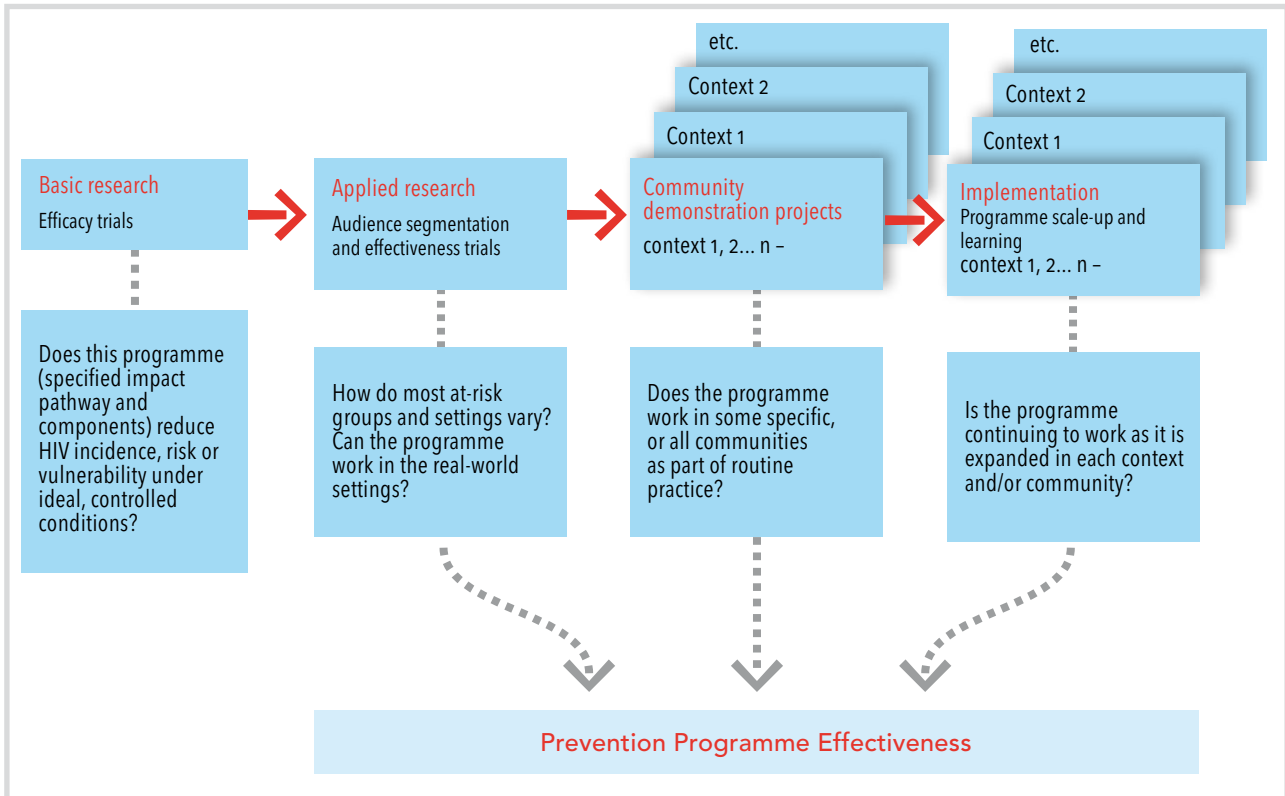


FIGURE 2.

Determining prevention effectiveness: the research to practice continuum

[Source: Adapted from Teutsch S. A Framework for assessing the effectiveness of disease and injury prevention. MMWR 1992;41 (No,RR-3)]



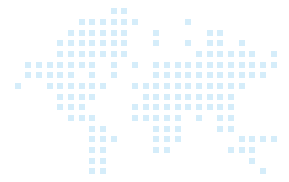
needs to be clarified as well as how they “add up” to averting HIV infections; and

- ▶ Measuring HIV incidence is particularly challenging. Currently, there are no ideal proxy measures for HIV incidence in populations.

Determining HIV prevention effectiveness: from research to practice

The effects of an intervention should be assessed at each stage of its development and implementation [Figure 2]. Ideally, the process begins with the development of the intervention based on

available evidence and appropriate theories of change. The intervention is then demonstrated to be efficacious (*Does it work?*) through research under carefully controlled conditions (i.e. in expert hands, fully resourced and under clearly defined conditions). As the intervention is applied at the community level, its effectiveness (*How well does it work in the real world?*) and cost can be assessed first in an applied research setting (i.e. for carefully selected target audiences) and then in community demonstration settings (i.e. as part of routine practice). Problems of access, follow-up, quality assurance and individual behaviour in the context of existing legal, health care and social systems are all elements of the evalua-



➔ **Important parameters in evaluation design are:**

Who is asking the question?

- > What do they want to know?
- > For what purpose?

What is the nature of the programme being evaluated?

- > What is the underlying programme logic?
- > What is the scope and size of the programme?
- > What is the maturity of the programme?
- > Was this programme or a similar programme ever evaluated? If so, what can we learn from the findings to improve the programme?

tion of effectiveness. Programme improvements should be incorporated iteratively as evaluation findings become available. Finally, evaluations of programme scale-up are conducted to determine if the programme continues to work under the conditions of widespread implementation.

This idealized model oversimplifies an iterative process. Often, there is pressure to move rapidly from basic and applied research to widespread implementation before appropriate evaluation studies can be completed. Consequently, there are often gaps in what is known about the efficacy, effectiveness, safety or economic impact of specific prevention strategies (MMWR 1992). Filling these gaps and maximizing our collective learning is what this strategic guidance is aimed at.

The evaluation methodology and designs need to adapt as one moves along the research to practice continuum based on the evaluation questions being asked, by whom, and for what purpose.

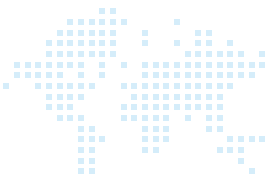
How complex and precise the evaluation must be, depends on who the decision-maker is and on what types of decisions will be taken as a consequence of the evaluation findings (Habicht et al, 1999). In addition, political, resource and time constraints as well as ethical considerations weigh in on the choice of evaluation method. Thus, methodology and context matter: not all methodologies are equally appropriate to answer a given evaluation question and no *one* methodology should be applied as a gold standard in all contexts (Julnes and Rog, 2007). Hence, the issue is: when, and under what circumstances, do various methodologies produce the most useful or actionable results?

Professional standards in evaluation

Evaluation does not stand alone as simply a logic or a methodology and it is certainly not free of values or interests. Rather, evaluation practices are firmly embedded in and inextricably linked to particular social and institutional structures and practices, which influence what is done within the study itself (House and Howe, 2000). Professional standards for programme evaluation were initially developed in the United States in 1975. Since then, they have been revised and adapted to different areas of investigation and specific local conditions. These standards are generally acknowledged to be good practice and should be routinely used in planning an evaluation, negotiating a contract to do an evaluation, and in reviewing progress during implementation of an evaluation.

The *Programme Evaluation Standards* address four main categories:

1. *Utility Standards* intend to ensure that an evaluation will serve the information needs of intended users.
2. *Feasibility Standards* intend to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.



3. *Propriety Standards* intend to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.
4. *Accuracy Standards* intend to ensure that an evaluation will reveal and convey technically adequate information about the features that determine the worth or merit of the programme being evaluated.

[American Evaluation Association, 1994. See Appendix 2 for a detailed overview of all standards]

Though not unique to HIV prevention, ethical conduct and due regard for the welfare of those involved in evaluation studies and those affected by their results are of utmost importance [Box 2]. Monitoring and evaluation must strike a balance between generating meaningful and useful information for programme managers while taking steps to ensure that data use does not worsen discrimination and stigma toward people who are HIV-positive (DeLay and Manda, 2004). Important ethical considerations include:

- ▶ Promoting social justice: Are the needs of those most disadvantaged addressed equally?
- ▶ Attending to procedural justice: Is there a commitment to autonomy and opportunity for meaningful input for a sufficiently broad range of stakeholders?
- ▶ Protecting study participants: Are the principles of respect for persons, beneficence, and justice honoured as guiding principles?

[Julnes and Rog, 2007]

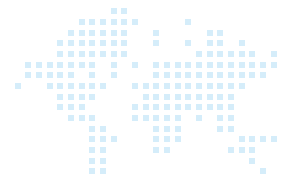
Another crucial matter in evaluation is that of participation of stakeholders – individuals, groups, or communities who have a decided stake or vested interest in the programme under evaluation. It can be assumed that many of those have minimal experience with and training in evaluation or formal methods of applied systematic inquiry (Cousins and Whitmore, 1998). Although still contested in

some quarters, the idea of stakeholder participation in evaluation is now widely accepted within the evaluation community and there is growing evidence that it improves the quality of the evaluation results (Whitmore, 1998). The purposes and meanings of participatory evaluation remain diverse and there are challenges to putting it into practice (including who should participate, assuring technical quality, dealing with objectivity and bias, resource constraints (especially time), ownership of the results, and differing evaluator roles).

Among the many forms of collaborative evaluation, *practical participatory evaluation* (also referred to as *utilization-focused evaluation*) is pragmatic and has as its central function the fostering of evaluation use. The core premise is that stakeholder participation will enhance relevance and ownership, and thus utilization of evaluation. This strategic guidance uses a utilization-focused approach based on the first, and perhaps most important, principle in M&E that data should be collected with the intention of being used. A truly utilization-focused M&E system will include closing the loop by evaluating actual data use and learning what factors enhanced use and what factors may have inhibited use, then using this learning to further enhance use (Patton, 1997).

The *how-to* process of stakeholder engagement in evaluation is outside the scope of this document, but it is certainly considered a critical factor. Many existing publications address participatory evaluation, including:

- ▶ Patton M.Q. (2007). *Utilization-focused evaluation. 3rd edition*. Thousand Oaks: Sage Publications.
- ▶ UNAIDS (2007). *Good participatory practice. Guidelines for biomedical HIV prevention trials*. Geneva.
- ▶ Whitmore E., ed. (1998). Understanding and practicing participatory evaluation. *New Directions for Evaluation* 80.



The UN International Guidelines on HIV/AIDS and Human Rights

The *UN International Guidelines on HIV/AIDS and Human Rights* (UN, 2006) promote the following safeguards in conducting HIV-related research and evaluation and use of study findings:

- Reforming public health laws to ensure that they adequately address public health issues raised by HIV, that their provisions applicable to casually transmitted diseases are not inappropriately applied to HIV, and that they are consistent with international human rights obligations.
- Enacting anti-discrimination and other protective laws that protect vulnerable groups, people living with HIV and people with disabilities from discrimination in both the public and private sectors, ensure privacy and confidentiality and ethics in research involving human subjects, emphasize education and conciliation, and provide for speedy and effective administrative and civil remedies.
- Enacting protective laws governing the legal and ethical protection of human participation in research, including HIV-related research, with specific attention to:
 - ▶ Non-discriminatory selection of participants (e.g. women, children, minorities);
 - ▶ Informed consent;
 - ▶ Confidentiality of personal information;
 - ▶ Equitable access to information and benefits emanating from research; and
 - ▶ Counselling, protection from discrimination, health and support services provided during and after participation.
- Establishing local and/or national ethical review committees to ensure independent and ongoing ethical review, with participation by members of the community affected, of the research project.
- Enacting general confidentiality and privacy laws.

See also *Guidelines for Privacy, Confidentiality and Security of HIV Information* (UNAIDS, 2008).

Strategic Guidance for Evaluating HIV Prevention Programmes



© Milosluz/Dreamstime

Setting realistic expectations for monitoring and evaluation of HIV prevention programmes

Monitoring and evaluation activities differ in purpose and design but complement one another. *Monitoring* provides information on *where* a programme is at any given time; it can provide a “snapshot” of the situation and programme status. *Evaluation* provides information about *whether or not* a programme is achieving specific objectives and *why* this is the case. Evaluation is intended to build on the findings from monitoring and provide additional information on the relevance and appropriateness, reach and coverage, quality, efficacy, effectiveness, and efficiency of specific programmes. A glossary of key monitoring and evaluation terms used in this document is provided in **Appendix 1**.

Not all M&E activities are appropriate for all programmes or for the stage of development at which a programme happens to be at a given time. However, *all* programmes are expected to participate in basic levels of M&E, including assessing needs and monitoring inputs and outputs once implementation begins. Expectations to conduct additional levels of M&E vary by the nature, size and maturity of the programme. Programme managers need to use their resources wisely, so the extent and costs of M&E activities should be commensurate to the size, reach and cost of the programme.

M&E should never compromise or overtake programme implementation. Based on experience, basic M&E should account for 5–10 per cent of the total programmatic budget. When rigorous special studies are to be conducted, 15–25 per cent of the programmatic budget may be needed.

Figure 3 reflects the expectations for M&E of programmes.

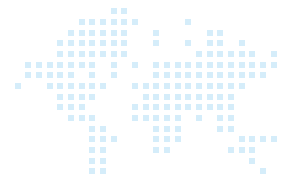
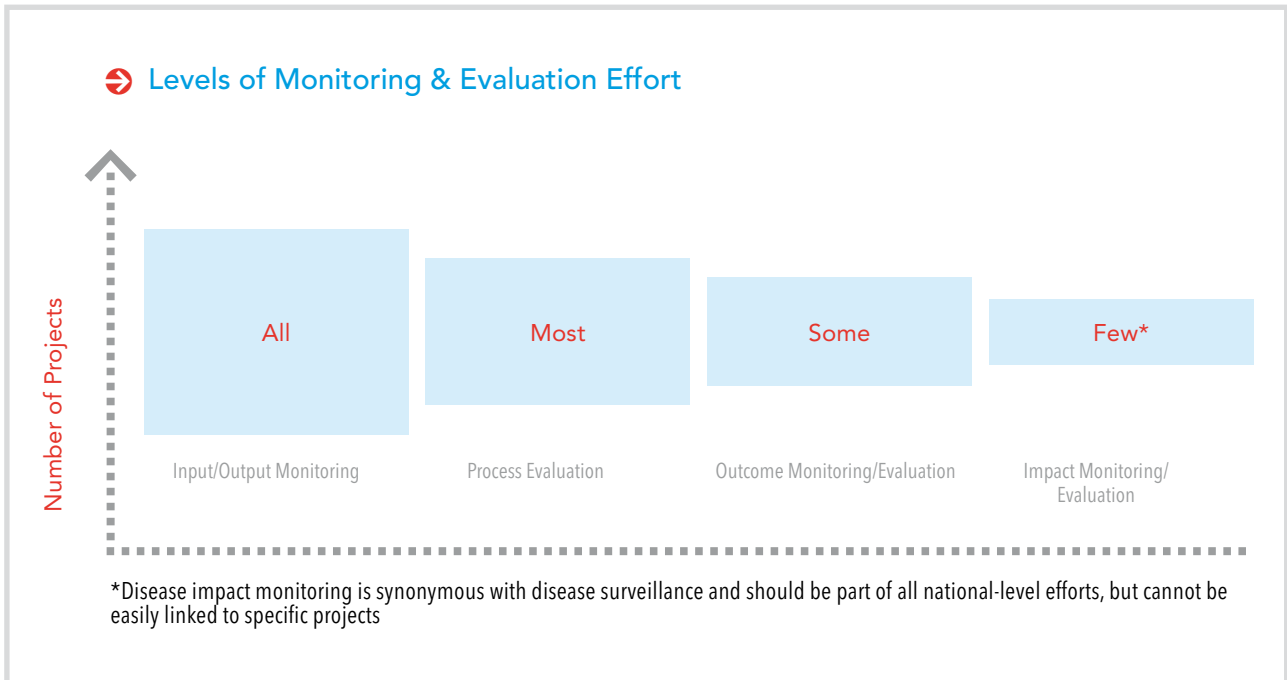


FIGURE 3.

Strategic planning for monitoring and evaluation: setting realistic expectations

[Source: Global AIDS Program, GAP. *Monitoring & Evaluation Capacity Building for Program Improvement. Field Guide*. Atlanta: Centres for Disease Control and Prevention, GAP, 2003; Rugg D, Peersman G, Carael M (eds). Global advances in HIV/AIDS monitoring and evaluation. *New Directions for Evaluation* 103, 2004]



- ▶ All programmes (national, subnational and service delivery levels) should conduct basic input and output monitoring for the purposes of good programme management and for selecting a few indicators to report to key stakeholders to whom the programme is accountable.
- ▶ Most programmes should also conduct process evaluations including implementation assessments, quality assessments, operations research, case studies and cost analyses.
- ▶ Only some programmes (usually the larger national or subnational programmes) will be able to conduct outcome monitoring and rigorous outcome evaluations, not only because of the additional time, expertise and resources these methods require, but also because they are only relevant to the more established programmes (outcome monitoring) or programmes for which there is insufficient evidence that they work (outcome evaluation) as they are new or simply have never been evaluated.
- ▶ Only in a few situations would impact evaluation be warranted in which an attempt is made to attribute long-term effects (impact) to a specific programme which is most often the result of collective effectiveness of all activities that constitute the national HIV response. These are usually done at national or subnational levels under the auspices of the government as they require large population sizes and considerable resources. Monitoring the unlinked distal impacts (impact monitoring) can feasibly be done through national surveillance systems and repeated population-based biological and behavioural

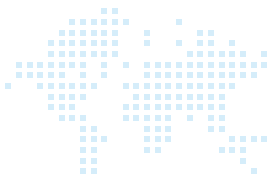


FIGURE 4.

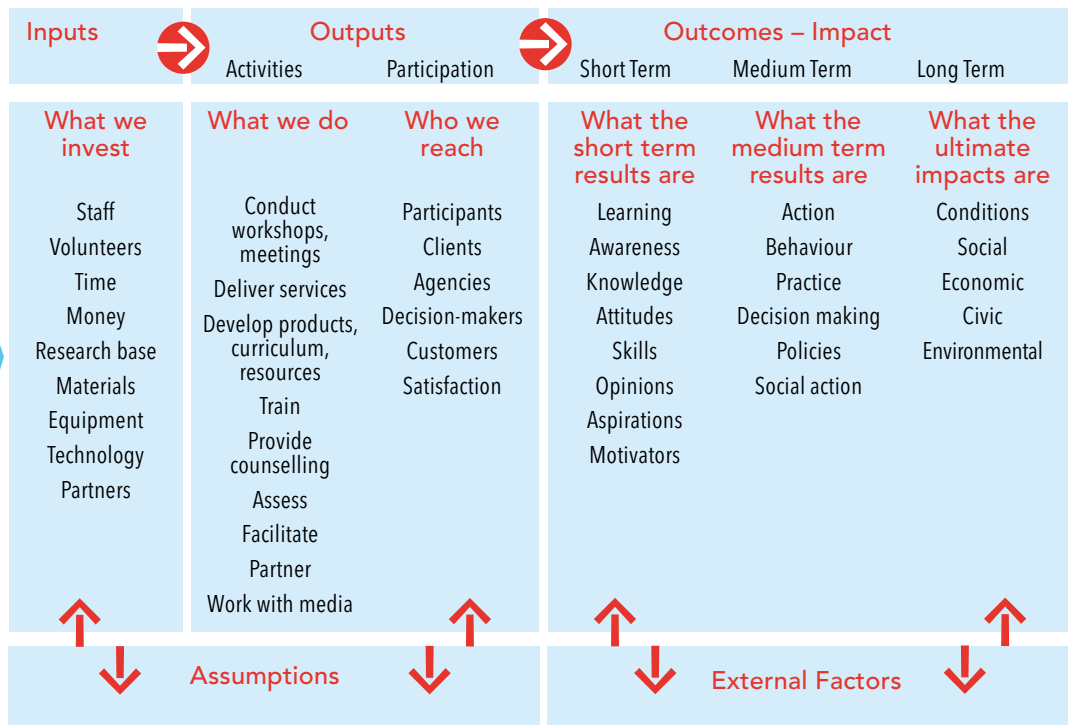
Key elements of a programme impact pathway (PIP)

➔ Programme Action – Logic Model

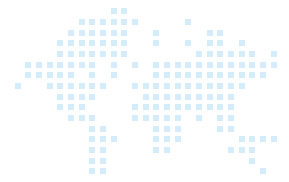
Situation
Needs and assets
Symptoms versus problems
Stakeholder engagement

Priorities
Consider:
Mission
Vision
Values
Mandates
Resources
Local dynamics
Collaborators
Competitors

Intended outcomes



Evaluation
Focus – Collect Data – Analyze and Interpret – Report



surveys using triangulation of multiple, existing data sources. It is important that long-term effects be interpreted in the context of results from process and outcome evaluations and from programme output monitoring data to ensure that findings are plausibly linked.

Recommendations for improving evaluations of HIV prevention programmes

The following five recommendations are considered key for improving the effectiveness of HIV prevention programmes. We believe that if they are more systematically and rigorously applied to programmes on the ground, they will not only benefit the programmes themselves, but will also contribute to our collective learning about what works and does not work in HIV prevention and why, thereby supporting more efficient use of resources. The recommendations are relevant to HIV prevention programmes in generalized as well as concentrated/low HIV epidemics.

1. Describe the programme impact pathway

Every programme manager should construct and regularly review the *programme impact pathway* (PIP) (also referred to as *programme logic model*) and use it throughout the design, implementation and evaluation of the programme. The programme impact pathway draws on existing evidence and experience with the programme to describe the main elements of a programme and how they are intended to work together to reach measurable objectives within the specific context. A programme impact pathway encompasses: (1) a programme impact theory which refers to the hypothesized cause-and-effect (not necessarily linear) pathways that connect a programme's activities to its intended outputs, outcomes and impact; (2) a service uti-

lization plan which relates to the assumptions of how and why intended recipients actually use the programme; and, (3) an organizational plan which relates to the implementation and operational aspects of the programme and its resources (Leroy et al., 2009). Thus, the programme impact pathway graphically presents the logical progression and relationship of the strategic programme elements (inputs, activities, outputs, outcomes, impact) and their causal relationships and the assumptions of risk that may influence success or failure of the programme [Figure 4]. Much of the benefit of constructing a programme impact pathway comes from the iterative process of discussing, analysing, and justifying the expected relationships between the different programme components and the feed-back loops.

➔ The recommended way forward

- ▶ Consult the literature and use available data bases with HIV prevention evidence as the foundational element in designing a programme.
- ▶ Construct a PIP drawing on existing evidence and theory, supplemented if necessary by new primary research. Ensure the programme design is informed by expert advice on the HIV, social, cultural and economic contexts and on the specific audiences that are likely to require different responses. Ensure the intervention approach adheres to human rights considerations including the participation of those infected/affected by HIV. If planning for a combination HIV prevention package, plan structural interventions in an integrated way with needed biomedical and behavioural interventions to reduce the specific causes of risk and vulnerability identified.
- ▶ The initial PIP is intended to represent the *ideal*, describing the way in which the programme is supposed to run and what results can be expected barring unexpected barriers and changes (i.e. if all goes as planned) [Figure 5].

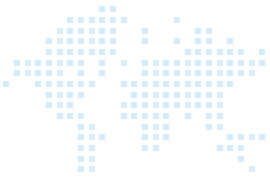
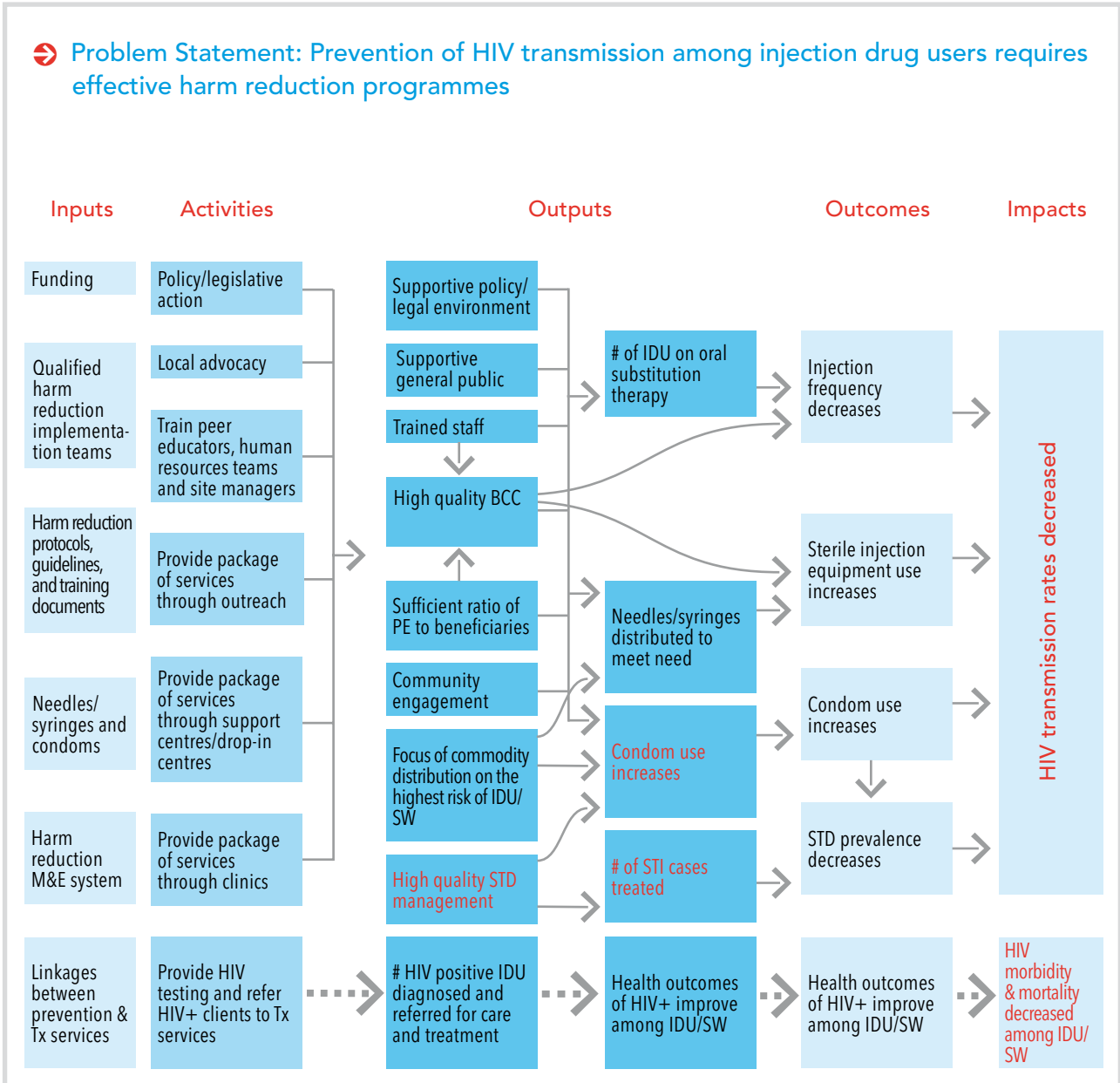


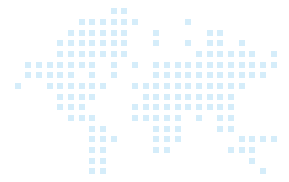
FIGURE 5.

**Programme impact pathway of a harm reduction programme:
Example from Viet Nam**

[Source: Viet Nam National AIDS Programme, 2009]

➔ **Problem Statement: Prevention of HIV transmission among injection drug users requires effective harm reduction programmes**





- ▶ Document any changes in the programme (such as changes in funding, shifting priorities and other stumbling blocks can lead to programme implementation and results that are different from what was intended) and update the PIP accordingly *during* and *after* programme implementation to describe what actually occurred and what results were achieved. Compare the planned and actual implementation PIP to help assess why differences may have occurred.

2. Determine what decisions need to be made and if an evaluation is warranted and feasible

The first and perhaps most important guiding principle for all M&E efforts is that information should be collected with the intention of being used. The main objective of evaluation is to influence decisions. Our basic assumption is that although the nature of what constitutes evidence and how it is applied in decision-making can be expected to differ, the ethos of being guided by evidence is strong.

Whether evaluation is *warranted* depends on what is already known about the programme (i.e. the level of uncertainty about its effects) and what programmatic decisions need to be taken. Whether evaluation is *feasible* depends on whether the programme is ready for evaluation and when programmatic decisions need to be taken, as well as whether adequate resources (human and financial) can be guaranteed to conduct the evaluation well.

➔ The recommended way forward

- ▶ Use the programme impact pathway to determine, in collaboration with the programme's stakeholders, the key questions about the programme and what decisions need to be made about the programme and when.

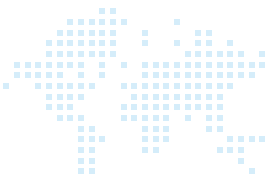
- ▶ Consult the literature and available evidence bases and determine if and what type of evaluation is needed to support decision-making. If evaluation is needed, secure adequate resources and develop terms of reference to guide the evaluation implementation and use.
- ▶ In case of a complex programme, a series of evaluation studies targeted at each of the key uncertainties in the programme design may be required to progressively refine the programme before embarking on a full-scale evaluation.

3. Select appropriate measures to assess programme effects

A central dilemma in HIV prevention is whether to hold all programmes accountable for reducing HIV incidence when many reduce HIV risk factors and/or vulnerabilities rather than averting HIV infections directly, and when measuring HIV incidence is particularly challenging. Typically, structural and social changes happen over longer timeframes than the current predominant short programme and funding cycles, demanding a new approach to programme planning, funding and evaluation.

➔ The recommended way forward

- ▶ Each programme or programme component should be judged for its effectiveness in delivering the outcomes appropriate to its place in the causal chain towards averting HIV infection.
- ▶ To justify any programme or programme component as essential parts of HIV prevention, it is critical for national programme managers to determine how the many component parts of the national programme "add up" to averting HIV infections. While there may not be a direct relationship between each of the component parts and HIV incidence, the programme manager needs to know if investment in these components is "reasonably associated with"



reduced HIV incidence at the population level. If not, they may be valuable activities in their own right, but the case is not being made that they are HIV prevention programmes.

4. Assess programme implementation as well as programme effects, using mixed methods

It is the combination of complementary data collection activities that help to answer the simple yet fundamental questions that must be answered in any public health response: “Are we doing the right things?”; “Are we doing them right?”; and, “Are we doing them on a large enough scale to make a difference?”.

Evaluations are often undermined by problems of acceptability, compliance and delivery of the intervention; recruitment and retention of study participants; and smaller than expected effect sizes. Thorough piloting and/or demonstration projects before embarking on a full-scale evaluation may help to reduce these problems (Craig et al., 2008). Based on the implementation of a sound programme impact pathway, it is unlikely that intended outcomes will be achieved unless a certain level of outputs is in place. Likewise, it is important to show that adequate outcomes have been achieved before starting to look for impact.

➔ The recommended way forward

- ▶ Programmes that are subjected to an outcome or impact evaluation should have implemented some level of process evaluation to examine: fidelity to the programme design, the quality of services provided, client recruitment and retention, programme reach, intensity of the programme delivered and received, client reaction/satisfaction, contextual changes, etc. This is critical to identify any implementation problems

which may negatively affect the programme’s effectiveness and to document important information for programme scale-up or replication elsewhere should this be warranted.

- ▶ Systematic collection of programme-related qualitative data assists in interpreting programme outcomes and impact and contributes to the understanding of what is or is not working and how to improve programme performance. Such information could also identify unexpected results and community perceptions that influence programme results but cannot be answered using indicator trend data alone.
- ▶ Using a combination of qualitative and quantitative mixed methods with nested designs and triangulation of different data sources (and if possible modelling) will most likely provide more complete information of HIV prevention effectiveness than applying one method as a definitive gold standard.

5. Focus on actionable results: a public health questions approach to HIV monitoring and evaluation

Eight basic questions serve to obtain a comprehensive understanding of the HIV epidemic and response [Figure 6]. They provide a simple and pragmatic way to organize the variety of data collection and analysis methods that need to be put in place to gather the right information and interpret it correctly. This utilization-focused approach addresses key issues in the programme design and management cycle and reflects on the role of smaller-scale projects, interventions and services as components of a large-scale national or subnational HIV prevention portfolios.

Each step in this investigative and analytic process is the foundation for the next step. However, in the real world, these steps are typically not conducted in the logical, sequential order laid out here.

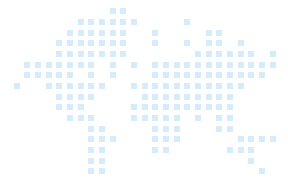
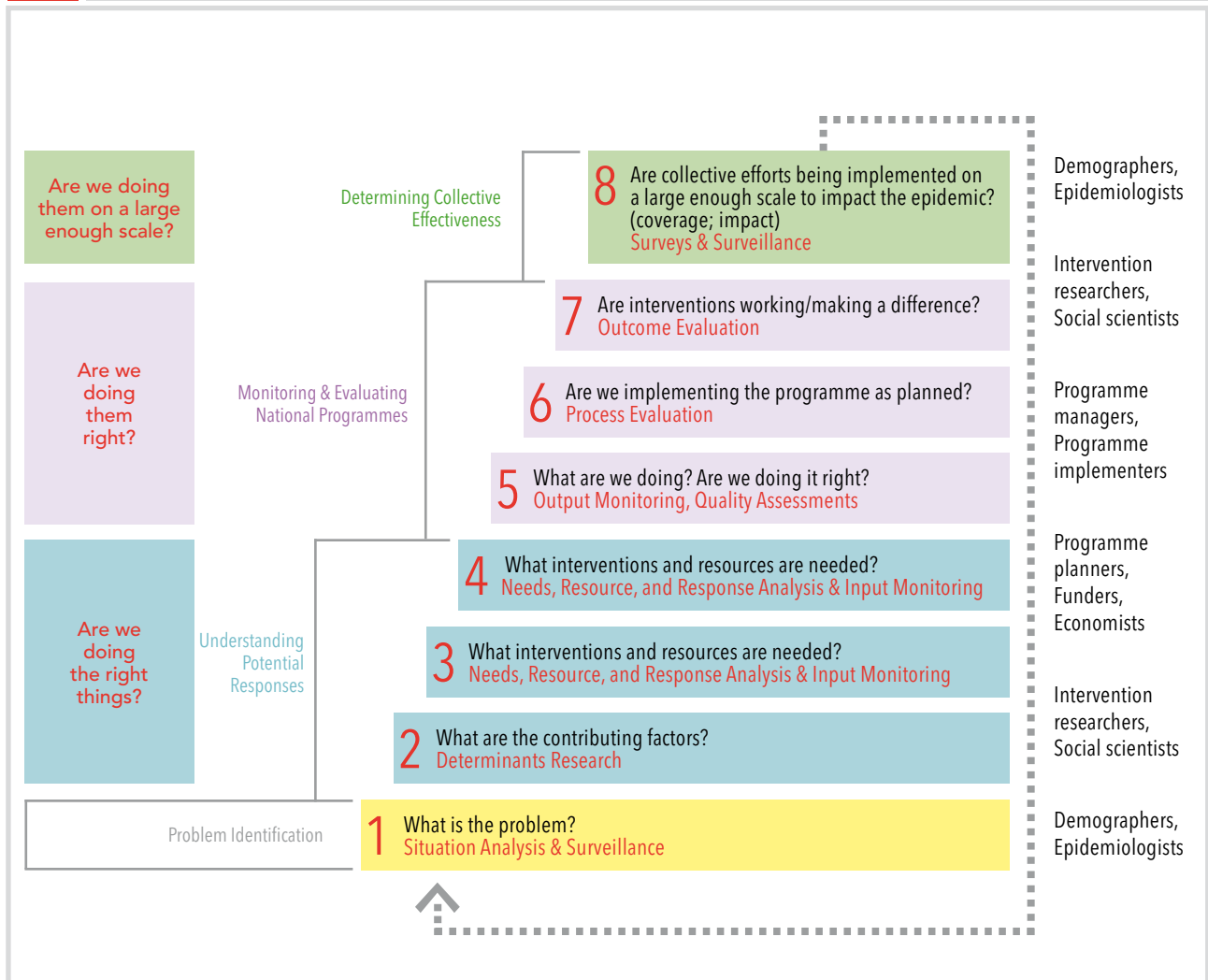
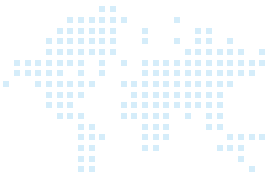


FIGURE 6.

A public health questions approach to HIV monitoring and evaluation

[Adapted from Rugg D., Carael M., Boerma J. T. and Novak J. Global Advances in HIV/AIDS Monitoring and evaluation: From AIDS Case Reporting to Program Improvement in Rugg D., Peersman G. and Carael M., eds. Global Advances in HIV/AIDS Monitoring and Evaluation, New Directions for Evaluation 103, 2004]





The basic message is that information in each of these areas is needed to obtain a *complete* picture of the HIV epidemic and its specific social and economic context and the effectiveness of the response. This is not a one-off effort but an ongoing systematic process of collecting, collating, analysing and using information to ensure the best possible programmes are put in place and identifying information gaps to ensure additional data are collected where needed.

Monitoring and evaluation involves a wide range of stakeholders (including programme implementers/service providers, programme beneficiaries, policy makers, funding agencies and the public at large) and a range of professionals play a role in its implementation (including epidemiologists, demographers, social scientists, programme planners and implementers, intervention researchers and economists) [Figure 6]. The following sections provide more detailed technical information for the professionals involved in implementing each step in the comprehensive monitoring and evaluation approach. Programme planners, managers and implementers (labelled “programme managers” for ease of reference) at all levels have an important role to play in maximizing programme effectiveness, even though they are most often not conducting the evaluations themselves. Their specific role in each of the steps is indicated.

Q1: What is the problem? What are the nature, magnitude and course of the HIV epidemic?

➔ The objective

- ▶ An effective national HIV response provides adequate HIV prevention information, services and support to those populations most likely to be exposed to and critical to the dynamics of the

epidemic and the response. Knowing the extent of HIV in various populations and geographical areas of the country (“Know Your Epidemic”) is key for planning the right mix of prevention strategies. This can be identified through second generation surveillance including national and subnational surveys, rapid assessments, participatory mapping of the HIV response and consultations with vulnerable populations and service providers (UNAIDS, 2008b).

HIV prevention programmes are intended to avert new HIV infections. To have an impact on HIV incidence, programme efforts must be directed to the appropriate populations and behaviours, in the appropriate locations and settings. Thus, programme managers should ideally have information on: HIV incidence in the country and any geographic variations within it; HIV incidence by different populations; the social and economic vulnerability of different populations (including gender, age, ethnicity and marginalized status); and trends in HIV incidence over time. In addition, data about the nature and magnitude of the programmatic response for each population are essential for identifying actual gaps between locations, scale and needs of the priority populations and the programmatic efforts currently underway to address these (“Know Your Response”).

➔ The experience so far

- ▶ HIV surveillance in designated sites (sentinel surveillance) has expanded and improved considerably, especially in sub-Saharan Africa and Asia, leading to more reliable estimates of the HIV epidemic and its impact, but important challenges with coverage and data quality remain (UNAIDS, 2008b). Many countries still lack the consistency required to follow trends over time in most-at-risk populations (Lyerla et al., 2008).
- ▶ HIV prevalence data collected in national pop-

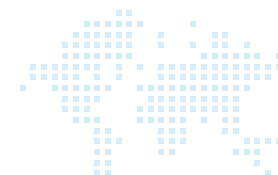
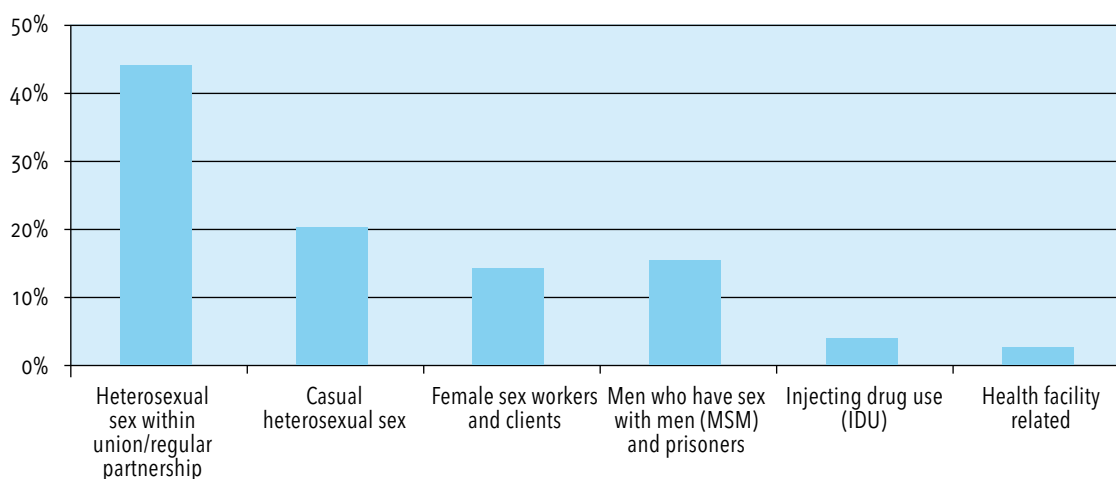


FIGURE 7.

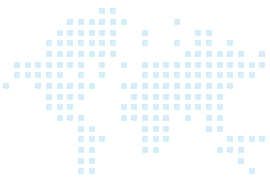
Studying the source of new infections to improve the focus of HIV prevention: Percent distribution of new infections by groups in Kenya

[Source: Kenya National AIDS Control Council, UNAIDS, World Bank. Kenya HIV prevention response and modes of transmission analysis. Final Report, March 2009. http://www.unaidsrstea.org/files/u1/Kenya_MoT_Country_Synthesis_Report_22Mar09.pdf]



Kenya's prevention programme has long defined its epidemic as generalized, based on previous second generation HIV surveillance, which demonstrated that since the late 1980s, Kenya has had more than 1% HIV prevalence in the general population in most parts of the country. A recent analysis, however, identified some data that could significantly improve the focusing of the prevention effort. This study estimated that a total of 76,315 new infections occurred in 2006 among the adult population aged 15–49 years. Nationally most new infections occurred in couples who engaged in heterosexual sex within a union/regular partnership, those who practise casual sex, are sex workers or are clients of sex workers, are among the prison population and men who have sex with men (MSM). Those who are in a union or regular partnership contributed 44.1% of new infections. Men and women who engage in casual sex contributed 20.3% of new infections, sex workers and their clients contributed 14.1% and MSM and prison populations contributed 15.2% of new infections. Generally the three main sources of new infections nationally and in the three

provinces are heterosexual sex in a union/regular partnership, casual sex and sex workers and their clients. These three categories contribute over 70% of new infections except in Nyanza province where they contribute over 90% of new infections. Injecting drug use (IDU) and health facilities contributed 6.3% of new cases (3.8% and 2.5%, respectively). The model estimates that the groups exhibiting the highest rates of transmission of infection are IDU (26%), prison population (13%), partners of IDU (8%) and MSM (7%). Although the number of cases in IDU is low, modelling results indicate that the incidence rate of the epidemic was highest among IDU at 256 per 1,000 followed by MSM in prison (126/1,000), partners of IDU (78/1,000) and MSM (67/1,000). This indicates that these populations are at high risk, the virus spreading among them at a very high rate compared to the other risk groups. Besides being due to very efficient transmission through sharing needles and anal sex, the high incidence rate among these groups may be an indicator of their marginalization and the lack of interventions directed towards them.



ulation-based surveys, in particular in countries with generalized epidemics, have improved the reliability of national HIV estimates (UNAIDS, 2008b).

- ▶ Direct measurement of HIV incidence through cohort studies is complex and expensive. At present, laboratory tests which aim to detect recent HIV infections (such as BED assay or STARHS⁴) are not recommended for routine surveillance applications or inclusion in national surveys, neither for absolute incidence estimates nor for monitoring trends (UNAIDS 2005). HIV prevalence remains the proxy measure for HIV incidence.
 - ▶ The models and assumptions in tools (such as the Estimation and Projection Package; WORKBOOK; Spectrum; the Asia Epidemic Model) to generate estimates (i.e. HIV prevalence over time, the number of people living with HIV, new infections, deaths due to AIDS, children orphaned by AIDS and treatment needs) are continually improved on the basis of latest available research (UNAIDS, 2008b). Estimates have been systematically calculated and are now available for most countries.
 - ▶ New modelling techniques aiming to estimate the number of new HIV infections by transmission category, are being applied to help countries prioritize their HIV prevention strategies better [Figure 7].
 - ▶ The systematic collection of standardized data about the HIV response, especially who is doing what and where in HIV prevention, is seriously lagging behind epidemiological analysis, hindering countries' ability to effectively set priorities for programme planning and resource allocation.
- ▶ Conduct sentinel surveillance among population groups representing the general population and groups with high-risk behaviour in an ongoing manner (every 1–2 years); regularly assess surveillance coverage and data quality and take corrective action.
 - ▶ Estimate the size of the identified most-at-risk populations and regularly update the estimates to ensure they are sufficiently accurate for service coverage planning.
 - ▶ Conduct nationally representative population-based surveys of the general population every 3–5 years, as appropriate to the epidemiological scenario in the country.
 - ▶ Routinely gather or tap data describing social and economic conditions.
 - ▶ Periodically conduct a data gap analysis to ensure additional data collection efforts can be put in place in a timely manner.
 - ▶ Apply appropriate modelling techniques to obtain estimates of the HIV prevalence and the rate of new HIV infections at national and subnational levels. Good modelling requires a minimum set of good quality data and includes explicit assumptions.
 - ▶ Regularly assess if the HIV prevention response and resource allocations match the epidemiological scenario(s).

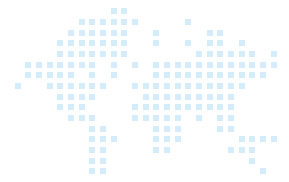
➔ The recommended way forward

National programme managers need to ensure the following actions are undertaken:

International organizations can help with:

- ▶ Continuing to improve methods for HIV incidence measurement.
- ▶ Continuing to improve the estimation models and expand training in their application.
- ▶ Providing a tool for describing the social, economic and political context of HIV programmes.
- ▶ Providing a glossary of HIV prevention activities to facilitate standardization of programmatic data collection and interpretation.

⁴ Serologic testing algorithm for recent HIV seroconversion.



Q2: What are the contributing factors and determinants of vulnerability and risk for HIV infection?

➔ The objective

Once we have information about *which* populations are most affected and *where* new infections are likely to occur, questions need to be answered about *why* these populations are most affected and *how* this occurred.

It is important to understand the relationships between the epidemiology of HIV infection, the risk behaviours that transmit HIV, and the complex environment of economic, legal, political, cultural and psychosocial factors that make people vulnerable to HIV infection. A national AIDS programme manager needs to determine which contextual factors for HIV risk and vulnerability are a priority, and what their social dimensions are.

This information is usually obtained from “determinants research” including both qualitative methods (e.g. rapid ethnography; participatory action research) to investigate and identify the relevant factors and to define them in local terms, and quantitative methods (e.g. knowledge, attitude, and behaviour (KAB) surveys; epidemiological risk factor studies) to measure the scale and distribution of the factors or determinants (see Vincent, 2009). People are socially embedded in particular social, cultural and economic contexts that give them and their behaviours meaning. These behaviours, meanings and contexts need to be understood when designing and evaluating prevention interventions. The results help to identify the multiple points of intervention required and to design appropriate and targeted intervention programmes that will remove biomedical, behavioural and structural barriers to safer sexual and health behaviour.

➔ The experience so far

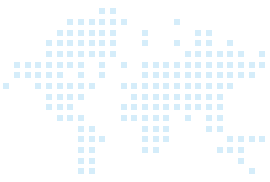
- ▶ In recognizing that social and structural factors influence HIV transmission by influencing attitudes and practices that can lead to infection, it is possible to intervene at the social level to alter those practices or the context in which they occur (Parker et al., 2000; Gupta et al., 2008; Auerbach et al., 2009)⁵.
- ▶ Social and structural factors operate at different levels (individual, interpersonal, community, institutional, legal/policy, public discourse and culture) and interact, but they have rarely been investigated and addressed in a concerted fashion in relation to HIV risk and vulnerability (Panos 2006). We need a better understanding of how these factors interact and can be exploited to synergistically support individual and societal change to avert new HIV infections.

➔ The recommended way forward

Key steps for *national* programme managers are:

- ▶ Commission a multi-disciplinary group to engage with affected communities and to conduct an in-depth situational analysis of the HIV epidemic context including the social factors that increase and protect against HIV risk and vulnerability. This should be informed by a broad review of the literature (including local qualitative research) to identify relevant social and structural issues at play [Box 3]. The review should pay special attention to age and gender differences, ethnic minorities, humanitarian emergencies and the

⁵ Structural factors are broadly defined to include physical, social, cultural, organizational, economic, legal or policy aspects of the environment that impede or facilitate an individual's efforts to avoid HIV infection (Sumartojo et al., 2000). Social factors are processes involving human relationships and influences among people including social roles, values, norms and institutions that structure social life, group practice and individual behaviour (Ingold, 1996). Economic factors are processes concerned with exchanges of resources and the organization of livelihoods. Political factors are matters of governance, decision-making and power. Issues of policy and legality relate to the formal regulation and codification of both political and economic factors. Cultural aspects emphasize meaning and the way differences in meaning become attached to aspects of social arrangements (Appadurai, 1996). All these processes are interrelated in ongoing social practice in any particular setting.

**BOX 3.****Example of social and structural influences on miners having unprotected sex in South Africa**

[based on Campbell, 2003]

Individuals are always socially located and embedded, since they are both shaped by and shape prevailing social norms and social practice which structures their dispositions and capabilities.

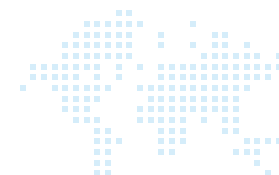
The following are pertinent influences on sexual risk behaviour of South African miners:

- **National gender norms and culture**
 - ▶ notions of masculinity imply men need regular sex and multiple partners
 - ▶ man needs to provide for family with paid work
 - ▶ masculinity involves sex standing for intimacy
 - ▶ miners' eclectic health beliefs – doubts over incurability of HIV
- **Characteristics of the mining industry**
 - ▶ migrant labour, miners living away from families in single hostels, no provision for families
 - ▶ gendered patterns of work include mining as a masculine job and restricted employment opportunities for women around mines
- **Characteristics of mining work**
 - ▶ dangerous, risky, arduous work, powerless work situation undermining sense of self-efficacy
 - ▶ relative economic power of miners in local economy compared to local women
- **Individual experience**
 - ▶ need to “forget” work and availability of alcohol
 - ▶ need for intimacy and comfort
 - ▶ intimacy means “flesh on flesh” and not using a condom
 - ▶ need to feel “in control” over others to compensate for lack of control in other aspects of life

inclusion or exclusion of people living with HIV and key populations such as sex workers, men who have sex with men, people who use drugs and prisoners.

- ▶ Commission additional determinants research where needed, using quantitative and/or qualitative approaches to seek a deeper understanding

of the identified gaps and to identify strategies for change. Methods such as rapid ethnography and participatory action research can help to examine the pathways linking social and structural factors (e.g. gender norms and legal frameworks) with access to and use of HIV services. Strategic studies can explore key social drivers in context



using whole system action research and positive deviance approaches can glean lessons from successful communities and programmes.

- ▶ Establish an expert group to advise on human rights, gender and other social factors in the national and local context, and to help segment audiences that are likely to require different responses. The group can track existing and emerging social research evidence and identify trends and priorities in social and cultural factors influencing particular sub-groups, including most-at-risk populations. They can then engage with programme managers to feed this information into (re)design of programmes.
- ▶ Develop a working hypothesis of the underlying causes or drivers of these social issues and the likely pathways and networks for influencing priority risk practices at the time. Identify “gate keepers” who influence the pathways and possible points of intervention at multiple levels. Map out causal chains that link social and structural factors to HIV transmission and the kinds of programmes that can influence different levels or links in each chain.

Q3: What interventions can work (efficacy and effectiveness)?

➔ The objective

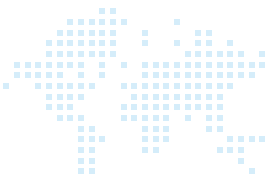
The focus in this evaluation step is on determining which interventions might work under ideal circumstances (in expert hands, fully resourced, and under controlled conditions) to establish efficacy, and subsequently, in its practical application in a real world setting (where practitioners may not be as expert, funds are usually less than ideal, and the intervention under evaluation is implemented as part of routine practice) to establish effectiveness. To facilitate the translation of research into practice, after efficacy and effectiveness studies have been done, demonstration projects are

needed to learn how to transfer lessons learned into routine practice.

When a programme has not been previously evaluated or there is limited evidence available in terms of its results, a rigorous evaluation study needs to be undertaken. How complex and precise the evaluation must be, depends on what types of decisions will be taken as a consequence of the findings and how confident the decision maker must be that any observed effects were in fact due to the intervention and not external factors (Habicht et al., 1999). Other important factors that realistically influence the choice of evaluation method include time and resource constraints, evaluation capacity and ethical considerations.

➔ The experience so far

- ▶ Behavioural change has been responsible for the prevention successes to date (Coates et al., 2008). Strategies to modify risk behaviours can focus on individuals, couples, families, peer groups or networks, institutions, and entire communities, and should ideally address multiple levels of influence at the same time. Several experimental and quasi-experimental studies and meta-analyses/systematic reviews have provided evidence of success for a range of behavioural strategies (see for example *Compendium of HIV prevention interventions with evidence of effectiveness*, http://www.cdc.gov/HIV/resources/reports/hiv_compendium/introduction.htm).
- ▶ Individual-focused concepts and methods of psychology and medical approaches have predominantly underpinned HIV prevention interventions and their evaluation, with an emphasis on individual behavioural change (Rugg et al. 2004; McKee et al. 2004; Coates et al. 2008). Recent recognition of the need for “combination prevention” in HIV responses, has led to renewed emphasis on the need to shift



the conditions that shape individual expectations, choices and behaviour with “structural” interventions. These include a broad range of activities – from national level policy reform to community dialogue to examine stigmatizing social norms. Additional impetus comes from the recognition that relative successes in some country responses to HIV, such as those in Uganda and Brazil, have involved widespread public communication, community mobilization and other contextual factors (Low-Beer and Stoneburner, 2004; Panos, 2006). The range of proven structural interventions – or more generally, evidence-informed strategies for social change that have reduced HIV incidence – is currently quite small. Developments in this arena have been hampered by lack of investment, weak theorizing, and methodological obstacles (Auerbach et al., 2009).

- ▶ Attempts to evaluate multi-component behavioural and/or social interventions using community randomized trials with HIV incidence as the impact measure have so far shown no effect (“flat results”), though positive behavioural outcomes were achieved (Kamali et al., 2003; Pronyk et al., 2006; Gregson et al., 2007; Ross et al., 2007; Cowan et al., 2008; Jewkes et al., 2008). The interpretation of these findings is difficult (i.e. is the intervention ineffective or was the method unable to detect an effect?) Possible methodological reasons include: poor adherence to the intervention under study, insufficient power, insufficient behaviour change to reduce HIV transmission significantly, the long pathway to biological impact, etc.) (Weiss et al., 2008). Nevertheless, important lessons were learned from these trials.
- ▶ For complex interventions, it is often not possible nor desirable to use experimental designs. Coker et al. (2004) make an interesting point: randomized trials make efforts to control for confounding and eliminate bias, yet programme implementers are seeking to understand con-

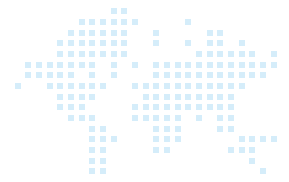
founding and live with confounding, because that is the day-to-day implementation reality. It is because trials control or “remove” these key variables that findings from such studies may not be viewed as relevant or applicable to people on the ground (Global Fund Background Paper, 2009) and may encounter difficulties in translating research into practice.

- ▶ In an ideal world, the programme manager can draw on an accessible and user friendly evidence base of efficacy and effectiveness studies of HIV prevention programmes and interventions. In reality, the programme manager may operate in uncertainty, because of methodologic challenges and gaps in evaluation, external validity or replicability of the results, and scarcity of interpretable data on the evaluation of complex HIV prevention programmes.
- ▶ Systematic reviews of existing evidence are often too narrowly focused on statistical analysis of experimental/quasi-experimental studies and fail to address external validity or replicability, or to draw lessons learned from a rigorous analysis of the range of methodologies used to elucidate what works in HIV prevention, how and why.

➔ The recommended way forward

Though programme managers are not conducting these evaluations themselves, they need to understand the following evaluation recommendations to be able to identify evaluation needs and oversee evaluation implementation and use:

- ▶ A clear and explicit rationale for the evaluation design and use based on the programme impact pathway (PIP) is essential. If the programme is complex, consider if the package of intervention components is to be evaluated together and/or by its individual components.
- ▶ To evaluate a *mature, existing* programme (i.e. already part of routine practice) which has



not been previously evaluated, use a mixed method approach. Include an experimental or quasi-experimental design if the programme has unknown effectiveness or is costly to implement in the population. Before embarking on this rigorous evaluation, each of the following prerequisites needs to be fulfilled: (1) the programme's characteristics are consistent with the most up-to-date information regarding the nature of the local epidemic; if it is not, the programme content needs to be revised; (2) the quality of programme implementation, coverage and duration are considered sufficient to expect a change in the specified outcomes for the evaluation; and, (3) check if any existing behavioural and biologic data are already available that can be used to compare the programme results over time.

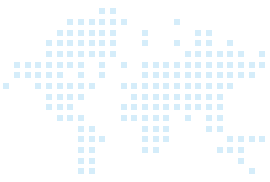
- ▶ To evaluate a *new* programme which has not been previously evaluated, use an experimental or quasi-experimental design if *any* of the following conditions apply: the programme has unknown effectiveness, *or* is risky politically or otherwise, *or* there is potential for negative effects. Before embarking on a full-scale evaluation, conduct a pilot study to obtain an understanding of the acceptability of the new programme, compliance with and delivery of the programme, participant recruitment and retention issues, the appropriateness of the outcome measures selected, and the anticipated effect size.
- ▶ Because of their increased complexity in design, implementation, analysis and interpretation, a *community* randomized trial should only be considered if there are: a well-defined, narrow hypothesis, so that the key to the intervention success or failure can be identified; *and* a measurable intervention, to be able to assess its implementation; *and* adequate statistical power; *and* well-defined, measurable outcomes (Susser, 1996). These considerations also apply to quasi-experimental designs.

- ▶ Most experimental and quasi-experimental designs require experienced evaluators for their implementation, analysis and interpretation. Some important considerations in these designs are provided in **Box 4**.

To evaluate a *structural* programme:

All the salient social factors that influence risk and vulnerability for HIV infection can never be completely known in advance. Thus, designing a structural programme should not be approached as a one-off event but closely integrated with continuing evaluation and programme adjustment in an iterative and participatory process. As experience with these evaluations is in its infancy, more detailed information on the key steps involved is provided here:

- ▶ Use the working hypothesis of the underlying social drivers and the *initial* programme impact pathway (PIP) (see above) to describe the kinds of conditions and interventions that are needed to influence the causal chain.
- ▶ Select some measures to evaluate the implementation as well as the effect of the programme, and the context in which it is being implemented. By engaging different stakeholders in this process, the selected measures are more likely to be the most relevant and important ones focused on different levels.
- ▶ Include members of the audience in both the programme and evaluation teams, and select participatory evaluation methods such as action research, appreciative enquiry, organizational learning, participatory evaluation, realistic evaluation, utilization-focused evaluation, social network analysis and story-based approaches of significant change (see for example, Henry et al., 1998; Whitmore, 1998; Ryan and DeStefano, 2000; Preskill and Coghlan, 2003; Davies and Dart, 2005).
- ▶ Use a range of methodologies (e.g. social and

**BOX 4.****Considerations in experimental and quasi-experimental designs**

Experimental and quasi-experimental studies should be preceded by formative research, a good programme design and explicit programme impact pathway. They should include an extensive process evaluation to fully understand the implementation of the programme and how it may affect expected outcomes.

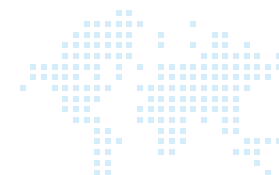
Randomization should be carefully considered in terms of:

- *Size and timing of effects:* Randomization may be unnecessary if the effects of the intervention are so large or immediate that confounding or underlying trends are unlikely to explain differences in outcomes before and after exposure. It may be inappropriate – for example, on the grounds of cost or delay – if the changes are very small or take a long time to appear; in these circumstances a non-randomized design may be the only feasible option.
- *Likelihood of selection bias:* Randomization is needed if the exposure to the intervention is likely to be associated with other factors that influence outcomes. Post-hoc adjustment is a second best solution: its effectiveness is limited by errors in the measurement of the confounding variables and the difficulty of dealing with unknown or unmeasured confounders.
- *Feasibility and acceptability of experimentation:* Randomization may be impractical if the intervention is already in widespread use, or if key decisions about how it will be implemented have already been taken, as is often the case with policy changes and interventions whose effect on health is secondary to their main purpose. On the other hand, if an intervention cannot be rolled out everywhere at once either because of limited resources or absorptive capacity, randomization is often the fairest way to determine who gets the intervention and when.

- *Cost:* If an experimental study is feasible and would provide more reliable information than an observational study but would also cost more, the additional cost should be weighed against the value of having better information.

[Source: based on Craig et al., 2008]

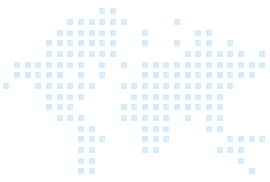
Experience with a range of alternative, adaptive designs and with sophisticated analysis techniques is growing and should be drawn on (see for example: Brown and Lilford, 2006). The mention of randomized trials often congers up visions of studies which are extraordinarily expensive, of long duration, complex and require detailed training for those who conduct them. However, there is a range of randomized approaches that are similar only in that beneficiaries of the intervention are selected randomly. Almost all other aspects of implementation differ: they are only marginally more expensive than a non-randomized approach; they can provide rapid feedback so that the intervention can be altered over the course of the study and they require no more training than what is required for programme implementation.



Considerations in experimental and quasi-experimental designs

These include:

- > *Large Simple Trials* (Peto et al., 1995): These trials typically include a large number of people (several thousand) and extend over a long period of time. They have broad eligibility criteria, simple enrollment procedures, collect minimal data, and use clearly defined, easy-to-assess outcomes as endpoints. They are intended to study interventions under “real world” conditions.
- > *Stepped Wedge Designs*: Rather than randomize based on intervention, individuals or communities are randomized based on time. Data are collected among all at baseline. Beneficiaries are randomized to receive the intervention at different defined time points over the course of the study so that at the end, all individuals receive the intervention. Assessments are made between those who do and do not receive the intervention at each time point. Because these assessments might provide information that suggests improving the intervention, the intervention is “better” at each time point, so those who are randomized to receive the intervention later receive a more robust intervention.
- > *Randomized Promotion or Encouragement Designs* (Bradlow, 1998): A randomly selected sub-sample of the beneficiaries of the intervention is selected and receives additional promotion, encouragement or incentives to participate. Such incentives can include: information, encouragement (small gift or prize), transport or the incentives. Almost certainly the promoted group will have higher enrollment or adherence to the programme. Likewise, those who do not receive the promotion receive a more diluted version of the programme and thus function as the control group. The fundamental assumption is that the promotion cannot directly affect the outcome.
- > *Adaptive Randomized Designs* (Bauer and Brannath, 2004; Chang, 2009): These are flexible designs that permit mid-trial modifications without compromising the ultimate statistical assessment of results. For example in “Drop-the-Loser”, multiple arms are compared to a control. Interim results suggest that some arms are inferior and they are dropped from the study. This is a particularly appealing design for combination HIV prevention where each component can be assessed independently and those that hold less promise are dropped.
- > *Hybrid studies*: It is also possible to have combinations of implementation designs. A randomized design in certain regions or age groups, for example, may be embedded within a nationwide campaign. Furthermore the randomized designs themselves may be combined. In particular, the principles of adaptive designs with frequent feedback and alterations in the intervention can be applied to all of these designs.



sexual network analysis; path analysis) from the behavioural and social sciences, in addition to traditional epidemiological methods, to assess the importance of different causal and mediating social factors and refine the programme impact pathway accordingly.

- ▶ Adjust the programme and adapt the evaluation throughout the programme's life cycle. It is particularly important to document what has changed in practice and any ways in which the programme and its implementation have unfolded differently from the *initial* programme impact pathway. Thus, the programme impact pathway should also evolve throughout this learning process. Document unexpected consequences, positive and negative, and collateral effects (e.g. on economic security, education, maternal and child health, social justice). Use the social research expert group to help document and interpret the findings for programme improvements and for adding to the social research evidence on social and structural factors influencing HIV transmission.
- ▶ Include key social and structural variables in HIV and related surveys, or glean the needed data from other sources (e.g. national economic surveys, census). Use these and bio- and behavioural survey data over time to track the influence of social and structural factors on individual attitudes, perceived social norms, content of influential media and communications, and other intermediate outcomes. Measure changes at the individual, group and organizational level that figured in the causal chain or programme impact pathway. Consider using Qualitative Comparative Analysis methods (Byrne, 2002) to identify social "control parameters" for priority policy action (Vincent, 2009).

International organizations can help with:

- ▶ Creating a compendium of programme impact pathways of structural interventions as an aid

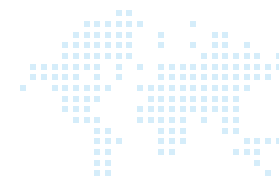
in the development and implementation of such interventions and a repository of the evaluation methods and results of structural interventions.

- ▶ Adding to the existing meta-analytic and systematic reviews of the literature, reviews that draw on the full range of methodologies used to elucidate what works in HIV prevention in specific populations and contexts, how and why. Rigorous methods are available to improve the interpretation and synthesis of complex intervention evaluation results and should be applied here (Oliver et al., 2004; Shepperd et al., 2009).
- ▶ Creating an accessible, consolidated and user friendly evidence base on HIV prevention including translation of evaluation findings into practical programme application guidance and providing examples of sound programme impact pathways for various types of programmes so they can be used easily by programme managers.

Q4: What specific interventions and resources are needed?

➔ The objective

This step involves determining what specific interventions are needed for particular populations and settings and what resources need to be available to implement them. This is an essential component of HIV prevention programme planning and needs to draw on "Know Your Epidemic and Response" information to determine the right mix of interventions and the scale and minimum level of quality with which they need to be implemented and where. Ideally, a programme manager will already be able to draw on the results from previous evaluation studies to select an appropriate intervention or package of interventions which have demonstrated the desired results in populations and settings similar



to his/her own. Systematic reviews which summarize the evaluation findings for specific interventions across different studies are an important resource to help shape the programme design. In addition, methods to estimate resource needs and to collect cost data need to be employed (such as National AIDS Spending Surveys (NASA), National Health Accounts (NHA), and/or resource tracking systems). Cost-effectiveness and cost-benefit analyses can also be conducted here.

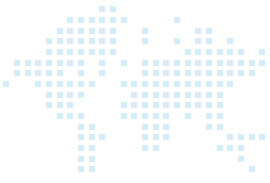
➔ The experience so far

- ▶ Data that allow for a thorough situation and gap analysis on HIV prevention are only recently being collected in a systematic way – especially information gathering about who is doing what, where, and at what cost, is in its infancy.
- ▶ Investing in HIV prevention programmes occurs as part of a decision-making process in which resources are allocated among competing demands.
- ▶ HIV prevention approaches with proven effectiveness are not fully employed where appropriate. Most countries, even those with generalized epidemics, are far from having achieved adequate coverage with basic and widely accepted interventions (Bertozzi et al, 2008). This may be due to political reasons, resource constraints, and/or practical reasons as translation of the available evidence into practical programme guidance is lagging behind.

➔ The way forward

Programme managers need to:

- ▶ Increase capacity building in the use and application of available standardized methods and tools to support effective planning processes (see for example: *Practical Guidelines for Intensifying HIV Prevention. Towards Universal Access. UNAIDS, 2008a*).
- ▶ Plan structural interventions in an integrated way with needed biomedical and behavioural interventions at different levels to address the identified social drivers. Ensure the intervention approach adheres to human rights considerations including participation of those infected/affected by HIV in needs analysis and programme design, and avoiding stigma or infringement of human rights of individuals or communities.
- ▶ Describe the *initial* programme impact pathway (PIP) including the selection of appropriate measures to monitor and/or evaluate the effects of the programme. Carefully assess the potential impact of each HIV prevention measurement strategy and consult with potential beneficiaries about their likely reception.
- ▶ When considering costs, consider that a costly measure that provides acceptable effectiveness and substantial benefits should take precedence over measures that are less expensive but have less impact on the epidemic. An analysis should be done of the relative value of interventions (such as comparative costs and effectiveness) in order to maximize results in resource-constrained settings. Provide realistic estimates of what it will cost to achieve the prevention targets. Costs for HIV prevention also must be analysed in comparison to costs of not preventing the spread of the HIV epidemic; of treatment for people who become infected; in loss of social capital with rising HIV infections; to the families and communities in terms of personal/individual, familial, social and economic loss; and of dealing with a larger epidemic later due to current inaction and inertia. Also estimate the impact that the HIV prevention measure is likely to have on broader health and social benefits.
- ▶ Ensure that national planning processes for scaling up HIV prevention programmes address the issues of needed human resources and commodity security. Work with existing human and organizational resources while planning



and investing to increase capacity to be able to expand services in the future.

International organizations can help with:

- ▶ Creating a typical intervention costs matrix.

Q5: What are we doing? Are we doing it right?

➔ The objective

Any attempt to evaluate the outcome or impact of a programme must first establish whether the programme was actually implemented. All programmes should therefore conduct input and output monitoring not only for programme management and accountability purposes but also for use in evaluations. Data on inputs (resources used in the programme) and outputs (results of the programme activities) usually exist in programme documentation (e.g. accounts, activity reports, logs) and client records which compile information about the time, place, type and amount of services delivered, and about the clients receiving the services.

➔ The experience so far

In many instances, routine monitoring data is insufficiently collected, of poor quality, and not well utilized. The barriers to routine monitoring include:

- ▶ Insufficient understanding of how to design a routine monitoring system that reflects programme impact pathways, and not just statutory recording of outputs (such as the number of leaflets handed out) (Lippeveld et al., 2000).
- ▶ Challenges at primary data collection level include: a lack of perceived utility of data; insufficient time and capacity of implementing staff; complicated and/or multiple data collection forms and registers with many elements of data

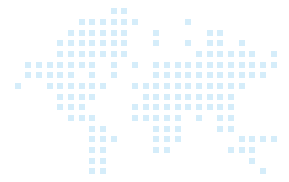
to be collected and cross-posted; and high-risk, marginalized or hidden communities may object to data collection and/or reporting (Napp et al., 2002; Kegeles et al., 2005; Otworld K et al., 2007; Garrib et al., 2008).

- ▶ At the collation and aggregation levels, major challenges seem to be: non-standardized definitions for data collected; duplication of data from multiple registers; lack of tools or inappropriate use of tools to extract and aggregate data into consistent indicators; and non-submission of consolidated indicators up the reporting chain even when available (Weeks et al., 2000; Otworld K et al., 2007; Garrib et al., 2008; Wilkins et al., 2008).
- ▶ Challenges with integration and coordination across multiple (similar) programmes: inability of national governments to enforce minimum data reporting from private (for profit) facilities; and lack of integration of monitoring data reported to donors from programmes funded directly by them (Pappaioanou et al., 2003; Otworld K et al., 2007; Garrib et al., 2008). Further, inability to deal with double counting can reduce the validity of reported data.
- ▶ Lack of a “data use” culture: frontline workers may collect and report data because of funding or administrative requirements but do not utilize the data for local decisions. Higher level officers may regard the receipt of monitoring reports from levels below as a “check in the box” with no material consequence for the programme (Pappaioanou et al., 2003; Bill and Melinda Gates Foundation, 2008; Garrib et al., 2008).

➔ The recommended way forward

Programme managers can improve routine monitoring by taking the following actions:

- ▶ Develop standards for output monitoring that address issues such as standardized reporting forms, how to count individuals (not only



contacts), the level of data disaggregation, methods for determining denominators, geo-referencing, feedback and use of the data at the point where it is collected.

- ▶ Collect expenditure data as part of routine monitoring data.
- ▶ Develop procedures which use field engagement to understand data quality issues and include external data quality assessments; ensure timely feedback mechanisms to improve routine monitoring where needed.
- ▶ Fund and support a responsible unit/person at the national level responsible for collating and analysing routine monitoring data for *all* HIV prevention interventions (not just HIV testing and counselling and prevention of mother-to-child transmission of HIV, or government-supported programmes only).
- ▶ Engage with donors to ensure that pertinent monitoring data that are collected from donor-funded programmes and reported to the donor are also reported to the government (e.g. Global Fund and PEPFAR monitoring data).
- ▶ Supplement routine programme monitoring data with data collected through surveys which include questions on programme exposure and with in-depth assessments of the quality of services provided.

International organizations can help with:

- ▶ Standardizing routine monitoring indicators for measuring coverage (i.e. inclusion of key interventions in the HIV prevention package and intensity of intervention exposure) and other outputs, and harmonizing across different donors. The global Monitoring and Evaluation Reference Group (MERG) could help lead this effort.
- ▶ Establishing mechanisms for sharing donor-requested data with respective governments.
- ▶ Including routine monitoring indicators in international reporting.

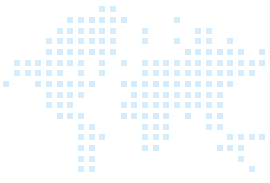
Q6: Are we implementing the programme as planned?

➔ The objective

Outcome and impact evaluations require information on whether the programme was implemented as intended. Most programmes should therefore periodically conduct a process evaluation to provide detailed information additional to routine monitoring data on programme implementation. This information may include, but is not limited to: access to services, whether services reach the intended population, how services are delivered, client satisfaction and perceptions about needs and services, management practices. Process evaluation assesses whether the programme was implemented according to quality standards and what the intensity of the programme exposure was for participants. In addition, a process evaluation might provide an understanding of cultural, socio-political, legal and economic contexts that affect implementation of the programme. Process evaluation is used to identify what is working well in programme implementation and where there are problems. Data gathered through a process evaluation help to document critical information for programme scale up or replication elsewhere, should the programme outcomes warrant these.

➔ The experience so far

- ▶ Many programmes are evaluated in terms of their results without adequate attention to implementation issues. This may lead to inconclusive results or an underestimate of programme effect, even null effects: a programme may be rejected not because of a failure of the approach, but because the programme was poorly implemented to the extent that it no longer has a detectable effect.
- ▶ The fact that a programme was implemented



in a particular way in the past is not sufficient grounds for assuming that the programme will be implemented in the same way when replicated elsewhere or at a later date. Fidelity (i.e. the implementation of a programme as originally designed) is expected to be reduced when implementing in a real world setting because of resource restrictions and/or other external factors (Rand publication, 2004). Moreover, some programmes are intended to be adapted to local circumstances. Therefore, it is critically important to assess what happens during programme implementation.

➔ The recommended way forward

Every programme manager should ensure that:

- ▶ The programme (and key projects within it) is periodically assessed using a process evaluation that examines fidelity to the programme design, the quality of services provided, client recruitment and retention, reach, intensity of programme delivered and received, client reaction/satisfaction, contextual changes, etc. This evaluation should include the following elements:
 - A programme impact pathway (PIP) as part of the evaluation plan to guide the process evaluation. For combination HIV prevention, some PIPs are conditional on having accomplished other programme elements, and some may work together synergistically. Their logical sequence or degree of integration will determine the process evaluation priorities. The more complicated the programme, the more essential a PIP is to help manage and evaluate the programme.
 - A data management system that reflects the PIP and informs the routine monitoring system.
 - An external (i.e. objective) evaluation component.

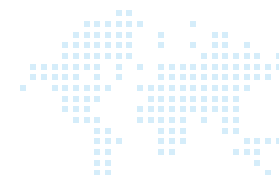
- ▶ Data and on-the-ground assessments are frequently looked at in order to identify problems and take corrective action in a timely manner. Also, provide timely feedback to stakeholders and to frontline staff providing the data and build capacity for using data to improve the programme.

Q7: Are interventions working/making a difference?

➔ The objective

Many of the basic questions about efficacy of HIV prevention programmes should already have been resolved. However, the effectiveness evidence base may be incomplete. In addition to scaling up HIV prevention interventions with known effectiveness, programme managers also have to take the risk of rolling out large-scale HIV prevention programmes of uncertain effectiveness. In any case, the programme should be designed according to the best available evidence and with an explicit programme impact pathway. In addition, due consideration must be given to any special circumstances (social, economic, political) that may limit the replicability of the intervention in other geographic areas or populations.

Collecting baseline data prior (or as early as possible) to the implementation of the programme is an important step. Where appropriate, existing data sources can be used here. If the data from routine monitoring and process evaluation show that no major implementation problems were encountered or corrective action was taken in case of problems, programme managers should assess if desired changes in intermediate outcomes (behavioural, social, structural) and changes in HIV incidence (if appropriate to the programme) are observed.

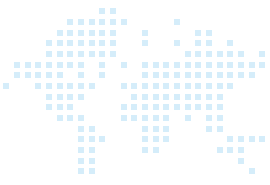


Bradford Hill considerations for determining plausible association*

[Source: Hill A (1965). The environment and disease: association or causation?
Proceedings Royal Society Medicine 58:295-300]

1. *Strength of association.* The stronger the association, the less it could merely reflect the influence of some other factor(s). This includes consideration of statistical precision and methodological rigour of the existing studies with respect to bias (selection, information, confounding).
2. *Consistency.* Replication of findings by different investigators, at different times, in different places, with different methods, and the ability to convincingly explain different results.
3. *Specificity of the association.* There is an inherent relationship between specificity and strength in the sense that the more accurately defined the disease and exposure, the stronger the observed relationship should be.
4. *Temporality.* The ability to establish that a cause in fact preceded in time the presumed effect.
5. *Biological gradient.* Incremental change in disease rates in conjunction with corresponding changes in exposure. The verification of a dose-response relationship consistent with the hypothesized conceptual model.
6. *Plausibility.* We are more willing to accept the case for a relationship that is consistent with our general knowledge and beliefs.
7. *Coherence.* How well do all the observations fit with the hypothesized model to form a coherent picture?
8. *Experimental evidence.* The demonstration that under controlled conditions changing the exposure causes a change in the outcome is of great value.
9. *Analogy.* We are more willing to accept arguments that resemble others we accept.

*These are a guide only and should not be used as criteria for scoring or weighting.



It is important to gather data that support a plausible link between the programme's operations and the observed outcomes. To increase the reliability and validity of programme evaluation at this stage, data triangulation methods should be applied (i.e. analysing data from multiple data sources; their combined and complementary use helps to overcome inherent weaknesses in the data sets) and can help create a "convergence of evidence" to draw plausible conclusions.

➔ The experience so far

- ▶ As mentioned above, HIV prevention approaches with proven effectiveness are not fully employed where appropriate; there are still some gaps in the evidence base for HIV prevention; and systematic reviews of existing evidence are often too narrowly focused and not translated into practical programming advice.
- ▶ Scaling up is possible only if a case can be made that programmes that have been successful on a small scale would work in other contexts (Duflo, 2004). Credible evaluations of scaling up programmes are increasing, but are still lagging behind.

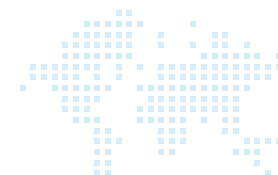
➔ The recommended way forward

National/subnational programme managers who need to make a decision to scale up a particular programme:

- ▶ Should commission a retrospective study assessing whether the programme achieved its intended results in a similar context elsewhere (if available, systematic reviews may be useful here). A programme is considered promising for scale-up if beneficial results were observed and their nature and extent are considered important. However, external factors that may affect the decision for scale-up or how to scale up should be carefully considered before

going ahead. These may include cost, funding limitations, cultural acceptability issues, sustainability constraints, etc. A rapid assessment of cost-benefit to scale up (e.g. modelling, rule of thumb, drawing on data from cost analysis in the prior stages of the programme's evaluation) and formative research on acceptability and other issues are typically needed prior to scaling up.

- ▶ If efficacy and/or effectiveness of the programme was well established and there are no major external factors affecting scale-up, it may then suffice to monitor if the programme is being scaled up according to plan and still achieving its intended effects.
- ▶ If there are important uncertainties about the effectiveness of the programme, then a strong prospective evaluation should be commissioned including the following key steps:
 - Construct a programme impact pathway (PIP) to lay out how the programme is expected to reduce HIV incidence (if appropriate) and other outcome measures.
 - Establish criteria for determining effectiveness with high plausibility that the observed trends are due to the programme and not other factors. The Bradford Hill list of considerations about causality, for example, may be used as a guide to establishing a coherent picture about the effects of the programme that are convincingly plausible [Box 5].
 - Document the implementation of the HIV prevention programme and its performance using data from routine programme monitoring, process evaluations and performance assessments.
 - Prospectively collect data that describe the programme in detail including intervention package, intervention coverage (who, where) and quality, duration, etc.
 - Establish (or if already in existence, use) a minimum package of data sources [Box 6]



BOX 6.

Key evaluation questions and data sources to evaluate Avahan, the India AIDS initiative

[Source: Chandrasekaran et al. 2008]

Evaluation questions	Methods	Primary data sources [scope]
----------------------	---------	------------------------------

Scale, coverage, quality, costs of Avahan implementation

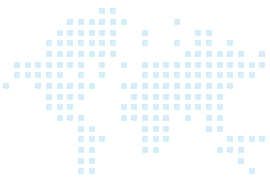
Are the geographic footprint, quality of coverage and service uptake (~80% of population) adequate over time?	<ul style="list-style-type: none"> ▶ Measure of coverage, utilization, intensity, quality of services delivered. ▶ Estimation of core group size using mapping and size estimation activities carried on by all state level partners. ▶ Coverage triangulation by exposure to intervention questions. 	<ul style="list-style-type: none"> ▶ Management Information System (MIS) [all Avahan implementation sites]. ▶ Quality assessments [sample of Avahan implementation sites]. ▶ Programme-generated mapping and size data [all Avahan districts]. ▶ Integrated Bio-Behavioral Assessment (IBBA) [29 districts].
What were the costs associated with the implementation?	<ul style="list-style-type: none"> ▶ Systematic collection of programme costs by category. 	<ul style="list-style-type: none"> ▶ Routine financial reports [all Avahan implementation sites]. ▶ Detailed costing studies in selected sites [23 districts]. ▶ MIS [all Avahan implementation sites].

Epidemic impact of Avahan

Has there been an increase in condom use in high-risk groups?	<ul style="list-style-type: none"> ▶ Reported condom use by partner type. 	<ul style="list-style-type: none"> ▶ IBBA [29 districts]. ▶ Special Behavioural Survey (SBS) [6 districts for sex workers; 4 districts for men who have sex with men; transgender]. ▶ Other data sources [sources vary by State].
Has there been a reduction in STI and new HIV infections in high-risk groups?	<ul style="list-style-type: none"> ▶ STI and HIV infection prevalence changes. ▶ Model estimates of HIV infections through mathematical modelling with available data to assess changes in HIV incidence in the presence and absence of HIV intervention. 	<ul style="list-style-type: none"> ▶ IBBA (2 rounds) [29 districts]. ▶ SBS [6 districts for sex workers; 4 districts for men who have sex with men; transgender]. ▶ MIS [all Avahan implementation sites].
Has there been a reduction in HIV infection in the general population?	<ul style="list-style-type: none"> ▶ Indirect measure of incidence through monitoring ANC prevalence in 15-24 years old. ▶ Model estimates of HIV prevalence/incidence. 	<ul style="list-style-type: none"> ▶ ANC surveillance [135 districts (2 sites/district)]. ▶ General Population Survey (GPS) [5 districts]. ▶ IBBA [29 districts]. ▶ SBS [6 districts for sex workers; 4 districts for men who have sex with men; transgender]. ▶ Other data sources [sources vary by State].
Can these changes be attributed to interventions in high-risk groups? What was Avahan's contribution to these changes?	<ul style="list-style-type: none"> ▶ ANC synthetic analysis. ▶ Modelling. 	<ul style="list-style-type: none"> ▶ ANC surveillance [135 districts]. ▶ District profiles of coverage by all interventions, MIS, IBBA, GPS [115 districts].

Cost effectiveness of Avahan

What was the cost effectiveness of the population (high-risk groups) reached?	<ul style="list-style-type: none"> ▶ MIS and cost data. 	<ul style="list-style-type: none"> ▶ MIS [all Avahan implementation sites].
What was the cost effectiveness of infections averted (high-risk groups, general population)?	<ul style="list-style-type: none"> ▶ Outputs of modelling and cost data. 	<ul style="list-style-type: none"> ▶ Population size estimates.
What was the cost efficiency of the various service components?		<ul style="list-style-type: none"> ▶ Routine financial reports [23 districts]. ▶ Detailed costing studies in select sites. ▶ Mathematical modelling.



including data from: antenatal care surveillance (annually); representative population-based surveys such as the Demographic Health Surveys (DHS) or AIDS Indicator Surveys (AIS) (every 3–5 years); integrated bio-behavioural surveys (IBBS) among most-at-risk populations (every two years).

- Conduct special studies (typically at a sub-national or lower level) to complement the minimum data sources in order to fill any data gaps. For example: in-depth social and behavioural studies of determinants of HIV vulnerability and risk; key contextual information; evaluations of interventions newly introduced in a specific setting or population, or testing specific assumptions in the programme impact pathway.
 - Involve programme beneficiaries and frontline service providers to validate the evaluation findings from their perspective. This is an important step in establishing credible evidence for the programme's effect.
 - Analyse all available data to determine if the observed changes can reasonably be attributed to the programme (use the criteria established earlier on). This process is referred to as data triangulation and should be conducted in a participatory manner (including decision makers, evaluators, programme managers, service providers and programme beneficiaries) to contribute different perspectives and minimize potential bias in the interpretation of the data. Mathematical modelling can be used to simulate control groups or control areas in the data analysis [Figure 8]. When data from different data sources converge (i.e. convergence of evidence), it provides sufficient evidence for a causal link. Ideally, the data sources and the analysis should allow for assessing some of the strengths and weaknesses of key components of the HIV prevention programme, so that adjustments in the programme can be made where needed.
- Weigh the robustness of the findings from the data triangulation analysis against the risk that the findings may be wrong and thus a decision based on them will have negative consequences (e.g. political, financial, health). A 2 x 2 table may be useful in assessing benefit versus risk (high/low robustness of evidence versus high/low risk).
 - An important additional question, especially for managers of large-scale programmes, is whether the efficiency of the HIV prevention programme can be improved. Conduct cost efficiency analyses of the different programme components.

Q8: Are collective efforts implemented on a large enough scale to impact the HIV epidemic?

➔ The objective

At this last step, the objective is to determine the effectiveness of the *national* HIV prevention programme. *National* programme managers will need to answer the following key questions:

- ▶ What are the level and trends of HIV prevalence and HIV-related behaviours (overall and by risk groups, time, person, place)?
- ▶ Do these trends relate to a change in risk or is it a reflection of the national history of the HIV epidemic?
- ▶ What factors (programmatic and contextual) may be associated with these trends?

The data collection methods used in step 1 (surveillance, surveys, modelling) and step 2 (determinant research) of Figure 6 are applied here again. The intent is to have a basic minimum package of comparable and consistent national data sets over time which allow for HIV trend analysis and how these relate to the characteristics of the national

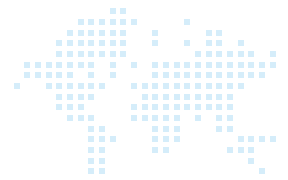
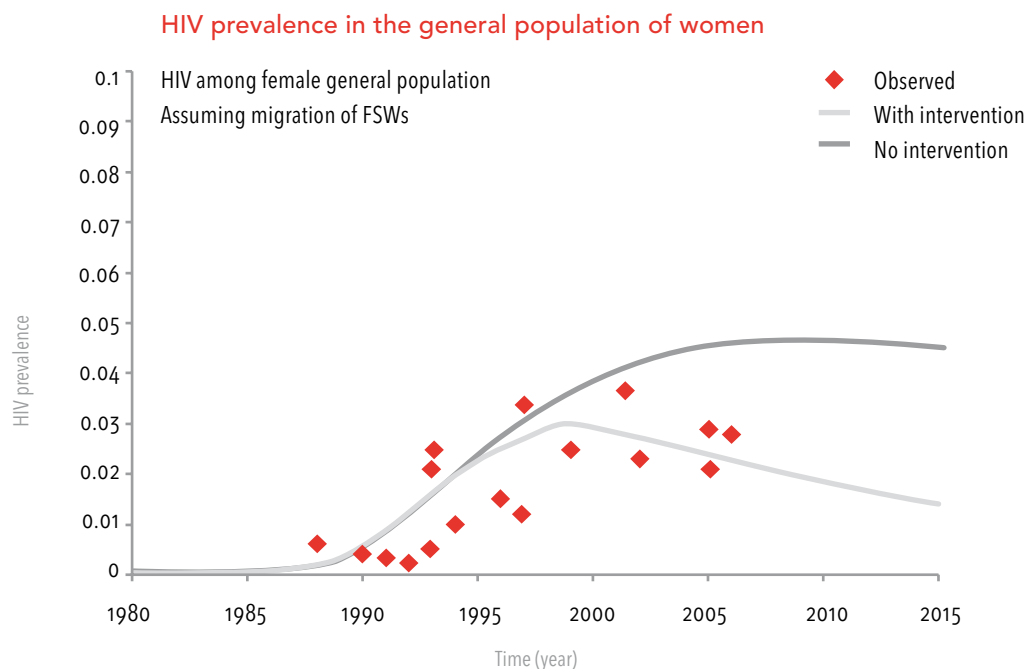
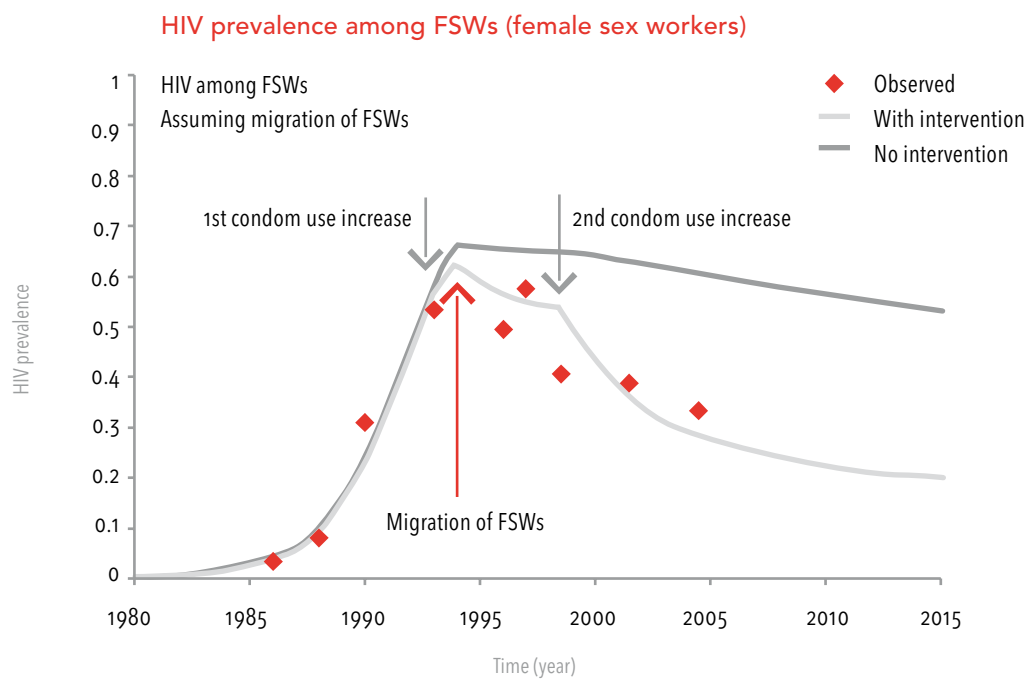
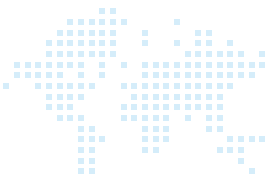


FIGURE 8.

Modelling of HIV prevalence to assess the impact of sex worker programmes: Example from Benin

Modelling of HIV prevalence in sex workers and the general population (interventions of Projets SIDA-1, SIDA-2, and SIDA-3)





HIV response, specifically the HIV prevention portfolio. Modelling can play an important role but needs to be applied in an appropriate manner and heavily relies on the availability of good quality data to obtain meaningful results.

➔ The experience so far

- ▶ The application of data triangulation methods to understand the course of the HIV epidemic and the national HIV response, including HIV prevention, is increasing. However, many countries still have important data gaps (data not collected and/or not made available to the national level), struggle with data quality issues, and neglect to analyse and use existing data (Peersman et al, 2009).
- ▶ Interpreting national HIV prevalence trends is a major challenge and has led to debates with regard to the explanations of decline, increase or stabilization of HIV epidemics. The issue is distinguishing between the expected saturation and decline in HIV prevalence without behaviour change and declines associated with reduced risk while accounting for improved survival on antiretroviral therapy. Recent models have allowed better insight into the possible causes and plausible attribution of observed declines in HIV prevalence (see for example, Hallett et al., 2009) [Figure 9].

➔ The recommended way forward

National programme managers need to:

- ▶ Focus on national (and subnational where available) HIV trends, behaviours, determinants and the mix of HIV prevention programmes. Note that attribution to specific contributions (financial or other) by any specific donor or programme partner is unrealistic and therefore not sought.
- ▶ Focus on data that are of critical impor-

tance for strategic planning and programme improvement.

- ▶ Focus on collecting a consistent, comparable data package and on analysis of determinants to understand both the programme and the context.

National programme managers need to ensure the following actions are undertaken:

- ▶ Collect a basic minimum package of comparable and consistent national data sets:
 - Sentinel surveillance among pregnant women attending antenatal clinics (ANC). Consider shifting to the collection of HIV prevalence data from prevention of mother-to-child transmission (PMTCT) programmes in countries and facilities with high PMTCT coverage, as per WHO criteria (WHO, forthcoming).
 - Sentinel surveillance among most-at-risk populations (both behavioural and HIV prevalence measures).
 - Nationally representative surveys collecting data on HIV prevalence, behaviours, antiretroviral therapy and exposure to HIV prevention interventions every 3–5 years (i.e. in high prevalence countries: conduct smaller surveys every 2–3 years and a larger survey every five years; in medium prevalence countries: conduct a large survey every five years). Conduct these surveys with large enough sample sizes to allow for comparisons between provinces (as appropriate) and to support precise HIV incidence measures. As appropriate, consider the application of laboratory assays for recent HIV infection, and triangulate the resulting incidence estimates with incidence estimates obtained through other methods – do not use results of incidence assays without validation/triangulation.
 - Trends in sexually transmitted diseases.
 - Routine programmatic data to fully document package of HIV prevention interventions: coverage of services, detailed description

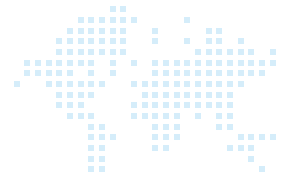
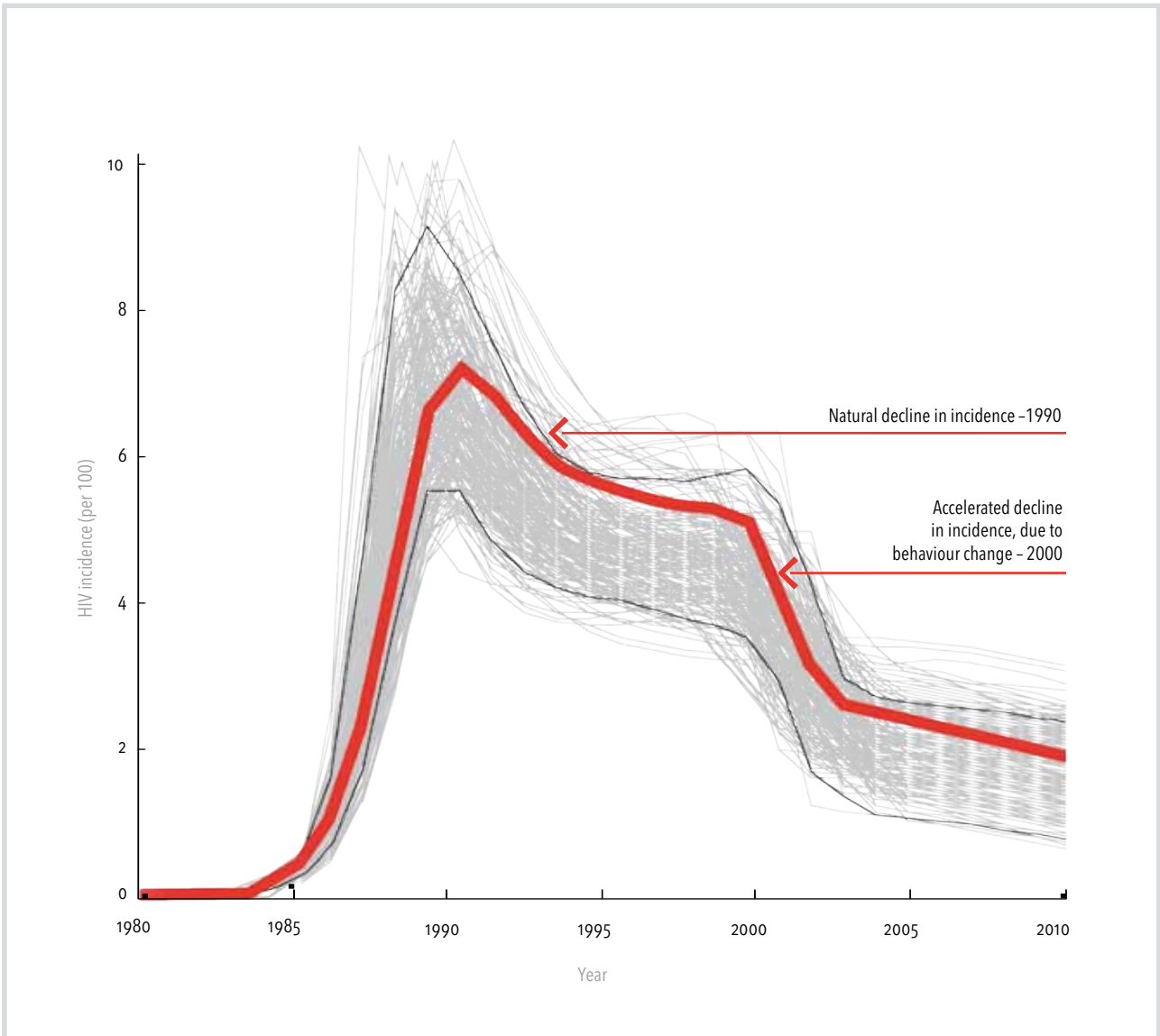
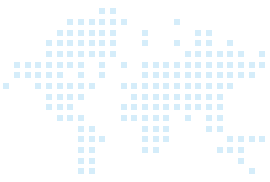


FIGURE 9.

Understanding national trends in the HIV epidemic:
Example from Zimbabwe





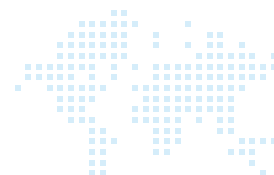
of programmes and their implementation, timing of roll out of different elements of the HIV prevention package (i.e. what, where, when, who, how).

- Special studies: in-depth social and behavioural studies using both quantitative and qualitative methods; key context data (see step 2 above for more details).
- ▶ Conduct analyses to determine if there is a significant change in HIV incidence trends:
 - Statistical analysis of HIV incidence trends to test for significant differences between trends (overall and by age, sex, region and urban/rural); and
 - Epidemiological significant trends to distinguish between trends related to a change in risk and trends reflecting the national history of HIV.
- ▶ Conduct analyses to explore “plausible attribution” to the national HIV prevention programme:
 - Descriptive triangulation of HIV incidence trends with behavioural trends.
 - Statistical analysis of trends by age, sex, region, urban/rural, risk groups.
 - Determinants analysis including:
 - Causal pathways/hypothesis-driven analysis: explicit hypothesis and competing explanations such as migration and mortality, to exclude external factors and establish programme associations.
 - Analysis of proximate determinants by age, sex, time, region (risk and behaviours; STI, circumcision, ART etc).
 - Descriptive statistics: associations with region, time, age; statistical analysis and sequence; exclusion of contextual factors.
 - Programme and contextual causal analysis: assessing competing hypotheses and excluding external factors. It is very important to involve programme managers and communities in these analyses.
- ▶ For analysing trends in HIV incidence, combine the following approaches (as none in itself is perfect):
 - Calculate HIV prevalence trends in young pregnant women (15–24 years old) at ANC clinics and triangulate these with HIV prevalence trends among young people (15–24 years old) in repeated population-based surveys.
 - Calculate HIV incidence from repeat cross-sectional studies (Hallett et al., 2008).
- ▶ Use models (e.g. Estimation and Projection Package, Asia Epidemic Model) to calculate trends in HIV incidence from cross-sectional HIV prevalence data. [Note that modelling is most valuable here for analysis of HIV trends and their possible association with changes in risk (i.e. triangulation of HIV incidence estimates and other trends and analysis of causal pathway hypotheses). See for example: Alary, 2009; Hallett et al., 2009]
- ▶ Communicate the results in understandable language to obtain continued support and to ensure programme managers understand how to further strengthen the HIV prevention programme.

International organizations can help with:

- ▶ Collating country examples of successful HIV prevention programmes with HIV impact.
- ▶ Cross-country analysis and comparisons of the effectiveness of HIV prevention programmes.
- ▶ Further developing and improving modelling techniques.
- ▶ Further developing and improving HIV incidence measures.

Conclusion

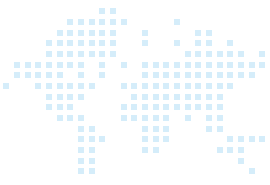


© A. Steblovsky/Dreamstime

There is growing recognition that greater investment in programme evaluation is needed to expand and solidify the evidence base for HIV prevention. To date, budgets for research and analytic work have not necessarily been used to generate data for improving high priority programmes or for scaling up of effective programmes. The recent push from the Global Fund, U.S. President's Emergency Plan for AIDS Relief (PEPFAR), UNAIDS, World Bank and WHO to focus more on evaluation is now putting these issues firmly on the table, and increased funding from international donors has recently become available. It is important for the HIV prevention community to take full advantage of this additional support and to direct evaluation efforts to where they are most needed and ensure they are conducted in a way that will maximize their utility for programme improvement and for our collective learning about successful HIV prevention.

This strategic guidance provided consensus recommendations for improving the design, implementation and analysis of HIV prevention evaluations with special attention to the current challenges:

- ▶ HIV prevention programmes are increasingly complex, multi-component and context-specific and the appropriate use of different evaluation methods needs to be clarified;
- ▶ The scientific evidence base to support the causal relationship between input/output and outcome/impact is incomplete and fragmented. The underlying behavioural theories leading to multiple behaviour changes and ultimately impact (i.e. reduction in HIV incidence) are difficult to assess;
- ▶ Many projects/interventions/services aim to affect HIV risk factors and/or vulnerabilities rather than averting HIV infections directly. The usefulness of different outcome measures needs to be clarified as well as how they "add up" to averting HIV infections; and



- ▶ measuring HIV incidence is particularly challenging. Currently, there are no ideal proxy measures for HIV incidence in populations.

All programme managers need to be able to identify what monitoring and evaluation activities are needed to guide programme management and improvement based on what is already known about the programme and the decisions that need to be made about the programme. Managers of *specific projects, interventions and services* also need to understand their programme's contribution to the national and subnational HIV prevention portfolio and the evidence base on HIV prevention. Managers of national and subnational HIV prevention portfolios also need to be able to coordinate a national or subnational evaluation agenda focused on actionable results for improving priority HIV prevention programmes.

Evaluation of HIV prevention programmes will substantially improve if programme managers take responsibility for:

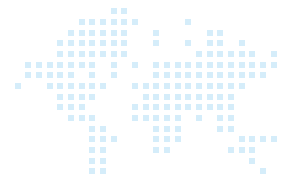
1. Describing the programme impact pathway and using it to guide programme implementation and evaluation.
2. Determining what decisions need to be made about the programme and selecting the most appropriate methods for collecting the necessary data.
3. Selecting appropriate measures to assess programme effects. Not all programmes should be responsible for collecting impact data, but all need to be justified as an essential component of the larger HIV prevention portfolio to avert HIV infections.
4. Ensuring that both programme implementation as well as programme effects are assessed, using a mixed-methods approach. Evaluation of the implementation of a programme and the context of implementation is essential for allowing a better understanding and interpretation of the programme results.

5. Focusing on actionable results. Monitoring and evaluation activities differ in purpose and design but complement one another. A utilization-focused approach focusing on eight basic questions addresses the key issues in programme evaluation: "Are we doing the right things?"; "Are we doing them right?"; and "Are we doing them on a large enough scale to make a difference?".

In addition, evaluation experiences and results need to be documented, compiled and shared more systematically and widely to inform programmes elsewhere.

International organizations can help with:

- ▶ Improving models for estimating HIV incidence and expand training on their application.
- ▶ Developing a reliable test for HIV incidence measurement.
- ▶ Providing a glossary for HIV prevention activities to facilitate standardization of programmatic data collection and interpretation.
- ▶ Providing a tool for describing the social, economic and political context of HIV programmes.
- ▶ Creating a typical intervention costs matrix.
- ▶ Adding to the existing meta-analytic and systematic reviews of the literature, reviews that draw on the full range of methodologies used to elucidate what works in HIV prevention in specific populations and contexts, how and why.
- ▶ Creating an accessible evidence base on HIV prevention effectiveness including translation of evaluation findings for practical programme application and providing examples of sound programme impact pathways for various types of programmes.
- ▶ Standardizing routine monitoring indicators for measure programme coverage.
- ▶ Establishing mechanisms for sharing donor-requested data with respective governments.



- ▶ Collating cross-country examples of national HIV prevention successes.
- ▶ Conducting cross-country analysis and comparisons of the effectiveness of HIV prevention programmes.

Evaluation funders can help with:

- ▶ Effectively targeting training, tools and other capacity building activities in support of building local evaluation capacity so that national

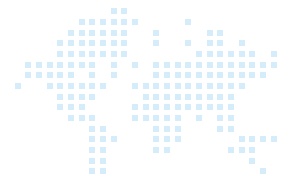
programmes can conduct and use their own evaluations.

- ▶ Agreeing on a prioritized global evaluation agenda in support of country-level needs in HIV prevention and a coordinated implementation approach.
- ▶ Ensuring that evaluation protocols and findings are more widely disseminated.
- ▶ Prioritizing support for research translation and the strategic use of evaluation findings in programme improvement.

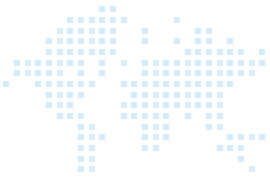


References

- Alary M. (2009). Impact of Sex Worker Programs in Benin. ICASA Report on West Africa.
- American Evaluation Association (1994). *Guiding Principles for Evaluators*. Fairhaven, MA.
- Apperdurai A. (1996). *Modernity at Large: Cultural Dimensions of Globalization*. Minneapolis: University of Minnesota Press.
- Auerbach J., J. Parkhurst, C. Caceras and K. Keller (2009). *Addressing social drivers of HIV/AIDS. Some conceptual, methodological and evidentiary considerations*. AIDS 2031 Social Driver Working Group.
- Barnett T. and A. Whiteside (2002). *AIDS in the Twenty-First Century: Disease and Globalisation*, Basingstoke: Palgrave MacMillan.
- Bauer P. and W. Brannath (2004). The advantages and disadvantages of adaptive designs in clinical trials. *Drug Discovery Today*, Vol. 9, No. 8, pp. 351–357.
- Bennett S. and R. Boerma (2007). Scaling up HIV/AIDS evaluation. *Lancet*, Vol. 367, pp. 79–82.
- Bertozi S., M. Laga, S. Bautista-Arredonda and A. Coutinho (2008). Making HIV prevention programmes work. *Lancet*, August 6 (HIV Prevention 5), pp. 64–77.
- Bill and Melinda Gates Foundation (2008). *Use it or lose it: How Avahan used data to shape its HIV prevention efforts in India*. New Delhi.
- Bradlow E. (1998). Encouragement Designs: An Approach to Self-Selected Samples in an Experimental Design. *Marketing Letters*, Vol. 9, No. 4, pp. 383–391.
- Brown C. and R. Lilford (2006). The stepped wedge trial design: a systematic review. *BMC Medical Research Methodology*, Vol. 6, pp. 54.
- Byrne D. (2002). *Interpreting Quantitative Data*. London: Sage Publications.
- Campbell C. (2003). *Letting them die: why HIV/AIDS prevention programmes fail*. Oxford: James Curry.
- Chandrasekaran P., G. Dallabetta, V. Loo, S. Mills, T. Saidel, R. Adhikary, M. Alary, C. Lowndes, M.C. Boily and J. Moore, for the Avahan partners (2008). Evaluation design for a large-scale HIV prevention programme: the case of Avahan, the India AIDS initiative. *AIDS*, Vol. 22 (suppl 5), pp. S1–S15.
- Chang M. (2009). *Adaptive Design Theory and Implementation Using SAS and R*. Boca Raton: Chapman & Hall/CRC.
- Coates T., L. Richter and C. Caceras (2008). Behavioural strategies to reduce HIV transmission: how to make them work better. *Lancet*, August 6 (HIV prevention 3), pp. 36–51.



- Coker R., R. Autun and M. McKee (2004). Untangling Gordian knots: Improving tuberculosis control through the development of “programme theories”. *International Journal of Health Planning and Management*, Vol. 19, pp. 217–226.
- Cousins J. and E. Whitmore (1998). Framing participatory evaluation. In Understanding and practicing participatory evaluation, E. Whitmore, ed. *New Directions for Evaluation*, No. 80, pp. 5–23.
- Cowan et al. (2008). The Regai Dzive Shiri Programme 2002–2007, presented at ICASA, Dakar.
- Craig P., P. Dieppe, S. Macintyre, S. Mitchie, I. Nazareth and M. Petticrew (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal*, Vol. 337. pp. 979–983.
- Davies R. and Dart J. (2005). The “Most Significant Change” (MSC) technique. A guide to its use. London: Care International.
- De Lay P. and V. Manda (2004). Politics of monitoring and evaluation: lessons from the AIDS epidemic. In Global advances in HIV/AIDS monitoring and evaluation, D. Rugg, G. Peersman and M. Carael, eds. *New Directions for Evaluation*, No. 103, pp. 13–31.
- Duflo E. (2004). *Scaling up and evaluation*. Annual World Bank Conference on Development Economics. Washington: World Bank.
- Garrib A., N. Stoops, A. Mckenzie, L. Dlamini, T. Govender, J. Rohde and K. Herbst (2008). An evaluation of the district health information system in rural South Africa. *South African Medical Journal*, Vol. 98, No. 7, pp. 549–52.
- Global Fund (2009). *Evaluation of complex health interventions and their scale up: a review. Background Paper*. London: LSHTM and Global Fund.
- Gupta Rao G., J. Parkhurst, J. Ogden, P. Aggleton and A. Mahal (2008). Structural approaches to HIV prevention. *Lancet*, August 6 (HIV Prevention 4), pp. 52–63.
- Habicht J.P., C. Victoria and J. Vaughan (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International Journal of Epidemiology*, Vol. 28, pp.10–18.
- Hallett T., Zaba B., Todd J., Lopman B., Mwita M., Biraro S., Gregson S. and Boerma J.T., on behalf of the ALPHA Network (2008). Estimating incidence from prevalence in generalised HIV epidemics: Methods and validation. *PLoS Medicine* (www.plosmedicine.org), April, Vol. 5, Issue 4, e80.
- Hallett T., S. Gregson, O. Mugurungi, E. Gonese and G. Garnett (forthcoming). Assessing evidence for behaviour change affecting the course of HIV epidemics: a new mathematical modelling approach and application to data from Zimbabwe.



Hill A.B. (1965). The environment and disease: association or causation? *Proceedings Royal Society Medicine*, Vol. 58, pp. 295–300.

Gregson et al. (2007). Impact and process evaluation of integrated community and clinic-based HIV-1 control: a cluster-randomised trial in eastern Zimbabwe. *PLoS Medicine*, Vol. 4, pp. e102.

Henry G., G. Julnes and M. Mark, eds. (1998). Realist evaluation: an emerging theory in support of practice. *New Directions for Evaluation*, No. 78.

Ingold T., ed. (1996). *Key debates in anthropology: the concept of society is theoretically obsolete*. New York: Routledge.

Julnes G. and D. Rog, editors' notes. Current Federal Policies and Controversies over Methodology in Evaluation. In *Informing Federal Policies on Evaluation Methodology: Building the Evidence Base for Method Choice in Government Sponsored Evaluation*, Julnes G. and D. Rog, eds. *New Directions in Evaluation*, Vol. 113, pp. 129–147.

Jewkes et al. (2008). Impact of Stepping Stones on incidence of HIV and HSV-2 and sexual behaviour in rural South Africa: a cluster randomized controlled trial. *British Medical Journal*, Vol. 337, No. a506.

Kamali et al. (2003). Syndromic management of sexually-transmitted infections and behaviour change interventions on transmission of HIV-1 in rural Uganda: a community randomised trial. *Lancet*, Vol. 361, pp. 645–52.

Kegeles S.M., G.M. Rebchook and S. Tebbetts (2005). Challenges and facilitators to building program evaluation capacity among community-based organizations. *AIDS Education & Prevention*, Vol. 17, pp. 284–99.

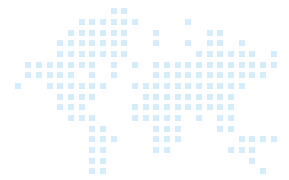
Leroy J., M. Ruel and E. Verhofstadt (2009). The impact of conditional cash transfer programmes on child nutrition: a review of evidence using a programme theory framework. *Journal of Development Effectiveness*, Vol. 1, No. 2, pp. 103–29.

Low-Beer D. and Stoneburner R. (2004). *Uganda and the challenge of HIV/AIDS*. In *The political economy of AIDS*, N. Pou and A. Whiteside, eds. Aldershot: Ashgate Publishing.

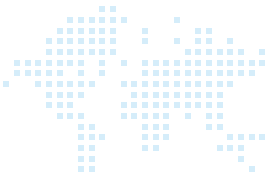
Lyerla R., E. Gouws and J. Garcia-Calleja (2008). The quality of sero-surveillance in low- and middle-income countries: status and trends through 2007. *Sex Transm Infect*, Vol. 84 (suppl. 1), pp. i85–i91.

McKee N., J. Bertrand and A. Becker-Benton (2004). *Strategic communication in the HIV/AIDS epidemic*. London: Sage Publications.

MMWR (1992). A framework for assessing the effectiveness of disease and injury prevention. *MMWR*, Vol. 41, No. RR–3.



- Napp D., D. Gibbs, D. Jolly, B. Westover B. and G. Uhl (2002). Evaluation Barriers and Facilitators among Community-Based HIV Prevention Programs. *AIDS Education & Prevention*, Vol. 14 (suppl. A), pp. 38–48.
- Oliver S., L. Clarke-Jones, R. Rees, R. Milne, P. Buchanan and J. Gabbay et al. (2004). Involving consumers in research and development agenda setting for the NHS: developing an evidence-based approach. *Health Technol Assess*, Vol. 8, No. 15.
- Otwombe K.N., J. Wanyungu, K. Nduku and M. Taegtmeier (2007). Improving national data collection systems from voluntary counselling and testing centres in Kenya. *Bull World Health Organization*, Vol. 85, pp. 315–318.
- Panos (2006). *Breaking Barriers: effective communication for HIV prevention treatment care and support by 2010*. London: Panos Institute.
- Pappaioanou M., M. Malison, K. Wilkins, B. Otto, R.A. Goodman, R.E. Churchill, M. White and S.B. Thacker (2003). Strengthening capacity in developing countries for evidence-based public health: the data for decision-making project. *Soc Sci Med*, Vol. 57, pp. 1925–1937.
- Parker R., D. Easton and C. Klein (2000). Structural barriers and facilitators in HIV prevention: a review of international research. *AIDS*, Vol. 14 (suppl 1), pp. S22–S32.
- Patton M.Q. (1997). *Utilization-focused evaluation. 3rd edition*. Thousand Oaks: Sage Publications.
- Peersman G., D. Rugg, T. Erkkola, E. Kiwango and J. Yang. (2009). Are the investments in national HIV monitoring and evaluation systems paying off? *J Acquir Immune Defic Syndr*.
- Peto R., R. Collins and R. Gray (1995). Large-scale randomized evidence: Large, simple trials and overviews of trials. *J Clinical Epidemiology*, Vol. 48, No. 1, pp.23–40.
- Preskill H. and A. Coghlan, eds. (2003). Using appreciative inquiry in evaluation. *New Directions for Evaluation*, No. 100.
- Pronyk et al. (2006). Effect of a structural intervention for the prevention of intimate partner violence and HIV in rural South Africa: a cluster RCT. *Lancet*, Vol. 368, pp1973–1983.
- Rand publication. *Estimating the magnitude of the scale-up degradation qualifier*. <http://www.rand.org/publications/MR/MR923/MR923.apph.pdf>.
- Ross et al. (2007). Biological and behavioural impact of an adolescent sexual health intervention in Tanzania: a community-randomized trial. *AIDS*, Vol. 21, pp. 1943–1955.
- Rugg D. and S. Mills (2000). Development of an Integrated Monitoring and Evaluation Plan. In *Evaluating Programs for HIV/AIDS Prevention and Care in Developing Countries*, T. Rehle, T. Saidel, S. Mills and R. Magnani, eds. Arlington: Family Health International.



Rugg D., G. Peersman and M. Carael, eds. (2004). Global advances in HIV/AIDS monitoring and evaluation. *New Directions for Evaluation*, No. 103.

Ryan K. and L. DeStefano, eds. (2000). Evaluation as a democratic process: promoting inclusion, dialogue, and deliberation. *New Directions for Evaluation*, No. 85.

Sheppard S., S. Lewin, S. Straus, M. Clarke, M. Eccles, R. Fitzpatrick, G. Wong and A. Sheikh (2009). Can we systematically review studies that evaluate complex interventions? *PLoS Medicine*, Vol. 6, No.8, pp.e1000086, 1–8.

Susser M. (1996). Some principles in study design for preventing HIV transmission: rigor or reality. *Am J Public Health*, Vol. 86, pp.1713–1716.

United Nations (2006). *International Guidelines on HIV/AIDS and Human Rights*. New York.

UNAIDS (2008). *Guidelines for Privacy, Confidentiality and Security of HIV Information*. Geneva.

UNAIDS (2007). *Practical guidelines for intensifying HIV prevention. Towards universal access*. Geneva.

UNAIDS (2008). *Report on the global AIDS epidemic*. Geneva.

UNAIDS Monitoring and Evaluation Reference Group (2008). *Monitoring and Evaluation Glossary*. Geneva.

UNAIDS (2009). *UNAIDS Taxonomy of HIV Prevention Activities*. Draft 22 May. Geneva.

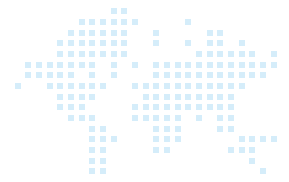
UNAIDS Prevention Reference Group (2009). Draft definition of “combination HIV prevention”, Montreux: PRG Meeting, 5 December.

Weeks R. M., F. Svetlana, S. Noorgoul and G. Valentina (2000). Improving the monitoring of immunization services in Kyrgyzstan. *Health Policy and Planning*, Vol. 15, pp. 279–286.

Weiss H., J. Wasserheit, R. Barnabas, R. Hayes and L. Abu-Raddad (2008). Persisting with prevention: The importance of adherence for HIV prevention. *Emerging Themes in Epidemiology*, Vol. 5, 8 doi: 10.1186/1742–7622–5–8.

Whitmore E., ed. (1998). Understanding and practicing participatory evaluation. *New Directions for Evaluation*, No. 80.

Wilkins K., P. Nsubuga, J. Mendlein, D. Mercer and M. Pappaioanou (2008). The data for decision making project: assessment of surveillance systems in developing countries to improve access to public health information. *Public Health*, Vol. 122, pp. 914–922.



Appendix 1. M&E Glossary

[Source: excerpted from UNAIDS. Glossary of Monitoring and Evaluation Terms. Geneva: UNAIDS, 2008c]

Activity. Actions taken or work performed through which inputs such as funds, technical assistance, and other types of resources are mobilized to produce specific outputs.

Baseline. The status of services and outcome-related measures such as knowledge, attitudes, norms, behaviours, and conditions before an intervention, against which progress can be assessed or comparisons made.

Coverage. The extent to which a programme/intervention is being implemented in the right places (geographic coverage) and is reaching its intended target population (individual coverage).

Effectiveness. The extent to which a programme/intervention has achieved its objectives under normal conditions in a real-life setting.

Efficacy. The extent to which an intervention produces the expected results under ideal conditions in a controlled environment.

Efficiency. A measure of how economically inputs (resources such as funds, expertise, time) are converted into results.

Evaluation. The rigorous, scientifically based collection and analysis of information about programme/intervention activities, characteristics, and outcomes that determine the merit or worth of the programme /intervention. Evaluation studies provide credible information for use in improving programmes/interventions, identifying lessons learned, and informing decisions about future resource allocation.

Impact. The long-term, cumulative effect of programmes/interventions over time on what they ultimately aim to change, such as a change

in HIV infection, AIDS-related morbidity and mortality. **Note:** Impacts at a population-level are rarely attributable to a single programme/intervention, but a specific programme/intervention may, together with other programmes/interventions, contribute to impacts on a population.

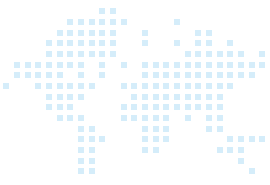
Impact evaluation. A type of evaluation that assesses the rise and fall of impacts, such as disease prevalence and incidence, as a function of HIV programmes/interventions. Impacts on a population seldom can be attributed to a single programme/intervention; therefore, an evaluation of impacts on a population generally entails a rigorous design that assesses the combined effects of a number of programmes/interventions for at-risk populations.

Impact monitoring. Tracking of health-related events, such as the prevalence or incidence of a particular disease; in the field of public health, impact monitoring is usually referred to as "surveillance".

Incidence. The number of new cases of a disease that occur in a specified population during a specified time period.

Inputs. The financial, human and material resources used in a programme/intervention.

Input and output monitoring. Tracking of information about programme/intervention inputs (i.e. resources used in the programme/intervention) and programme /intervention outputs (i.e. results of the programme /intervention activities). **Note:** Data on inputs and outputs usually exist in programme/intervention documentation (e.g. activity reports, logs) and client records which compile information about the time, place, type and amount of services delivered, and about the clients receiving the services.



Intervention. A specific activity or set of activities intended to bring about change in some aspect(s) of the status of the target population (e.g. HIV risk reduction, improving the quality of service delivery).

Monitoring. Routine tracking and reporting of priority information about a programme/project, its inputs and intended outputs, outcomes and impacts.

Outcome monitoring. Tracking of variables that have been adopted as valid and reliable measures (i.e. indicators) of the desired programme/intervention outcomes. Outcome monitoring does not infer causality; changes in outcomes may be attributable to multiple factors, not just a specified programme/intervention. **Note:** With national AIDS programmes, outcome monitoring is typically conducted through population-based surveys (i.e. representative of the target population, not necessarily the general population).

Outcome. Short-term and medium-term effect of an intervention's outputs, such as change in knowledge, attitudes, beliefs, behaviours.

Outcome evaluation. A type of evaluation that determines if, and by how much, intervention activities or services achieved their intended outcomes. An outcome evaluation attempts to attribute observed changes to the intervention tested. **Note:** An outcome evaluation is methodologically rigorous and generally requires a comparative element in its design, such as a control or comparison group, although it is possible to use statistical techniques in some instances when control/comparison groups are not available (e.g. for the evaluation of a national programme).

Outputs. The results of programme/intervention activities; the direct products or deliverables of programme/intervention activities such as the

number of HIV counselling sessions completed, the number of people served, the number of condoms distributed.

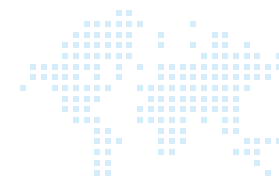
Prevalence. The total number of persons living with a specific disease or condition at a given time.

Programme. An overarching national or subnational response to a disease. A programme generally includes a set of interventions marshalled to attain specific global, regional, country or subnational objectives; involves multiple activities that may cut across sectors, themes and/or geographic areas.

Process evaluation. A type of evaluation that focuses on programme/intervention implementation, including, but not limited to, access to services, whether services reach the intended population, how services are delivered, client satisfaction and perceptions about needs and services and management practices. In addition, a process evaluation might provide an understanding of cultural, socio-political, legal, and economic contexts that affect implementation of the programme/intervention.

Programme evaluation. A study that intends to control a health problem or improve a public health programme or service. The intended benefits of the programme are primarily or exclusively for the study participants or the study participants' community (i.e. the population from which the study participants were sampled); data collected are needed to assess and/or improve the programme or service, and/or the health of the study participants or the study participants' community. Knowledge that is generated does not typically extend beyond the population or programme from which data are collected.

Programme logic model or programme impact pathway (PIP). Management tool used to improve the design of interventions. It involves identifying



strategic elements (inputs, outputs, activities, outcomes, impact) and their causal relationships, indicators and the assumptions of risks that may influence success and failure. It thus facilitates planning, execution, and monitoring and evaluation of an intervention.

Project. An intervention designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader programme.

Second-generation surveillance. HIV surveillance that not only tracks HIV prevalence but also uses additional sources of data to increase the understanding of trends of the epidemic over time. It includes biological surveillance of HIV and other sexually transmitted infections as well

as systematic surveillance of the behaviours that spread them.

Sentinel surveillance. Ongoing, systematic collection and analysis of data from certain sites (e.g. hospitals, health centres, antenatal clinics) selected for their geographic location, medical speciality and populations served, and considered to have the potential to provide an early indication of changes in the level of a disease.

Surveillance. The ongoing, systematic collection, analysis, interpretation and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health. Surveillance data can help predict future trends and target needed prevention and treatment programmes.



Appendix 2.

Programme Evaluation Standards

A pioneer project to develop professional standards for programme evaluation was initiated in the United States in 1975. Its goal was to improve the evaluation of educational and training programmes in a variety of settings. Since then, these “Programme Evaluation Standards” have been revised and adapted by a range of national evaluation associations to make them relevant to other areas of investigation and to specific local conditions.

The Standards are generally acknowledged to be good practice and are routinely used in planning an evaluation, negotiating a contract to do an evaluation and in reviewing progress during implementation of an evaluation.

The Programme Evaluation Standards are:

1. Utility (U) Standards are intended to ensure that an evaluation will serve the information needs of intended users:

U1. Stakeholder Identification. Persons involved in or affected by the evaluation should be identified, so that their needs can be addressed.

U2. Evaluator Credibility. The persons conducting the evaluation should be both trustworthy and competent to perform the evaluation, so that the evaluation findings achieve maximum credibility and acceptance.

U3. Information Scope and Selection. Information collected should be broadly selected to address pertinent questions about the programme and be responsive to the needs and interests of clients and other specified stakeholders.

U4. Values Identification. The perspectives, procedures, and rationale used to interpret the findings

should be carefully described, so that the bases for value judgments are clear.

U5. Report Clarity. Evaluation reports should clearly describe the programme being evaluated, including its context, and the purposes, procedures, and findings of the evaluation, so that essential information is provided and easily understood.

U6. Report Timeliness and Dissemination. Significant interim findings and evaluation reports should be disseminated to intended users, so that they can be used in a timely fashion.

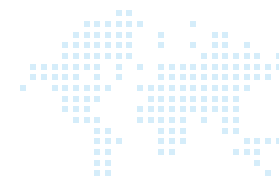
U7. Evaluation Impact. Evaluations should be planned, conducted, and reported in ways that encourage follow-through by stakeholders, so that the likelihood that the evaluation will be used is increased.

2. Feasibility (F) Standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal:

F1. Practical Procedures. The evaluation procedures should be practical, to keep disruption to a minimum while needed information is obtained.

F2. Political Viability. The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.

F3. Cost Effectiveness. The evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified.



3. Propriety (P) Standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results:

P1. Service Orientation. Evaluations should be designed to assist organizations to address and effectively serve the needs of the full range of targeted participants.

P2. Formal Agreements. Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or formally to renegotiate it.

P3. Rights of Human Subjects. Evaluations should be designed and conducted to respect and protect the rights and welfare of human subjects.

P4. Human Interactions. Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation, so that participants are not threatened or harmed.

P5. Complete and Fair Assessment. The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the programme being evaluated, so that strengths can be built upon and problem areas addressed.

P6. Disclosure of Findings. The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation, and any others with expressed legal rights to receive the results.

P7. Conflict of Interest. Conflict of interest should be dealt with openly and honestly, so that it does

not compromise the evaluation processes and results.

P8. Fiscal Responsibility. The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible, so that expenditures are accounted for and appropriate.

4. Accuracy (A) Standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the programme being evaluated:

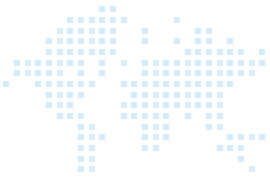
A1. Programme Documentation. The programme being evaluated should be described and documented clearly and accurately, so that the programme is clearly identified.

A2. Context Analysis. The context in which the programme exists should be examined in enough detail, so that its likely influences on the programme can be identified.

A3. Described Purposes and Procedures. The purposes and procedures of the evaluation should be monitored and described in enough detail, so that they can be identified and assessed.

A4. Defensible Information Sources. The sources of information used in a programme evaluation should be described in enough detail, so that the adequacy of the information can be assessed.

A5. Valid Information. The information gathering procedures should be chosen or developed and then implemented so that they will assure that the interpretation arrived at is valid for the intended use.



A6. Reliable Information. The information gathering procedures should be chosen or developed and then implemented so that they will assure that the information obtained is sufficiently reliable for the intended use.

A7. Systematic Information. The information collected, processed, and reported in an evaluation should be systematically reviewed and any errors found should be corrected.

A8. Analysis of Quantitative Information. Quantitative information in an evaluation should be appropriately and systematically analysed so that evaluation questions are effectively answered.

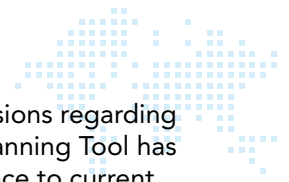
A9. Analysis of Qualitative Information. Qualitative information in an evaluation should be appropriately and systematically analysed so that evaluation questions are effectively answered.

A10. Justified Conclusions. The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them.

A11. Impartial Reporting. Reporting procedures should guard against distortion caused by personal feelings and biases of any party to the evaluation, so that evaluation reports fairly reflect the evaluation findings.

A12. Metaevaluation. The evaluation itself should be formatively and summatively evaluated against these and other pertinent standards, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses.

[Source: American Evaluation Association, AEA. The Program Evaluation Standards. Summary of the Standards. <http://www.eval.org/EvaluationDocuments/progeval.html>, accessed on 30 March 2009]



List of MERG Documents 2007–2009

1. **A Framework for Monitoring and Evaluating HIV Prevention Programmes for Most-At-Risk Populations (2007):** Provides an overview of M&E methods and approaches for most at-risk populations; it covers the use of strategic information for programme planning, M&E. Its focus is on the M&E of targeted HIV prevention programme.

2. **Additional Recommended Indicators. Addendum to UNGASS Monitoring the Declaration of Commitment on HIV/AIDS, Guidelines on Construction of Core Indicators (2008):** Presents the 40 core national indicators that provide minimum necessary information for national-level monitoring of the HIV epidemic and response, and provides detailed specifications and guidance on the 15 indicators recommended in addition to the 25 UNGASS indicators.

3. **Organizing Framework for a Functional National HIV M&E System (2008):** This framework describes 12 main M&E system components and defines a performance goal and results for each component. The framework helps countries to define an agreed set of national performance objectives and measures for the HIV M&E system and to guide strategies for building capacity, where needed, to reach these objectives.

4. **Glossary of M&E Terminology (2008):** Contains an alphabetical listing of M&E terms and their definitions, often with more in-depth explanations than would customarily be provided by dictionary definitions. The Glossary will facilitate and improve dialogue and understanding among all those who are involved in M&E of development activities. It should also serve as a valuable reference guide in M&E training. The selection of terms and their definitions have been carefully discussed and endorsed by the Global UNAIDS Monitoring and Evaluation Reference Group (MERG).

5. **Indicator Standards and Assessment Tool (2009):** Consists of a set of agreed indicator standards that are relevant at the national level for programme managers and service providers who need to select, revise and use indicators to monitor, manage and implement their country's response to the epidemic effectively. This will ensure that indicators provide decision-makers and key stakeholders with useful, feasible and relevant information. An additional aim is to reduce the burden of global reporting on countries by harmonising global level indicators across multilateral and bilateral organisations.

6. **Planning Tool for Developing a Digital Library of M&E Resources (2009):** This will help assure that users of a digital library can successfully locate

resources and make informed decisions regarding the quality of the materials. The Planning Tool has two purposes: 1) to provide guidance to current owners and future developers of a digital library on the range of issues to be addressed in usability and user-friendliness of the library and 2) to provide a list of questions to help organizations brainstorm if they can and should invest their resources in developing a digital library.

7. **Guidance on Capacity Building for HIV Monitoring and Evaluation (2009):** Provides practical advice for national AIDS programmes that are planning and implementing capacity building activities as part of their effort to develop a unified and effective national HIV M&E system. The Guidance is relevant to the wide range of individuals and organisations involved in the national HIV M&E system; it is particularly relevant for the health sector, given its central role in M&E of HIV.

8. **12 Components Monitoring and Evaluation System Assessment – Guidelines to support preparation, implementation and follow-up activities (2009):** These Guidelines provide information on the preparation for and implementation of an assessment of a national HIV M&E system. It also includes key steps to take after an assessment to facilitate implementation of M&E system strengthening activities. The Guidelines are built around the 12 main components of the HIV M&E system, which define the Organizing Framework for a Functional National HIV Monitoring and Evaluation System (UNAIDS, 2008). Consequently, the Guidelines also focus on using the 12 Components Monitoring and Evaluation System Strengthening Tool (UNAIDS, 2009a) to ensure a comprehensive and successful assessment.

9. **12 Components Monitoring and Evaluation System Strengthening Tool (2009):** Is a tool for assessing how well each of the 12 components of a national HIV M&E system is functioning. The tool facilitates the identification of strengths and weaknesses in the national HIV M&E system and the prioritization of system strengthening activities.

10. **Guidelines for Developing Terms of Reference for Prevention Evaluation (2009):** The Guidelines aim to foster a systematic approach to the evaluation of prevention programmes by focusing on an often overlooked yet critical step in evaluation planning: the preparation of terms of reference (TOR). It can be used to facilitate the planning of evaluations for HIV prevention, discussions on the design of these evaluations, and the drafting of TOR to guide such assessments. It is intended for use by anyone who prepares or reviews TORs for evaluations of HIV and AIDS prevention programmes and projects.



Uniting the world against AIDS

UNAIDS
20 AVENUE APPIA
CH-1211 GENEVA 27
SWITZERLAND

T (+41) 22 791 36 66
F (+41) 22 791 41 87

www.unaids.org