Data Triangulation

for HIV Prevention Program Evaluation in Low and Concentrated Epidemics

June 2010

Data Triangulation
for HIV Prevention Program Evaluation in
Low and Concentrated Epidemics
Data Triangulation
for HIV Prevention Program Evaluation in
Low and Concentrated Epidemics

USAID FROM THE AMERICAN PEOPLE | ASIA

In July 2011, FHI became FHI 360.

THE SCIENCE OF IMPROVING LIVES

FHI 360 is a nonprofit human development organization dedicated to improving lives in lasting ways by advancing integrated, locally driven solutions. Our staff includes experts in health, education, nutrition, environment, economic development, civil society, gender, youth, research and technology – creating a unique mix of capabilities to address today's interrelated development challenges. FHI 360 serves more than 60 countries, all 50 U.S. states and all U.S. territories.

Visit us at www.fhi360.org.

This document is made possible by the support of the American people through the United States Agency for International Development (USAID) through the USAID/Regional Development Mission Asia/Technical Assitance and Support Contract 3 (TASC 3), Task Order #2 (GHS-I-02-07-00007-00).

Copies of this publication can be obtained from:

FHI, Asia/Pacific Regional Office Tel. +662 263 2300, Fax. + 662 263 2114

ISBN 978-616-90553-0-3

Acknowledgements

Primary Authors: Tobi Saidel and Dimitri Prybylski

Those individuals who provided data and inputs for the country level case studies are gratefully acknowledged. These include Laxmi Bilas Acharya, Tasnim Azim, Ashish Kumar Datta, Philippe Girault, Kate Killberg, Virginia Loo, Robert Magnani, Nurholis Majid, Guy Morineau, Shanthi Noriega, Satish Raj Pandey, Dimitri Prybylski, Mahesh Shrestha, and Rizky Ika Syafitri.

Special thanks are given to Virginia Loo for extensive review and feedback on the document.

Table of Contents

List of Figures, Boxes, and Tables

Boxes

Figures

Tables

List of Acronyms

ANC	Antenatal Clinic
BCI	Behavior Change Intervention
BSS	Behavioral Surveillance Survey
DIC	Drop-In Center
FHI	Family Health International
FSW	Female Sex Worker
HCT	HIV Counseling and Testing
HSS	HIV Sentinel Surveillance
IBBS	Integrated Bio-Behavioral Survey
IEC	Information, Education, and Communication
IDU	Injecting Drug User
IPC	Interpersonal Communication
MARP	Most-At-Risk Population
M&E	Monitoring and Evaluation
MSM	Men who Have Sex with Men
MSW	Male Sex Worker
NEP	Needle Exchange Program
NGO	Non-Governmental Organization
OR	Odds Ratio
RSA	Rapid Situation Assessment
STI	Sexually Transmitted Infection
UNAIDS	Joint United Nations Programme on HIV/AIDS
VCT	Voluntary Counseling and Testing

Chapter I: Introduction

"Are we making progress?"

Those of us working on Surveillance, Monitoring and Evaluation in the field of HIV and AIDS face this question on a daily basis: "Are we making any progress in the fight against HIV?" We are asked this by global and international institutions spending billions of dollars to stop the spread of HIV globally. We are asked by governments who are trying to respond to the threat of HIV spread in their countries. We are asked by Ministers of Health who are responsible for allocation of limited resources to save lives and improve the health of the citizens they serve. We are asked by program managers who seek to implement the most effective programs. We are asked by communities affected by the spread of HIV in their midst. And we are asked by colleagues, peers and friends alike…are we making any progress?

When people ask this question, what they really want to know is whether we have found a "cure" for HIV and AIDS so that it goes away permanently. But once they understand that such cures are still far away, they want to know whether we are preventing HIV from spreading to those not yet infected. This is a difficult question, because short of a vaccine or a magic bullet that would stop new infections entirely, the answer requires us to understand whether fewer people are becoming infected as the epidemic progresses: This is not something that is easy to measure.

Rationale for this guide

The purpose of this guide is to explore how HIV surveillance can be used in the process of evaluating HIV prevention programs in low and concentrated epidemics, using the Asian context as an example, and to measure progress toward the goal of reducing HIV transmission. Asia is home to over half of the worlds population. While HIV prevalence in the region is lower than in some other regions, the HIV epidemics in Asia pose special challenges because they are diverse, rapidly evolving, and concentrated among most-at-risk populations (MARPs) who, because they often suffer high levels of stigma and discrimination, are difficult to identify and access. This situation not only makes it challenging to design and implement HIV prevention programs, but to evaluate the effectiveness of those programs as well.

A major limiting factor in evaluating HIV prevention programs at scale is that in most routine service delivery settings, evaluations using rigorous community-randomized experimental designs are not a feasible or ethical option, given the resources required and the delay or denial of services to communities assigned to a control arm. Instead, HIV prevention programs often rely on cross sectional surveys conducted for surveillance purposes and on routinely collected program monitoring data to assess effectiveness. In an era where there are far more resources going into the collection of data rather than into its use, there is a lot that can be done to improve the situation with the data we already collect [1]. It is not the intention of this guide to provide extensive coverage on the topic of monitoring and evaluation (M&E) for HIV programs since there is ample material on that subject already available [2-5]. Instead, the guide focuses on the synthesis and triangulation of data that are routinely collected as part of local and national level M&E and surveillance systems to provide evidence of program success.

Chapter 2: Approaching Program Evaluation

Key points in this chapter

- Outcome monitoring is an essential first step in evaluating program effectiveness
- In the absence of experimental or quasi-experimental evaluation studies (which are difficult to implement as part of routine program evaluation), we cannot establish definitively whether programs caused changes in outcomes
- By triangulating data, we can examine the evidence that programs contributed to changes in outcomes

Outcome and impact evaluation in the context of HIV prevention programs

For HIV prevention programs, "success" can be looked at in many different ways because program evaluation takes place at many different levels. We use monitoring and evaluation systems to guide the measurement of program performance according to the prescribed logic of the programs in question. These systems cover details as small as the number of new targeted condom outlets opening in a particular place every month (an output level indicator), and as large as the number of HIV infections prevented among clients of sex workers and their wives (impact level indicators). Obviously the complexity of tracking the first is significantly less than the second, which is why the first can be done on a fairly routine basis, as part of input and output project monitoring. But when it comes to outcome and impact evaluation, these are done more rarely, and they require more data and a different level of effort. This is the focus of the guide.

Evaluation terminology

Before delving further into the topic, it is instructive to review some of the standard terminology used in conjunction with outcome and impact evaluation (see Box 1). The differences in these definitions and how they apply to situations in which we use them can be subtle, but for the purposes of this guide, evaluation is considered mainly in the context of measuring the effectiveness of interventions implemented under routine field circumstances, as opposed to more formal efficacy research [6].

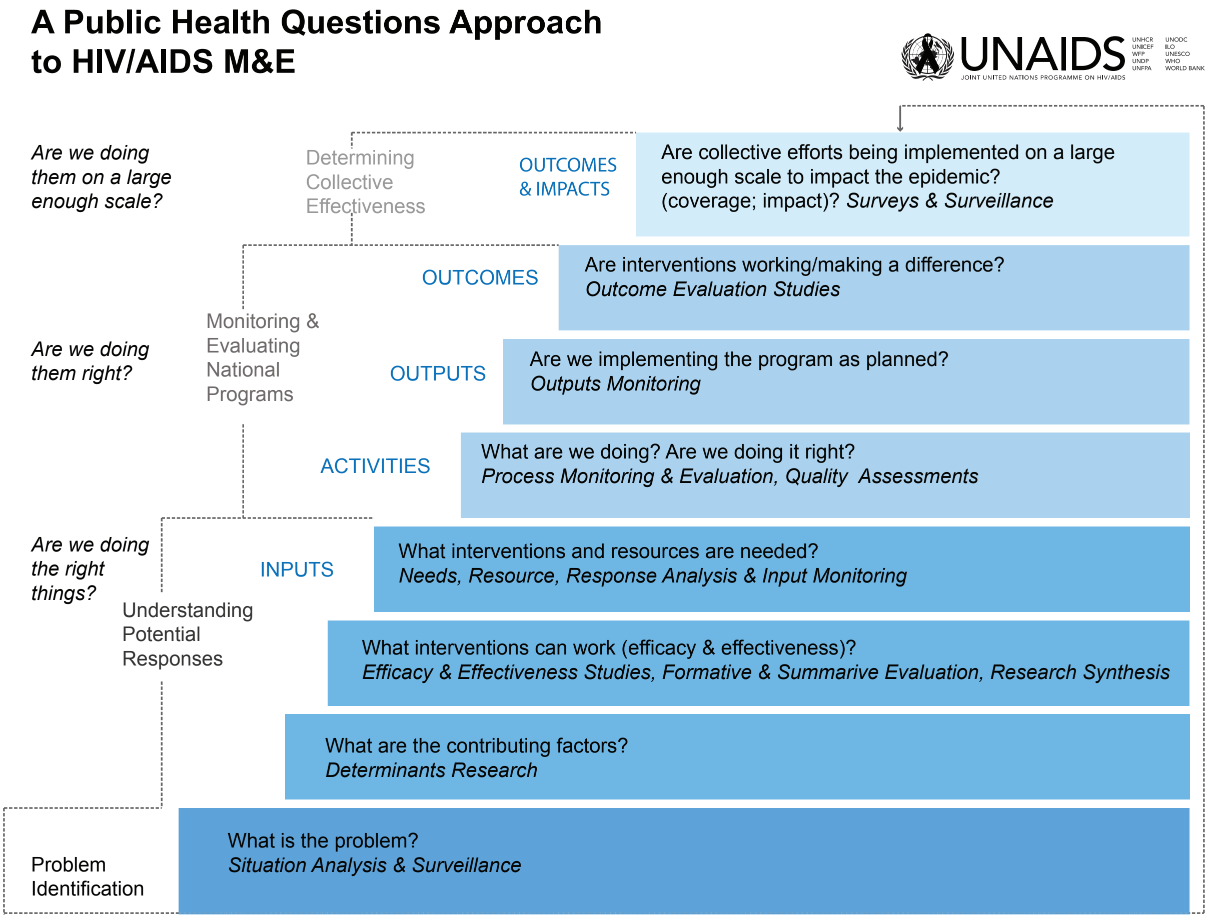
Box 1: Glossary of terms related to program evaluation[1]

- **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 programs/ interventions. Impacts on a population seldom can be attributed to a single program/intervention; therefore, an evaluation of impacts on a population generally entails a rigorous design that assesses the combined effects of a number of programs/interventions for most-at-risk populations.
- **Outcome evaluation**—a type of evaluation that is concerned with determining if, and by how much, program activities or services achieved their intended outcomes among the targeted population. Whereas outcome monitoring is helpful and necessary in knowing whether outcomes were attained, outcome evaluation attempts to : a) attribute observed changes among the targeted population to the intervention tested; b) describe the extent or scope of program outcomes; and c) indicate what might happen in the absence of the program. An outcome evaluation is methodologically rigorous and generally requires a comparative element in design, such as a control or comparison group, although it is possible to use statistical techniques in some instances when control groups are not available (e.g., for a national program).
- **Outcome monitoring**—the basic tracking of variables that have been adopted as measures or "indicators" of the desired program outcomes. Outcome monitoring does not infer causality; changes could be attributable to multiple factors, not just the program. With national AIDS programs, outcome and impact monitoring is often conducted through disease surveillance data reporting and population-based surveys (representative of the target population, not necessarily the general population) to track trends of outcome and impact level indicators.
- **Input and output monitoring**—the basic tracking of information about program inputs or resources that go into a program, and about outputs of the program activities. Data sources for monitoring inputs and outputs usually exist in program documentation (e.g., activity reports, logs) and client records, which offer details about the time, place, and amount of services delivered as well as the types of clients receiving services.
- **Program evaluation**—a systematic assessment of the means and the ends of some or all stages of a program, including planning, implementation, and outcome, to determine the value of and to improve the program.

This guide deals mainly with the realm of Outcome Monitoring and Outcome Evaluation with the understanding that impact evaluation takes place at a higher level than what most programs can measure directly. This level of evaluation is represented in the top two squares of the UNAIDS M&E Framework (see Figure 1), which is supported by major international partners and donors [3].

[1] UNAIDS Monitoring and Evaluation Reference Group. Monitoring and Evaluation Terminology Glossary, 2008

Figure 1: UNAIDS M&E Framework



Understanding the distinction between the "M" and the "E" in M&E

At the most basic level, program evaluation can be said to begin with monitoring programs. When we monitor programs, we are asking whether program activities have been implemented as planned.

- Assuming that the various inputs have been put in place (i.e., resources put into programs, staff hired, facilities arranged, supplies procured, etc.), then the relevant monitoring questions are in relation to outputs, such as "Are services are being delivered as intended?"

- Answering these questions involves input and output monitoring and quality assessments on an ongoing basis during project implementation. See examples of input and output indicators in Box 2.

Box 2: Examples of input and output indicators

- Number of outreach events (e.g., workshops conducted, support group meetings held, drama or street theatre shows)
- Number of products distributed (e.g., male and female condoms, lubricant, needles and syringes)
- Number of individual clients receiving products through distribution
- Number of clients participating in outreach events (e.g., condom demonstrations, educational sessions, workshops, support meetings)
- Number of staff trained (e.g., in outreach, couple counseling, STI treatment)
- Number of information contacts (e.g., calls to a hotline, hits on a website, pamphlets distributed)
- Number of media events (e.g., radio programs, television announcements)
- Number of services provided (e.g., HIV tests conducted, methadone doses provided)
- Number of project sites (e.g., STI clinics, drug treatment centers, drop-in centers)
- Number of clients contacted (e.g., through community outreach, at drop-in centers)
- Number of clients receiving services (e.g., IDUs in drug treatment, FSWs screened for STIs)
- Number of clients referred to services (e.g., HIV counseling and testing, HIV treatment services, psychosocial support services)

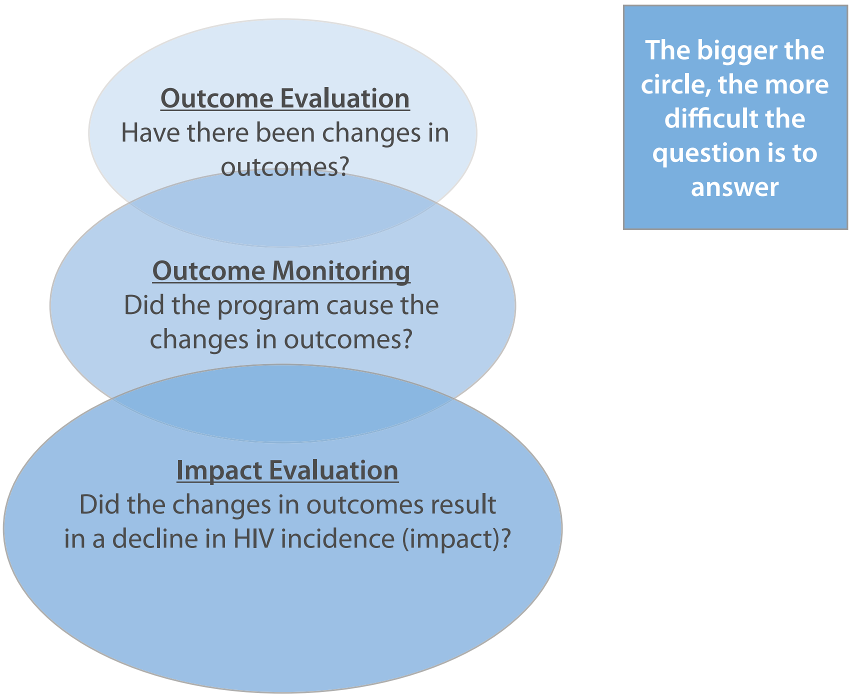
Outcome and impact level evaluation questions

Outcome and impact evaluation look at a more complex set of questions that ultimately will assess whether the program is performing "successfully".

Outcome monitoring and evaluation: At the outcome level, the overarching evaluation questions are in relation to whether expected outcomes have been achieved (outcome monitoring), and what caused them to be achieved (outcome evaluation). These two levels are represented in the top two circles in Figure 2.

Impact evaluation: This evaluation addresses the question of whether there is evidence that changes in outcomes (caused by the program) resulted in the desired impact. This level is represented in the bottom circle in Figure 2.

Figure 2: Outcome and impact evaluation questions



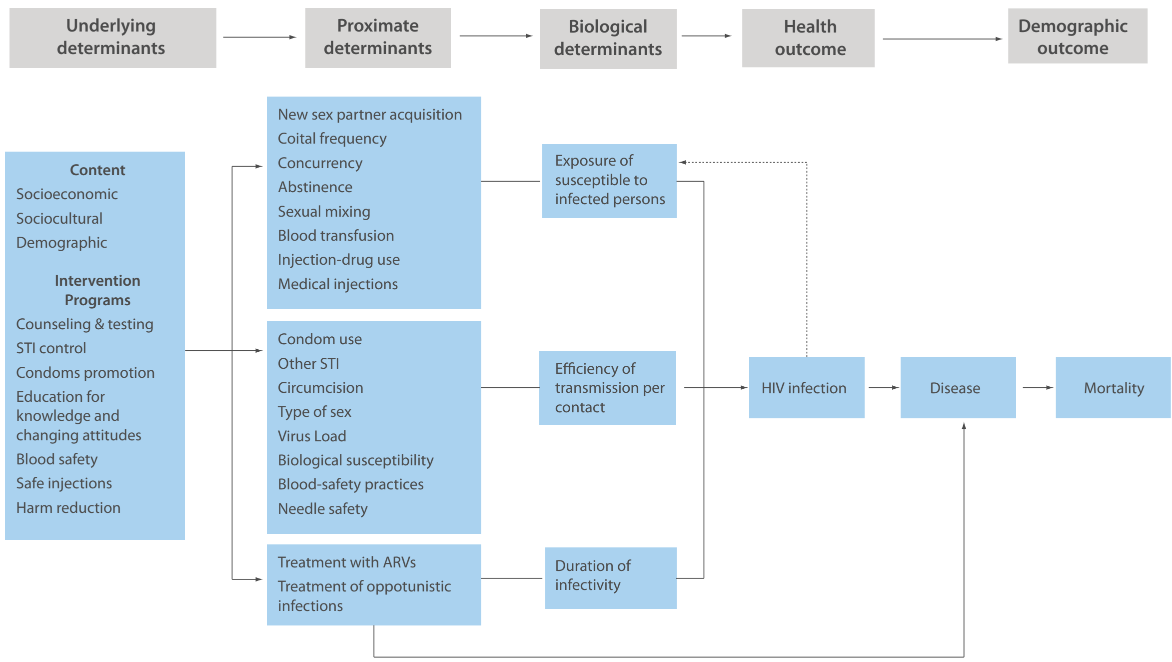
Steps involved in answering outcome and impact level evaluation questions

Starting with the question of whether expected outcomes were achieved, Chapter 3 discusses the basis for selection of outcomes that are relevant in the context of HIV prevention programs, and ways to measure them. Then it moves to the question of whether changes in outcomes are caused by programs and the challenges of attributing effects to specific programs. Chapter 4 explores the challenges involved in conducting impact evaluation, due, in large part, to the difficulties in interpreting HIV prevalence trends, and isolating the part of the trend that reflects the effects of the intervention. In Chapter 5, a framework is introduced to guide the process of triangulating data to explore evidence of program contribution to changes in outcomes (i.e., outcome evaluation). Chapter 6 then explores the kinds of data sources and data collection tools that are typically used to provide this type of information, with a focus on surveillance as the main source of outcome level data, and its strengths and limitations for the purpose of outcome evaluation. This role of surveillance is in line with the "Third One" of the "Three Ones" principle which is supported by UNAIDS and key donor partners. This principle is the implementation of one integrated national M&E plan with dedicated personnel and resources for a single M&E organizing committee and national indicator database, which is the basis for National level M&E in the context of the "Three Ones" [7]. Finally, in Chapter 7, a series of case studies using data from the Asia region illustrates how, in the absence of "experimental research", we can rely on the strength of evidence that comes from synthesizing data from many sources to interpret the situation.

Chapter 3: The Proximate Determinants Framework

If we begin with the question of how we measure changes in outcomes, we need to decide which outcomes are important to measure. The proximate determinants framework shown in Figure 3 [8] provides a good basis for defining which outcomes are important to track, as a measure of progress toward lowering HIV transmission.

Figure 3: HIV transmission and risk factors: The proximate-determinants framework



Source: Proximate determinants conceptual framework for factors affecting the risk of sexual transmission of HIV, ARVs, antiretrovirals; STI, sexually transmitted infection.

The determinants in this framework (Figure 3) correspond to the main outcomes of interest for prevention programs for high risk groups. In the framework, the determinants are grouped according to which biological determinant they affect (for example, number of new sex partners affects how often an uninfected person is exposed to an infected person, and condom use affects efficiency of transmission during the contact between the infected and uninfected person).

Proximate determinants

- Some proximate determinants affect exposure of uninfected to infected persons
  - Number of new sex partners
  - Number of sex acts (especially with high risk partners)
  - Number of concurrent partners
  - Frequency of injecting drug use, unsafe medical injections or transfusions
  - HIV prevalence in the population (and high risk sub-population)

- Some proximate determinants affect the efficiency of transmission per contact
  - Condom use
  - Use of clean needles during injection
  - Presence of other sexually transmitted infections (STIs)
  - Circumcision
  - Type of sex (e.g., anal sex more risky than vaginal sex)
  - Viral load
  - Biological susceptibility (e.g., greater area of genital exposure in women)
  - Blood safety practices

- Some proximate determinants affect the duration of infectivity of people with HIV
  - Treatment with ARVs
  - Treatment of opportunistic infections

Underlying determinants

Underlying determinants are not directly "causal" but are critical because of their potential influence on proximate determinants. HIV prevention programs work by influencing changes in proximate determinants which are affected by the country-specific environment. Measuring the extent to which outcome indicators change as result of exposure to intervention programs is one of the main ways of substantiating the plausibility of program effects.

In addition to programmatic factors, other underlying factors include structural, societal and environmental factors (e.g., infrastructure of a comprehensive National AIDS Program):

- Contextual factors
  - Socioeconomic
  - Sociocultural
  - Demographic
- Availability of and exposure to comprehensive and effective HIV prevention, care and treatment interventions

Measuring outcomes

Conversion of proximate and underlying determinants into outcome indicators

A few examples of indicators that can measure changes in proximate and underlying determinants are shown in Box 3.

Box 3: Converting determinants to indicators

Proximate Determinants	Example Indicators
• # of new sex partners	• Decrease in % of men who have sex with men (MSM) who had more than 3 anal sex partners in the past one month
• Condom uses	• Increase in % of sex workers who used a condom the last time they had sexual intercourse with a client
Underlying Determinants	
• Socio-cultural factors	• Decrease in % of sex workers who have been harassed by police for carrying a condom in the past 6 months
• STI Control Program	• % of sex workers who visited the NGO clinic and were screened for STIs

Sources of outcome and impact data

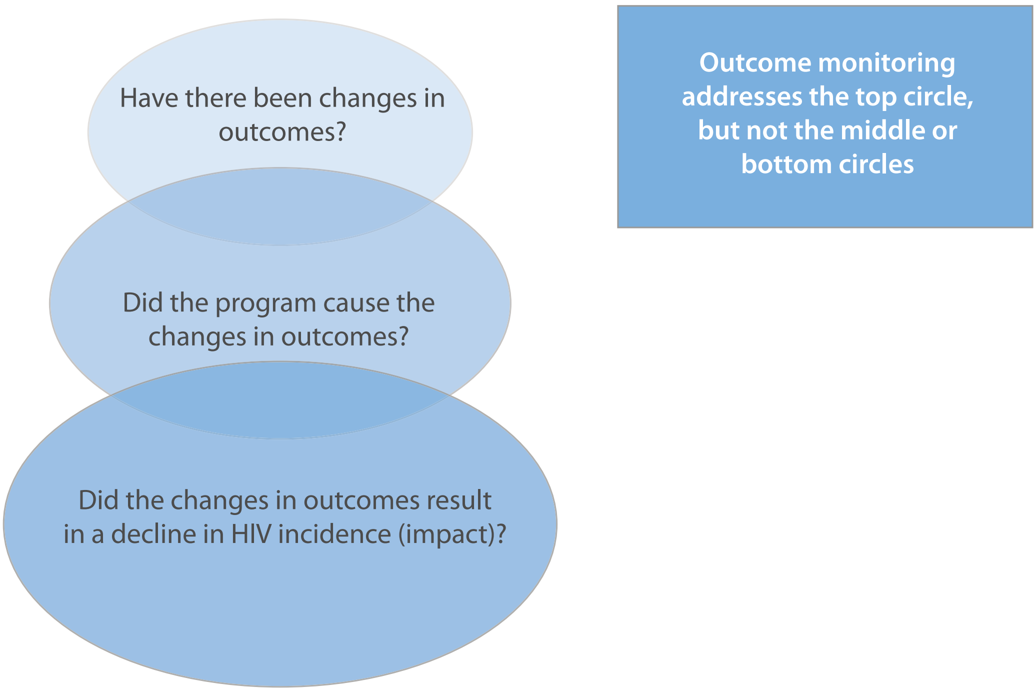
Many outcome indicators can be directly measured through behavioral surveys (e.g., number of new sex partners, number of unsafe sex acts [especially with high risk partners or anal sex acts], number of concurrent partners, or frequency of unsafe injection. Or they may be measured through biological surveys (e.g., prevalence of syphilis, gonorrhoea, chlamydia, herpes simplex virus (HSV), hepatitis C virus (HCV)). HIV prevalence trends are also measured through biological surveys.

Behavioral trend data often come from population based surveillance surveys such as behavioral surveillance surveys (BSS) or integrated bio-behavioral surveys (IBBS) among high-risk populations. Biological surveillance may come either from facility-based surveys, usually called HIV Sentinel Surveillance (HSS), or from integrated bio-behavioral surveys among high-risk populations (IBBS). Facility-based surveillance surveys may come from antenatal clinics (ANC), sexually transmitted infection (STI), or non-governmental organization (NGO) service delivery points (e.g., clinics or drop-in centers). Sometimes surveillance among IDUs is conducted in drug rehabilitation facilities or in prisons.

Attribution

As has already been suggested, outcome level monitoring is not meant to establish or prove causal links between programs and changes in outcomes (i.e., it cannot answer the question "did the program cause the changes in outcome?"). This is because there may be multiple programs operating in the same location working with the same population, or factors other than the program(s) may be influencing changes in outcomes; these factors make it difficult to separate out the effect of a specific program. For this purpose, appropriately designed evaluation research (i.e., randomized trials or experimental studies with control groups) are required (see Figure 4).

Figure 4: What questions can be answered with outcome monitoring?



- It does not establish/prove casual links between programs and changes in outcomes
  - » There may be multiple programs operating in the same location, working with the same populationl; or
  - » Factors other than the programme(s) may be influencing change
- It does not measure impact of programme on HIV transmission

For these purposes, appropriately designed evaluation research (i.e., randomized trials or experimental studies with control groups) are required.

Box 4 summarizes the major study designs that can establish causality, starting with experimental designs, which are the most rigorous, and working down to non-experimental studies, which are the least rigorous.

Box 4: Study Designs

Strength of evidence of program effectiveness

<u>Study Designs</u>

Experimental studies

- Randomized controlled studies

Quasi-experimental studies

- Controlled trials (subjects assigned to interventions but with no randomization)

Observation/non-experimental studies

- Cohort studies with concurrent controls
- Cohort studies with historical controls
- Case-control studies
- Repeated cross-sectional surveys with no control group (open populations)
- Pre- and post-intervention groups (intervened group only)

Strongest ↓ **Weakest**

The main difference among these designs is in the strength of evidence they provide about program effectiveness. Outcome monitoring that uses the types of data typically available through repeated cross-sectional surveys, such as the kind that are usually available to programs through surveillance systems, fall into the weakest category. In an ideal world, assessing program effectiveness would involve experimental evaluation research. The main measure of effectiveness in the context of HIV prevention would be the reduction of HIV incidence, and evaluation would involve comparing HIV incidence in intervened and non-intervened susceptible populations over time. If incidence declined more quickly in the intervention group than in the control group, this would suggest that the program was effective.

However, in the real world, rarely do we have the ability to use experimental designs to evaluate program effectiveness. Rarely do we have control groups to serve as the basis for comparison, and rarely do we have the ability to directly measure HIV incidence. Instead we rely on behavioral and HIV prevalence trend data from repeat cross-sectional surveys to provide evidence of effectiveness, and we reserve more rigorous research for important unanswered questions. The strength of evidence from these types of surveys is compromised, but because it is often the best we have, we need to understand how to work with it. This means understanding how to interpret trends using data triangulation to understand the truth beneath the data.

Chapter 4: Interpreting Trends and Evaluating Impact

Key points in this chapter

- Impact evaluation requires evidence of declining HIV incidence, which is almost never available. Instead we rely on prevalence trends, which are influenced by many factors, in addition to HIV incidence rates.
- It is crucial to understand the factors that can influence prevalence trends, besides incidence, including sampling bias, population turnover, and mortality and treatment effects.
- Population movement and sampling bias are moving parts of the "trend equation" that are unknown, immeasurable, and may be exerting considerable influence on HIV prevalence trends, sometimes in opposite direction.
- Interpreting prevalence trends requires assessing these biases.

Before attempting to link programs to declines in incidence, we must first establish the likelihood that incidence actually is declining, by ruling out other explanations.

We rely heavily on behavioral and biological trend data from repeat cross-sectional surveys to provide evidence of effectiveness. In this chapter the challenges involved in interpreting those trend data, especially biological trends are discussed. Trends are important to understand, because before asking whether changes in outcomes resulted in declines in HIV incidence (the impact evaluation question in the bottom circle of Figure 2 and Figure 4), we must first examine the evidence that there actually was a decline in incidence. This is not a straightforward task, which is why it is not undertaken on a routine basis.

The answers to the core impact evaluation questions for HIV prevention programs all require trend data:

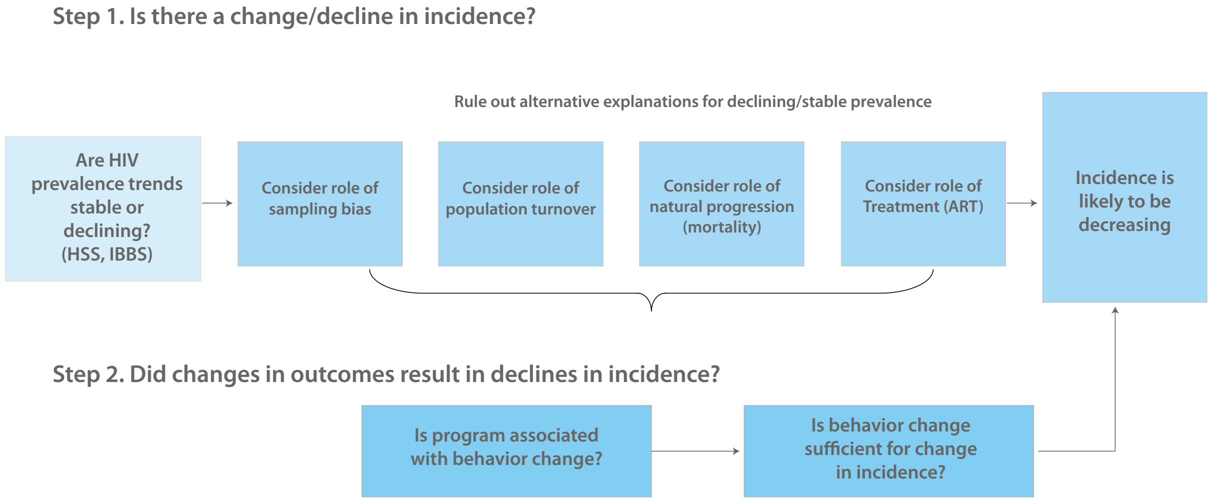
- Did the rate of new HIV infections decline?
- Was the decline caused by changes in outcomes?
- Were the changes in outcomes caused by programs?

The interpretation of trends (particularly HIV prevalence trends) can be complex and therefore challenging in the context of data triangulation.

Challenges faced in measuring and interpreting trends

Establishing incidence declines is a challenge. Since it is very difficult to measure incidence directly, we often use HIV prevalence trends to understand the direction of the epidemic. We tend to interpret stable or declining prevalence trends as a sign that HIV incidence is also declining. However, as depicted in Figure 5, the declines in HIV prevalence can also reflect other factors, including sampling bias, population turnover, mortality, and treatment effects. Any one of these factors can significantly alter the prevalence curve, with or without changes in HIV incidence, making it very difficult to isolate actual HIV transmission tendencies.

Figure 5: Factors influencing HIV prevalence trends



All of these factors must be taken into account when assessing trends. It is only after ruling out these other factors that we can rightly ask the question of whether HIV incidence has declined, and if so, whether programs have contributed to those declines. We do this by first looking at the evidence that programs influenced changes in outcomes, and then whether the changes in outcomes can be linked to declines in incidence.

Practically speaking, how do we measure these factors? Two of the most difficult issues to contend with are sampling bias and population turnover. These are a difficult phenomenon to deal with because they are not directly measurable, and they can dramatically affect the prevalence trends.

How can population movement affect trends?

If we think of the population as a "pool" of people actively engaged in a risk behavior (e.g., selling sex, injecting drugs), over time, the pool may be dynamic or static (i.e., people may get in and out of the pool, or they may stay in the pool throughout). During successive rounds of surveillance, the number of people in the pool might increase, decrease, or stay the same. If the size of the pool increases, the people in the pool may be a mixture of people who were in the pool before plus some new people, or it might be all new people. Likewise, if the size of the pool decreases, the people in the pool may be only people who were in the pool before, it might be all new people, or a mixture of the two. If HIV positive or HIV negative people are not equally likely to get in or out of the pool, this can cause "spurious" increases or declines in the HIV prevalence trends, (spurious, meaning for reasons other than actual declining incidence or increasing mortality).

How does sampling bias affect trends?

Sampling bias affects prevalence trends in similar ways as population movement, but instead of affecting who is in the pool, it affects who from the pool is selected into the sample. The sample should be a representative subset of the pool, but depending on the way sampling takes place, that may or may not happen. If HIV positive or HIV negative people in the pool are not equally likely to get into the sample, this can cause spurious increases or declines in the HIV prevalence trends.

Population movement and sampling bias are moving parts of the "trend equation" that are unknown, immeasurable, and may be exerting influence on HIV prevalence trends in opposite directions at the same time.

Other factors that affect ability to interpret trends: phase of the epidemic

Aside from problems with sampling and population movement, underlying epidemic dynamics play a strong role in determining the shape of HIV prevalence curves. This is related to the natural epidemic progression. In early epidemics, the prevalence of HIV can rise even if the incidence of HIV is slowing. This is because the time period from infection to death can be several years, and it continues to increase because of ARV treatment effects. This means that in early epidemics, HIV prevalence trends will not necessarily reflect reduced incidence. In mature HIV epidemics, high rates of new infections (incident infections) can be balanced by high rates of people dying of HIV disease (mortality) so that the actual number of people living with HIV (i.e., prevalent HIV infections) might not change much at all and appear stable. Since people with HIV being treated effectively with ARVs can be expected to stay alive much longer, prevalence may actually increase even if prevention programs are reducing the number new HIV infections.

Interpreting trends in light of the large potential for bias

Interpreting behavioral and biological trends is a process that requires familiarity with local context. Behavioral trends are somewhat more straightforward to interpret than biological trends. Although sampling bias and population movement can still affect behavioral trends, the underlying epidemic dynamics are less of a factor. Interpreting biological (e.g., HIV) trends requires assessing sampling bias and bias that may have resulted from population movement. It is important also to consider the phase of the epidemic.

Chapter 5: From Surveillance to Evaluation: A Framework for Triangulation

Key points in this chapter

- The term "triangulation" is used in this guide to describe the process of using multiple data sources to examine hypotheses about program effectiveness.
- The SUCCESS framework is a systematic approach for triangulating data for program evaluation that serves as "checklist" for making sure the major factors for outcome evaluation have been assessed.
- Behavioral and biological surveillance surveys are a common source of outcome level trend data used for outcome and impact evaluation.

Defining triangulation

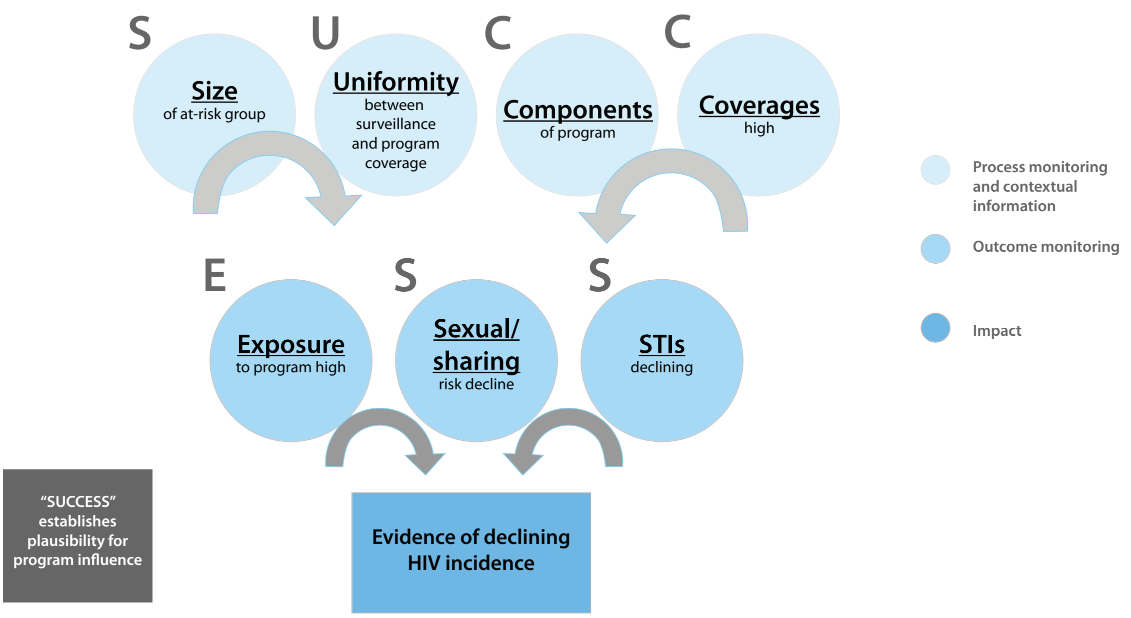
What is it and why do we do it?

Triangulation, as it is used in this guide, is the process of using multiple data sources to provide evidence of program effectiveness in the absence of experimental studies. The goal of triangulation in the program evaluation context is 1) to make plausible linkages between program activities and changes in outcomes, and 2) to make plausible linkages between changes in outcomes and declines in incidence.

The SUCCESS framework

The "SUCCESS framework" depicted in Figure 6 suggests a systematic approach to triangulation for understanding program effects in standard service delivery contexts where priority is placed on program evaluation and documentation of progress and contributions rather than on rigorous evaluation (which can prove to be more resource intensive).

Figure 6: The SUCCESS framework



The components of SUCCESS

The next section describes the components of SUCCESS. In brief, the letters of the acronym SUCCESS stand for **S**ize, **U**niformity, **C**omponents, **C**overage, **E**xposure, **S**exual and Sharing behavior, and **S**TIs. This list serves as a kind of checklist, with each letter representing an important aspect of the triangulation analysis (not necessarily in order).

Size refers to the number of people in a risk population. Population size is the piece of the triangulation puzzle that puts all the other pieces in perspective, which makes it very important.

In the context of SUCCESS, specifically, size plays a role in providing a denominator for measuring program coverage for the population at-risk. Some programs aim to reach 100% of a particular risk population in a specific geographic area, and some only a portion. Either way, size is essential for determining the number of people to be reached, and for understanding the proportion that has been covered.

Uniformity refers to the degree of overlap between the different sources of data used for triangulation, with regard to time, place and definition of the population intervention. For example, you may have very good data showing that your program had excellent coverage against your program targets. But if the outcome data you have available do not represent the population covered by your program, it is difficult to state that the program was responsible for the changes in outcomes. Likewise you may have very good data showing changes in outcomes in your program coverage area, but if the timing of the changes preceded program implementation, or happened too soon after the intervention started to have been caused by the intervention, it is difficult to establish the plausibility of program effect.

The problem of uniformity can come into play when surveillance data are used for outcome monitoring, because site and group selection for surveillance is not necessarily designed with intervention coverage in mind. Similarly, when monitoring trends, definitions of populations and definitions of indicators must be uniform over time, otherwise the trend information becomes difficult to interpret.

Components refers to the appropriateness of the mix of interventions included in the response for a particular group, given their stage in the epidemic. For example, a program for sex workers that included only HIV counseling and testing (HCT) might not be expected to make the same kind of impact as one that included also condom distribution and STI treatment.

Coverage refers to the proportion of the target population that has been reached by the prevention program. "Reach" can be defined in many different ways, so it is important to spell out the program targets, in terms of the number of people to be reached by different activities and the level of intensity with which they should be reached. Clarifying the relevant denominator is the second key aspect of examining coverage. For example, a program may be contracted to provide services to 1000 sex workers, measured in terms of who visits the NGO STI clinic at least once every quarter. If, at the end of the program, it can be shown that 700 sex workers visited the clinic every quarter, the program can be said to have achieved 70% of its target. This sounds like a lot, and it is, in terms of the program's goals. But good coverage at the program level does not necessarily translate into good coverage at the population level. It is critical to know what the overall size of the vulnerable population is -- not just the size of the population the program is expected to cover. If the program exists in a city where there are 5,000 sex workers, then the program would have achieved only 14% (700/5,000) of the coverage needed to make an impact on the epidemic. So although the program may be doing quite well in terms of its own goals, it may not be well positioned to make a large impact on the epidemic. Overall coverage cannot be determined without estimating the population sizes of the right geographic scope, to use as a denominator (refer also to discussion on size).

Exposure, like coverage, also refers to the proportion of the target population that has been reached by the program. The difference is in the way it is measured. Population based measures of exposure can be obtained through behavior surveys by asking respondents a series of questions about their contact with interventions and use of services. Program exposure is considered an underlying determinant, because it influences the proximate determinants[8]. To the extent that surveillance surveys can be tailored to measure exposure to specific aspects of the programmatic response, as well as frequency and duration of the exposure, they can also be used to provide evidence of program effectiveness

Sexual and Sharing Risk: Unsafe sex and unsafe injecting are two of the most important proximate determinants of HIV transmission. Thus changes in these behaviors are a central part of outcome level monitoring. Measures of risk behavior can be obtained through program specific behavioral surveys or through behavioral surveillance surveys (BSS) or integrated bio-behavioral surveys (IBBS). The outcome variables include number and type of sexual partners, number of sex acts by partner type, condom use by partner type, concurrency in sexual partners, type of sex, and frequency of sharing injecting equipment, among others. These variables can be measured at the population level through program-specific surveys (i.e., in coverage areas specific to program) or through surveillance surveys. In the latter case, the extent to which the surveys reflect the reach and effectiveness of programs depends on the degree of uniformity between the surveillance coverage and program coverage in terms of geographic areas and population definitions.

Sexually Transmitted Infection - The final "S" in SUCCESS stands for risk related to STIs. STIs are another important proximate determinant because they directly affect the probability of sexual transmission of HIV. Declines in the incidence of curable (e.g., bacterial) STIs can be difficult to measure through surveys, because the prevalence at any given point in time may not reflect the incidence over a longer period of time such as the past year. This is especially true for men, who tend to be symptomatic more often than women, and may therefore be less likely to have prevalent infection. For this reason it is important to also use reported case data from STI clinics, if acceptable quality data are available, as part of the SUCCESS data triangulation process. Other indicators that are also useful because of their close relationship with STIs are those related to STI treatment seeking, including type of treatment sought and time between onset of symptoms and seeking treatment. These variables are sometimes available through behavioral surveys.

Chapter 6: Preparing for Data Triangulation

Key points in this chapter

- Priority populations for HIV prevention interventions in the Asia-Pacific region is largely comprised of individuals with higher risk behaviors and their sexual partners.
- Types of data that can be used for a SUCCESS analysis are reviewed in this chapter, as well as some of the measurement challenges.

Who are the priority populations in Asia?

The Asia-Pacific region is characterized by low-level and concentrated epidemics [3]. HIV infections are largely confined to individuals with higher risk behaviors and their sexual partners. Behaviors that put people at greater risk for HIV infection, as described in the proximate determinants framework in Figure 3 from Chapter 2, include multiple sexual partnerships, unprotected sex (especially anal sex) with multiple partners, and injecting drugs with shared injecting equipment. Throughout the region HIV infections tend to be highest among:

- *Female sex workers (FSWs)*
- *Injecting drug users (IDUs)*
- *Men who have sex with men (MSM), including male sex workers (MSWs) and transgenders (TGs)*
- *Clients of female sex workers*

These populations are often linked behaviorally. In many locations throughout Asia, they are the priority populations for HIV prevention efforts. When implemented at scale targeted interventions focused on these populations can result in far fewer HIV infections, both in these higher risk populations, and in the population at large [3, 9-12].

When should we do triangulation?

Triangulation can take place at any time, but some key junctures when it may be most useful for program evaluation purposes include:

- When multiple data sources exist with sufficient quality for analysis;
- When program activities have been realized for a significant time period, and there is a desire to examine effectiveness;
- When changes in HIV prevalence or HIV-related outcomes observed through surveillance or other surveys suggest that there is a decline in incidence, and you want to explore program contribution.

In the last case, it can be very tempting to attribute observed changes to prevention programs, but "examining outcome/impact indicators without assessing the process of program implementation can lead to erroneous conclusions regarding the effectiveness of the intervention" [13].

Given that these guidelines address most-at-risk populations, a general rule of thumb is that triangulation analysis should be undertaken when interventions have been in place for a sufficient duration of time (e.g., 1-2 years) to reasonably expect that changes in proximate determinants at the population level may be attributable to program interventions. The timeframe of one to two years of intervention may be sufficient to be able to look at changes in behavioral trends. For HIV prevalence trends among high risk groups, a longer period of time (e.g., 4 years or more) is needed before such changes can plausibly be attributed to program interventions.

What data will you need?

Before conducting triangulation analysis, you will need to gather several types of data from different sources which may include both program-level and surveillance or other outcome/impact data on most at risk populations, such as:

Program data

- Information on program geographic coverage (planned program components and targets);
- Size estimates for the targeted populations in the program coverage area and in the overall area;
- Input and output monitoring data;
- Program coverage estimates;
- Information about quality of intervention implementation.

Surveillance data

- Behavioral trends from BSS or IBBS;
- STI trends from routine clinic data and from IBBS prevalence surveys;
- HIV trends from HIV Sentinel Surveillance (HSS) or IBBS.

Other

- Information on phase of the epidemic;
- Information on patterns of HIV spread in different geographic areas.

What should be the geographic area covered by the triangulation analysis?

It is important to pay attention to epidemic phase and geographic zone when triangulating data to explore the link between programs and outcomes. This issue relates to the **uniformity** aspect of measuring SUCCESS. Epidemics progress at different times in different places, depending on when the virus is introduced (i.e., start time of epidemic), and how rapidly it spreads (i.e., volume of risky contact between susceptible and infected individuals). When data are aggregated over geographic areas that are in different epidemic phases, it is easy to miss important local patterns because it becomes more difficult to discern how much change in prevalence is due to epidemic phase, as opposed to possible program induced change. Likewise, if data are aggregated across areas with and without effective interventions, or across areas where interventions were introduced at different times, similar difficulties can arise in interpreting the trends.

Data sources for SUCCESS

Size estimates

Sources: Size estimates of different high risk populations normally come from a variety of sources including:

- **Mapping** of high-risk populations, undertaken by NGOs and/or National AIDS Programs using a variety of methods (e.g., geographic mapping, social mapping, and rapid situation assessments). Such mapping is usually done to help design intervention programs and help identify coverage needs, but it also provides useful information about the size of the population;
- **Surveys** incorporating multiplier methods or network scale-up methods;
- **Capture-recapture** surveys.

Measurement issues: Different methods tend to give different size estimates because of the inherent difficulties of implementing the methods and the inevitable bias that comes from trying to identify and access hidden populations. Biases commonly result from over-reliance on NGOs who serve the communities involved, or from a variety of other implementation issues. One way to avoid bias is to consider a range of high and low estimates when you use the data to assess coverage [14].

Coverage estimates

Sources: Routine program monitoring data (inputs and outputs) are the major source of triangulation data for estimating intervention coverage. They are used to help capture whether the program is meeting its targets according to plan. These data are also used for triangulation purposes in order to assess the extent to which the risk population has been reached by the program, and with what types of services. Without this type of data it is difficult to substantiate program effectiveness (see Box 2: Examples of input and output indicators in Chapter 2).

Measurement issues: The type of indicators collected by programs depends on the activities they implement, and the type of information they need to manage. If the data are to be used to measure coverage, there are several measurement issues that must be addressed.

To illustrate the point, consider a program for male sex workers with the following targets:

- Weekly target: outreach workers contact at least 5 new MSW clients (per week)
- Annual target: at least 50% of male sex workers will have been reached by the end of the first year of program implementation.

Several issues are involved in measuring these targets.

- The weekly target involves measuring the number of clients reached per outreach worker. This is an absolute number that is useful for the program, but it does not help measure the overall contribution made by the program at the population level (i.e., the annual target).
- The annual target is a proportion that needs to take the entire population into account. Calculating the annual target requires the summation of the weekly targets over a year's time for the numerator. But it also requires a denominator which is the size of the total target population in the catchment area of the program.

Measurement issues for the *numerator*:

- Requires clear eligibility criteria to be used for the client population (i.e., who is included and who is excluded);
- Requires clear guidance on what constitutes a client contact;
- Requires method of verification that clients contacted by each outreach worker are actually new to the program, and not just new to that particular outreach worker, otherwise the numerator could be overestimated

The numerator is calculated assuming that the number of individuals contacted by the program can be counted and the numbers can be accumulated over a one year time period. But calculating the proportion covered (i.e., 50% coverage) requires a denominator, which comes from the size estimate measurement.

Measurement issues for the *denominator*:

- Ensures accuracy and consistency of available size estimates (usually there will be a range);
- Ensures the population for the size estimate is defined in the same way as "client" is defined by the program (i.e., good match between the numerator and the denominator). For example, if a size estimate is for all men who have sex with men, and the program targets only male sex workers, the calculation for proportion of male sex workers reached will not be accurate;

- Allows for flexibility as far as coverage area is concerned, as it is not strictly necessary that the coverage area for the size estimate and the coverage area for the program to have 100% overlap. For example, a program may intend to cover only half of the male sex workers in a city, knowing that the other half will be covered by other NGOs. In this case, the program would need either a separate size estimate for the portion of the city they were covering, or they would need to account for their intended coverage in the calculation. In this example, covering 50% of their target population would mean covering 25% of the total number of male sex workers in the city.

Components

Sources:

- Data on mix of services can come from program documentation combined with routine monitoring data. Does the program offer a comprehensive range of services? Are these the "right" services, considering the epidemic situation, and are clients being reached with all the different services?
- Data on intensity can also come from routine monitoring data, depending on the type of information tracked, and whether it can count the number or frequency with which clients are reached with services. Data from behavioral surveys that measure exposure to interventions can also be helpful for measuring intensity, if they track the number of different components the respondents are exposed to, and the number of times respondents receive the service.
- Quality of services usually requires special data collection efforts. Sometimes this is in the form of qualitative research, for example, one might want to assess the acceptability and accessibility of services to the client population, or the skills of those providing the services. Other times the data collection may be more quantitative, as in measurement of the provision of all aspects of STI services, including examination, diagnosis, and appropriateness of treatment through facility-based surveys. When doing data triangulation, you will need to look for reports from such data collection activities.

Measurement Issues: One of the major challenges with tracking the type of information described above (including coverage data) is collecting it and documenting it in a way that it can be analyzed and used easily. Well designed computerized management information systems (CMIS) are an important feature for insuring that information is accessible in an easily usable format.

Exposure

Sources*:* Information on exposure to programs (and changes in exposure over time) come mostly from population-based surveys. Carefully designed "exposure to intervention" modules can be added to BSS and IBBS surveys to obtain population-based measures of exposure to programs. Sometimes the exposure data can also be used to assess whether people who report exposure to interventions, also report less risky behavior. This is an attractive feature since it can be done with a single round of data. Intensity of exposure can also be looked at with a single round of data, however it is important to remember that sampling just from program sites will not be representative of the broader population.

Measurement Issues*:* If such data are used to compare levels of risk behavior between people who are more exposed or less exposed to interventions, attention must be paid to the temporal relationship between when the behavior was adopted, and when the person was exposed to the intervention program. The "cause and effect" relationship is not always clear in a single cross-sectional survey.

In addition to yes/no questions about program exposure, questions about the frequency of exposure can help assess intensity of intervention coverage.

In order to obtain program-specific coverage data, there have been some cases where respondents are asked to the name of the organization that provided the intervention services or commodities they received. This can be problematic because respondents often have difficulty differentiating the identities of organizations they were exposed to, particularly in areas where many different programs are offering similar services. The use of locally appropriate branded logos or symbols associated with specific intervention programs can sometimes help increase visual recognition, but this type of information is sensitive to recall bias, limiting its utility.

Finally, surveillance surveys may not be useful for measuring exposure when there is not good correspondence between the surveillance coverage area and the program coverage area. This is an issue that relates to the **uniformity** aspect of SUCCESS.

Sexual and sharing risk

Sources*:* Behavioral surveys (BSS or IBBS) are the most common sources of information for measuring changes over time in proximate determinants related to sexual and sharing risk. When careful sampling is employed, these surveys have the advantage of being generalizable to the populations frequently targeted by prevention programs; this type of sampling allows for inferences to be drawn from the results, helping with the SUCCESS analysis. This requires **uniformity** in the way the populations are defined and at least partial overlap in the coverage areas of the program and the survey (the less overlap, the less useful the surveys are for conducting this type of analysis). Sometimes special surveys conducted by programs can also be used for this purpose; although if these surveys include only the beneficiaries of the program, they will be less useful for looking at factors like exposure and overall contribution to changes at the population level.

Measurement Issues: The types of measurement error to be aware of in using these data sources are those related to sampling, data collection, data quality and data analysis. Much has been written on these topics, but a few things worth mentioning are the importance of probability sampling (as opposed to convenience sampling) when you want to infer results to the whole population [15]. The potential for selection bias is always an issue with survey sampling, but the extent and direction can be harder to assess with convenience samples. For this reason, facility-based surveys and non-probability surveys are of limited value for obtaining reliable trends. Poorly designed questionnaires, poorly qualified/trained survey teams, and inadequate supervision can adversely affect the quality of both behavioral and biological data. Failure to analyze data using analysis techniques that are appropriate to the sampling and survey design, and inadequate attention to denominators can be major problems for analysis of quantitative survey data. Given the widespread use of skip patterns in most behavioral survey instruments, it can be difficult to use appropriate denominators or to adequately document who is in the denominator for different variables in a report. Needless to say this can dramatically affect the trends, so it is a detail that requires a lot of attention in data triangulation.

Sexually transmitted infections

Sources: The two major sources for data on incidence and prevalence of sexually transmitted infections are routine clinic data and surveys such as the IBBS. Information on treatment seeking behavior can also be obtained from BSS or IBBS surveys. Indicators related to the adequacy of laboratory testing and STI drugs also provide information related to the potential effectiveness of STI treatment programs.

Measurement issues: There are many potential biases, especially in routine clinic data, that need to be recognized and taken into account. With clinic data, one can never be sure who is being reached and who is being missed. STIs treated at pharmacies or through the informal health sector are difficult to monitor. Even STIs treated by private practitioners can be difficult to monitor if providers do not report systematically. STIs that are treated in public health or NGO clinics can be difficult to diagnose with available tools, so they may be misdiagnosed or underreported, especially in women. Using trends in STI prevalence as a proxy measure for the effectiveness of HIV prevention programs is also problematic because such trends may not be solely attributable to behavior change. Complex interrelationships and processes related to coverage, quality, and effectiveness of treatment guidelines and programs can all impact STI incidence and prevalence, making the trends difficult to interpret.

Chapter 7: Triangulating Data for Program Evaluation: Examples and Case Studies

In this chapter, case studies from several countries in Asia are used to explore the strength of evidence for intervention program effects in different settings. The analysis involves triangulating and synthesizing multiple sources of data/information to explore what programs have achieved.

We begin with a hypothetical example that illustrates a stepwise approach to building a case for program success, using multiple data sources to answer a series of questions and to explore the various biases for consideration when interpreting the information. Then we present a series of case studies selected specifically to demonstrate how surveillance data are used in outcome and impact evaluation. In all cases the working hypothesis is that the program played a role in preventing HIV transmission, and the case studies examine the evidence that supports that hypothesis.

For each case study, the framework presented in Figure 4 in Chapter 2 is the basis for structuring the relevant questions for this type of analysis. These questions are once again summarized in Box 5.

Box 5: Triangulation questions

Broadly speaking, the relevant sequence of questions for demonstrating program contribution to declines in HIV incidence is:

- What is the evidence that outcomes have changed?
- What is the evidence that programs caused the changes in outcomes?
- What is the evidence that changes in outcomes contributed to declines in incidence?

Case Study # 1: Police in Cambodia

Background information

Commercial sex has been the major driver of HIV transmission in Cambodia since the epidemic started in the early 1990s. The Government has prioritized prevention efforts for sex workers and their clients for many years. Since police and military are known to be frequent clients of sex workers, prevention efforts among these two groups was scaled up starting in the late 1990s. Interventions among sex workers started earlier. The interventions in these two groups (sex workers and clients) go hand in hand because changes in outcome indicators in one group are likely to reflect similar changes in the other.

The program data presented in this case study come from a large-scale police program conducted by the Ministry of the Interior. This program started in the year 1999. Peer education activities commenced in 2000, but there were most likely other smaller-scale interventions with police prior to this program. The program by the Ministry of the Interior started in the provinces of Phnom Penh, Kandal, Kompong Chnnang, and Siem Reap. Similar activities were conducted by the Cambodian Red Cross (CRC) in the provinces of Kompong Cham, Battambang, Banteay Meanchey and Pailin. Starting in 2003, the program made efforts to increase in scale and geographic scope and by 2007 it had covered all 24 provinces.

Components

- Between 1999 and 2008, the program with the police focused on outreach, peer education and condom distribution, using a combination of sensitization workshops, interpersonal communication, and counseling by peer educators.
- Provision of STI services serving large numbers of sex workers and clients were also part of the effort during the same time period.
- The approach of the program was to reach police on a quarterly basis with structured educational sessions and to include HIV/AIDS messages during regular weekly meetings.
- Police were also encouraged through peer counseling to go for regular STI screening, whether or not they were symptomatic.

Examining the evidence

Data available for this case study

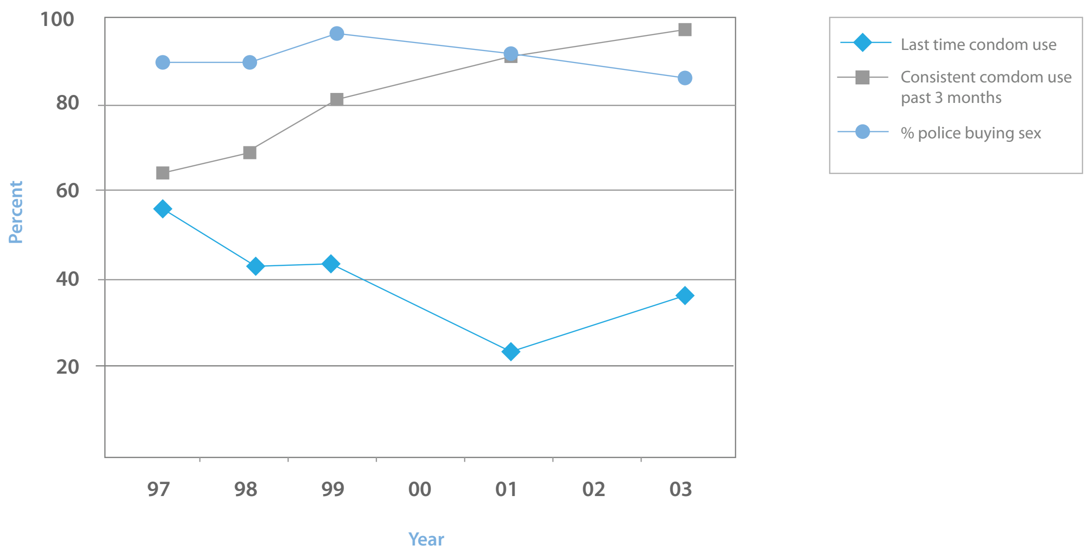
- **Outcome**
  » BSS data from 1997 to 2003 among police in five major cities (Phnom Penh, Kandal, Kompong Chnnang, Battambang, and Siem Reap)
- **Size**
  » Limited information on number of police in 2005 and 2007
- **Output**
  » Limited program data from 1999-2008
- **Impact**
  » HIV prevalence trends for police, sex workers and ANC from sentinel surveillance in virtually all provinces from 1997 to 2003

What is the evidence that outcomes changed?

Sexual Risk Behavior

The BSS data in Figure 7. Condom use among police suggest that between 1997 and 2003 police in Cambodia were adopting safer behaviors. Fewer of them reported buying sex and more reported using condoms when they did. In particular, the proportion of police who reported being clients of sex workers declined from 60% in 1997 to 40% in 2003 and reports of consistent condom use went from just over 60% in 1997 to more than 90% in 2003. To the extent that these reports were accurate, they reflected significant increases in the factors that prevent HIV transmission.

Figure 7: Condom use among police between 1997 and 2003 in Cambodia (Cambodia BSS)



What is the evidence that programs caused the changes in outcomes?

Assessing this evidence involves having good information on the size of the concerned risk populations at the time the program was being implemented, and on the extent to which these populations were being reached or otherwise influenced by programs during the same time period.

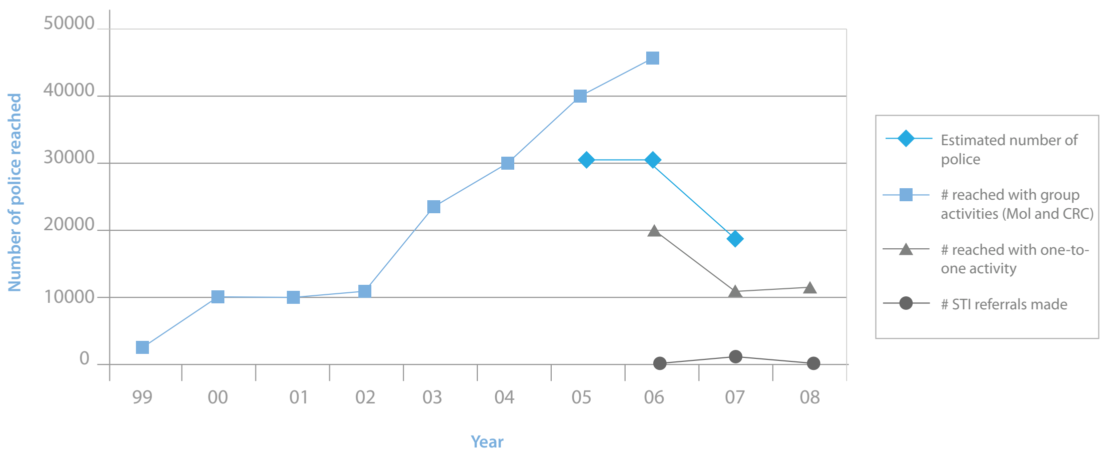
Size

Available information on size of the police force is limited. Some reports put the number at around 31,000 in 2005-2006. One source estimated that there were 17,785 police in 2007, with attrition, combined with a freeze on new recruits being the reason for the shrinking size. The extent to which these estimates correspond to the real size is not known. Information about the size of the police force at an earlier time (i.e., when the intervention began) is also not available.

Coverage

The available program monitoring data were also inconsistent. Only contact through group activities is available for the intervention period prior to 2006. The consistency between program data and size estimates (as in Figure 8) is also problematic. The number of people the program reported reaching through group activities in 2006 far exceeded the total estimated number of police, which calls into question the accuracy of both the program data and the size estimates. However, even taking these quality issues into account, available data demonstrate that there was a very large scale-up of the program between 2002 and 2007, with the vast majority of police being reached (at least with group activities) by 2003. Far fewer men were reached with one-to-one services and very few were reportedly referred for STI services.

Figure 8: Program coverage among police in Cambodia (Program data)

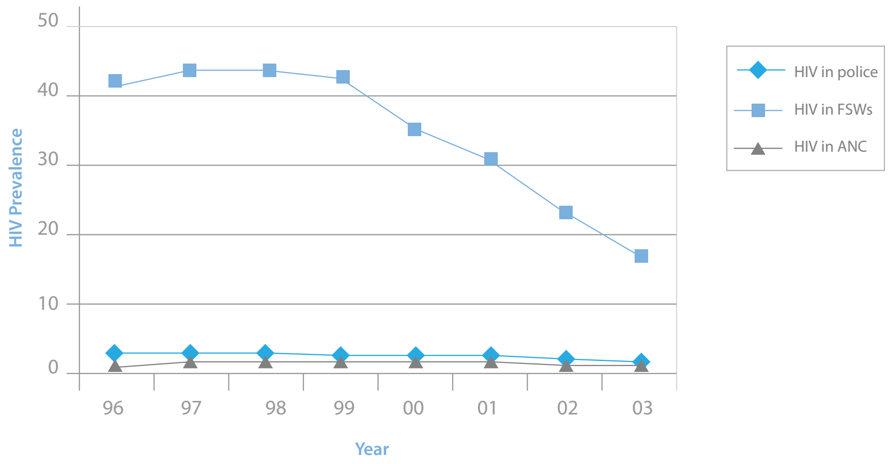


What is the evidence that changes in outcomes contributed to declines in incidence?

HIV surveillance trends

Figure 9 shows that HIV prevalence among police declined from 4.3% to 2.7% between 1996 and 2003 (data from 22-24 provinces). Sentinel surveillance data after 2003 were not available. We see that prevalence among FSWs declined from 40% to 17% during the same time period, and that prevalence among Antenatal Clinic (ANC) attendees started to decline slowly after peaking at around 2.1% in 2000/2001. This evidence is consistent with declining incidence in sex workers, police and possibly their partners (reflected by ANC women) starting in the mid-1990s.

Figure 9: HIV prevalence declined in police, sex workers and ANC women between 1996 and 2003 in Cambodia (Cambodia HSS)



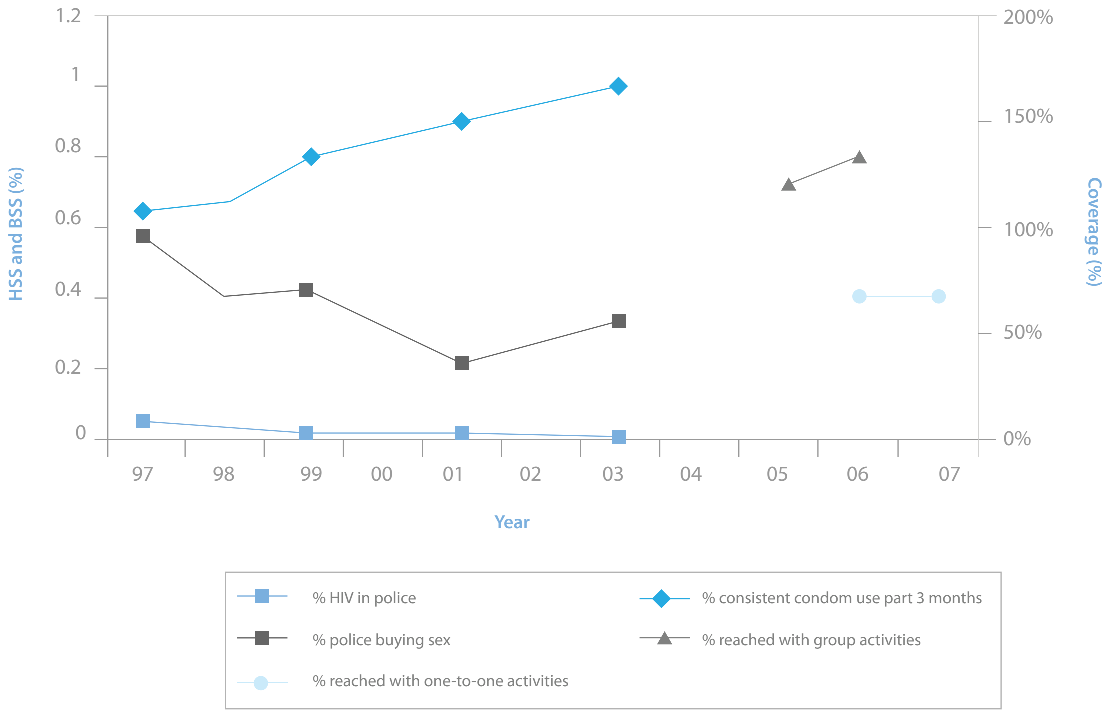
It is important to assess the many potential explanations for the decline in prevalence before concluding that it was due to a drop in incidence.[2] For example, recruitment of police was reportedly frozen during the period of the program and resumed only later (in 2006). So the decline in prevalence could have been due to HIV infected police dropping out of the police force at a higher rate than uninfected police (due to HIV-related illness or death). A better assessment of the likelihood of declines in incidence in all three of these populations might be better achieved with modeling tools, which would account for the effects of epidemic progression (e.g., mortality and treatment effects), and also in-and out-migration, in addition possible declines in incidence. However, application of models is fairly labor intensive and not usually done on a routine basis. The more common practice is to use declining HIV prevalence trends as a "proxy" for declining incidence, which is why it is important to corroborate the evidence by looking for changes in risky behaviors or changes in the composition of the population.

Are the data sources Uniform?

The degree of overlap between the intervention program and the surveillance data -- in terms of the profile of the population being targeted, the geographic location, and the timing of data collection -- is critical for establishing the plausibility of the link between the program, the changes in outcomes, and the presumed incidence declines. The national surveillance system in Cambodia collected behavioral and biological data routinely on police until 2003, and the program relied on those surveys for outcome monitoring. However, the behavioral data were collected in only five cities, whereas the program, which started in nine cities in 2000, had expanded to all 24 provinces by 2007. Furthermore, there were no behavioral or biological data available for this case study after 2003 and most of the program monitoring data was not available until after 2003; therefore there was almost no time point when both biological/behavioral outcome data and program monitoring data were available, including during the time of the biggest scale-up of the project (see Figure 10).

[2] Sources of bias in trend data are discussed in more detail in Chapter 2 and Chapter 6 of this guide.

Figure 10: Lack of uniformity between coverage data and bio-behavioral data



Putting the pieces together - Summary of the evidence

Table 1: Police in Cambodia: Did the program make a difference? What is the evidence?

Evidence For	Evidence Against	Inconclusive
• Consistent condom use climbed to nearly 100% after the program began, and the proportion of police buying sex declined only slightly	• Declines in HIV prevalence in police, sex workers and ANC from 1999 onward suggest that incidence declines began prior to the intervention program • This is consistent with available behavioral trend data which show that by 1997, reported consistent condom use was already above 60% in police	• There is a lack of uniformity between the various data sources • Outcome data were available from only a handful of sites where the program was active, and this was prior to the biggest scale-up of the project • Impact data (i.e., HIV prevalence trends) were available from all sites, but this was at a time when the intervention was active in only a subset of sites • From the available data, it is unclear whether possible program effects were stronger in some places than others, or whether there is evidence that more intense program exposure was associated with more behavior change or incidence declines

Conclusions

Province-specific surveillance data were not available for this case study and the use of national level data may not have captured the diversity of the situation with respect to the timing of the epidemic in different parts of the country, and the phasing of interventions with police and sex workers. All of these factors make it very difficult to establish evidence of program impact. However, it does seem fairly clear from available data that declines among the police and the FSWs began quite awhile before this police intervention started and before the biggest scale-up of the project.

What would make the evidence stronger?

- If the data could be analyzed by site, it may be possible to tease out more information about plausible program effects in a more quasi-experimental manner.
- Better and more accurate information on size of the police force in different locations and times would also help, although the fact that police are mobile and change locations periodically makes it more difficult to sort out effects at different sites.
- Data from the behavioral surveys on **Exposure** to the program would help corroborate information about the extent of program coverage.

Case Study # 2: Female sex workers in Nepal

Background information

The Safe Highway Project, also known as the ASHA project, is an HIV/STI prevention program. It started in nine Terai highway districts from 1993 to 1997, and expanded to 16 districts from1997 to 2002, and 22 districts from 2002 to 2007. During the last five years, the project covered the entire length of the highway and nearly all of the districts (22 of 26 districts). The project has been comprehensive, targeting both sex workers and client groups (i.e., truck drivers, laborers and policemen).

Components

The program included interpersonal communication, outreach and peer education, condom social marketing and distribution through retail and non-traditional outlets, awareness-raising in the community, and mass media. Drop-in centers, STI, and voluntary counseling and testing (VCT) services were also part of the program. All of these activities were phased in over time and the project grew from one implementing NGO in 1993 to 44 NGOs by 2006.

Examining the evidence

Data available for this case study

- **Outcome**
  - » Three rounds of IBBS data among sex workers in 1999, 2003 and 2006
  - » BSS data from 1998-2002

- **Size:**
  - » Limited data from the Government of Nepal

- **Output:**
  - » # of sex workers reached annually by peers in 2000 to 2004
  - » # of FSWs examined annually for STIs from 2000 to 2008

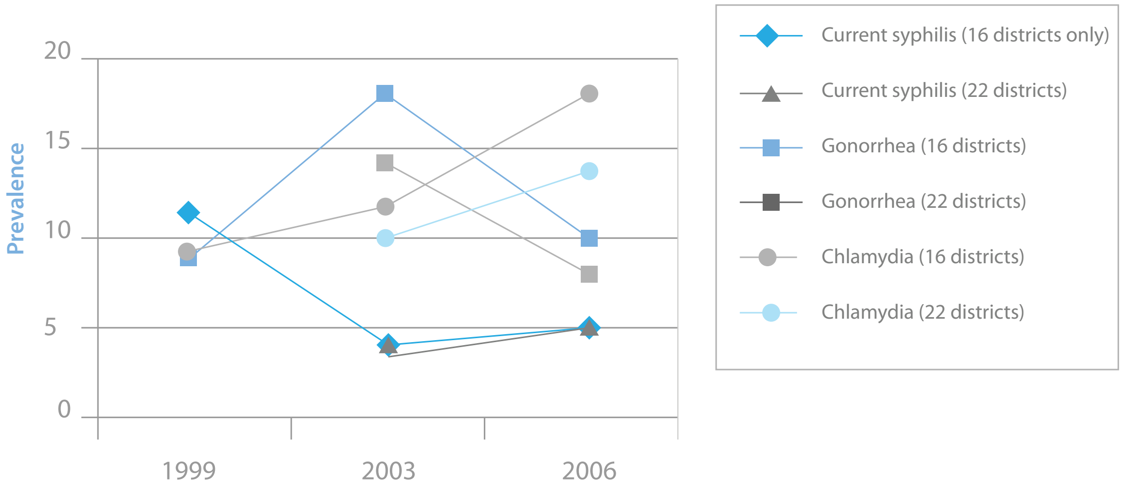
- **Impact:**
  - » IBBS data for HIV prevalence in 1999, 2003 and 2006

What is the evidence that outcomes changed?

STIs

The data on STI trends among sex workers on the highway (shown in Figure 11) are consistent with the hypothesis of declining incidence. Both syphilis and gonorrhea prevalence declined, which is indicative of reduced risk behavior. However, the increase in chlamydia is difficult to explain since it is not consistent with the rest of the data. One possible explanation is that while gonorrhea is short-lived and can cure spontaneously, chlamydia is more difficult to interpret because of a substantial proportion of long-lasting asymptomatic carriers.

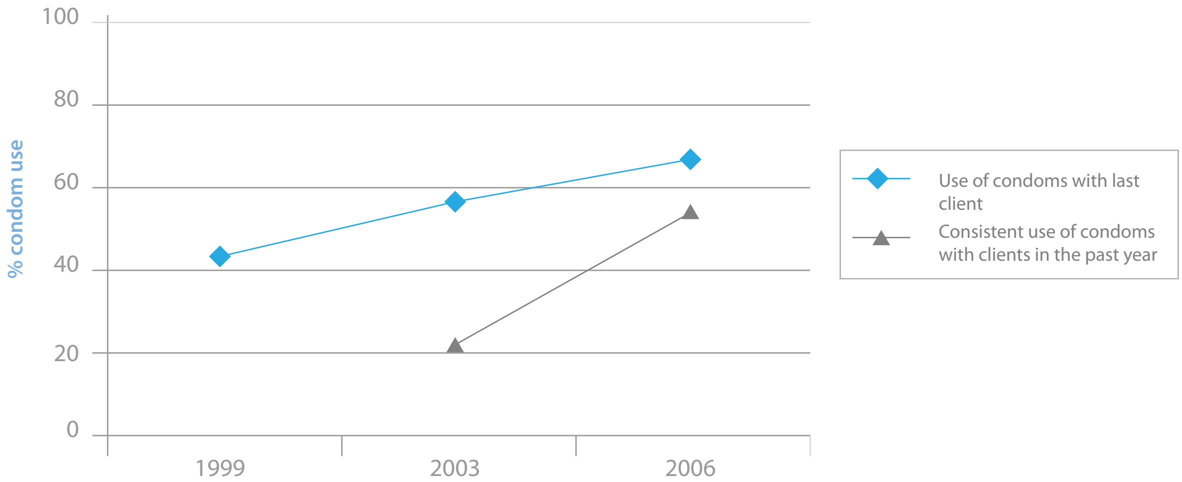
Figure 11: STIs trends among FSWs along the highway in the Terai between 1999 and 2006, New Era, STI/AIDS Counseling and Training Service (SACTs), FHI



Sexual Risk Behavior

Both reported last time and consistent condom use with clients increased significantly between 1999 and 2006 according to the IBBS data. However, consistent condom use was still below 60% and last time condom use below 70%, indicating that 30-40% of commercial sex encounters may be unprotected by condoms (see Figure 12).

Figure 12: Condom use among FSWs and their clients from 1999 to 2006 (IBBS), New Era, SACTs, FHI



What is the evidence that programs caused the changes in outcomes?

This question is explored by looking at program coverage based on size estimates, program monitoring data, and exposure information from behavioral surveys.

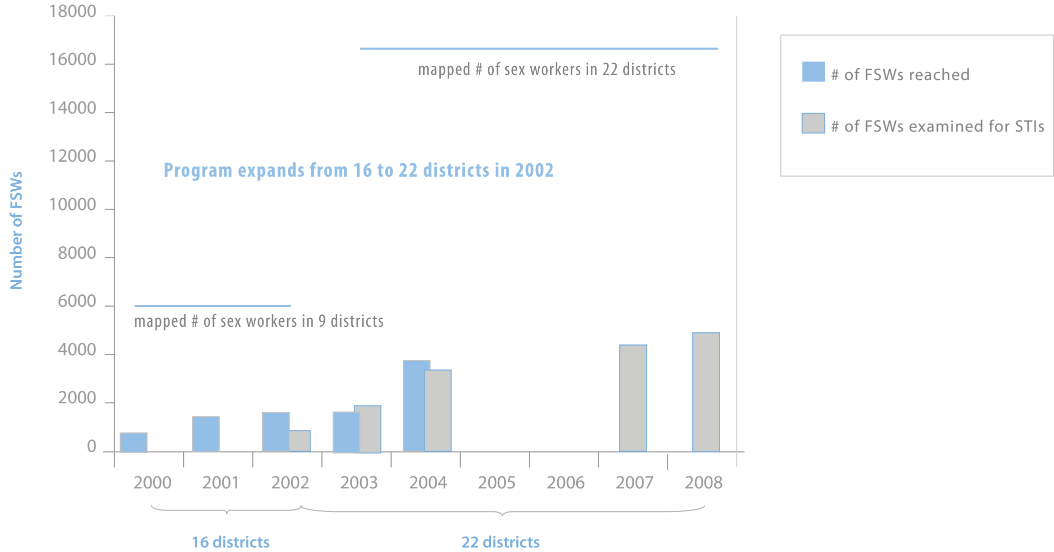
Size

Initially the Safe Highway project covered nine Terai districts, in which there was a working size estimate of 6,000 FSWs in those districts. The program expanded from nine to 16 districts to a total of 22 districts of the Terai between 1993 and 2007. In 2005, the Government of Nepal estimated that there were 14.000 to 18,000 FSWs in the Terai highway districts. In 2007, the estimate was 13,000 FSWs in the 22 districts covered by the Safe Highway project. Earlier estimates had put the number as high as 24,000. For the purposes of this case study the number is set at 6,000 for the period before 2002 when the program was covering fewer districts, and at 16,000 after 2002 as a reasonable compromise between the high and low estimates for all districts. Population turnover must be factored in when interpreting the data.

Coverage

Figure 13 gives a synopsis of program coverage against the estimated number of FSWs. From the available data we see that the number of FSWs being reached by the program expanded over time, as did the number referred and treated for STIs. But the numbers are nowhere near the total number of sex workers. These data reinforce the STI and sexual risk behavior trends, i.e. they underscore the need for continued scale-up of coverage, better control of STIs, and further reduction of exposure to HIV by increasing condom use.

Figure 13: Coverage data for Safe Highway Project with FSWs (program data from 2000-2008)



Exposure

Exposure data from behavioral surveys reinforce the information on reach and coverage of intervention programs. Both the program data (Figure 13) and the surveillance data (Figure 14) suggest that the program had wide and increasing coverage, but that there was still a significant proportion of women not being reached, and a large proportion of not accessing STI treatment through the program. The discrepancy between the coverage based on size of the population versus the exposure data from the IBBS points to the possibility that the IBBS survey may have been biased toward including people who had been exposed to the program.

Figure 14: FSWs in Nepal reported exposure to Safe Highway program (BSS and IBBS data)

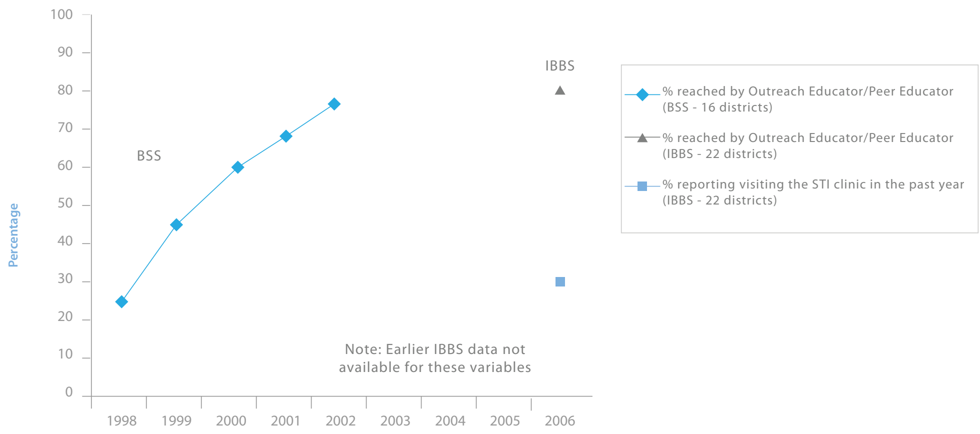
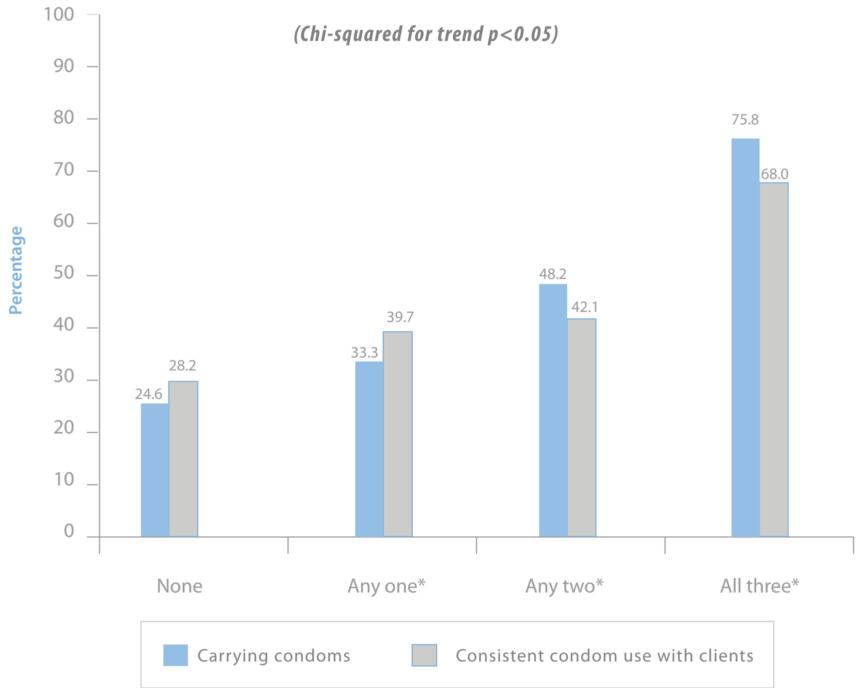


Figure 15 illustrates the existence of a "dose-response" relationship between the number of interventions accessed and outcomes of interest. The proportion of FSWs who reported both carrying condoms and using condoms consistently with their clients increased as a function of the number of interventions to which they were exposed in the previous year. These programmatic activities were receiving condoms, receiving Information, Education, and Communication (IEC) materials such as brochures and pamphlets, and Interpersonal Communication (IPC) information about HIV.

Figure 15: FSWs in Nepal: dose-response effect of exposure on condom use, Terai Highway districts, 2002



*1) Received condoms, 2) received brocures/materials, 3) received IPC information about HIV/AIDS

Table 2 is another analysis that shows the dose-response relationships between the frequency of FSW exposure to HIV prevention interventions and consistent condom use with clients during the last 12 months. Statistically significant dose-response relationships were identified for peer education and outreach work, drop-in center (DIC) visits, and the use of STI services, and participation in community awareness events.

Table 2: Relationship between the frequency of exposure of FSWs to HIV prevention interventions and consistent condom use, Terai Highway Districts, 2006

	Consistent condom use with clients in the last 12 months					
	Yes		No			P-value
	n	%	n	%	Odds Ratio	(χ^2 test for trend)
No. of contacts with peer educators and/or outreach workers during the last 12 months						P<0.001
None	30	24.0	95	76.0	1.00	
Once	8	22.2	28	77.8	0.90	
2-3 times	71	31.3	82	33.5	2.74	
4-6 times	49	44.1	62	55.9	2.48	
7-12 times	46	52.9	41	47.1	3.52	
>12 times	53	62.3	32	37.7	5.19	
No. of drop-in center visits during the last 12 months						P<0.001
None	148	39.9	223	60.1	1.00	
Once	10	43.5	13	56.5	1.16	
2-3 times	43	46.7	49	53.3	1.32	
4-6 times	27	48.2	29	51.8	1.40	
≥7 times	30	51.7	28	48.3	1.61	
No. of visits to STI clinic during the last 12 months						P<0.001
None	160	38.8	252	61.2	1.00	
Once	45	43.3	59	56.7	1.20	
2-3 times	43	60.6	28	39.4	2.42	
≥4 times	10	76.9	3	23.1	5.25	
No. of visits to VCT clinic during the last 12 months						P=0.47
None	253	65.7	132	34.3	1.00	
Once	70	53.4	61	46.6	0.60	
2-3 times	46	65.7	24	34.3	1.00	
≥ 4 times	9	69.2	4	20.8	1.17	
No. of times participated in HIV/AIDS community awareness events						P<0.001
None	126	31.6	273	68.4	1.00	
Once	31	53.4	25	44.6	2.69	
2-3 times	67	69.1	30	30.9	4.84	
≥ 4 times	32	69.6	14	30.4	4.95	

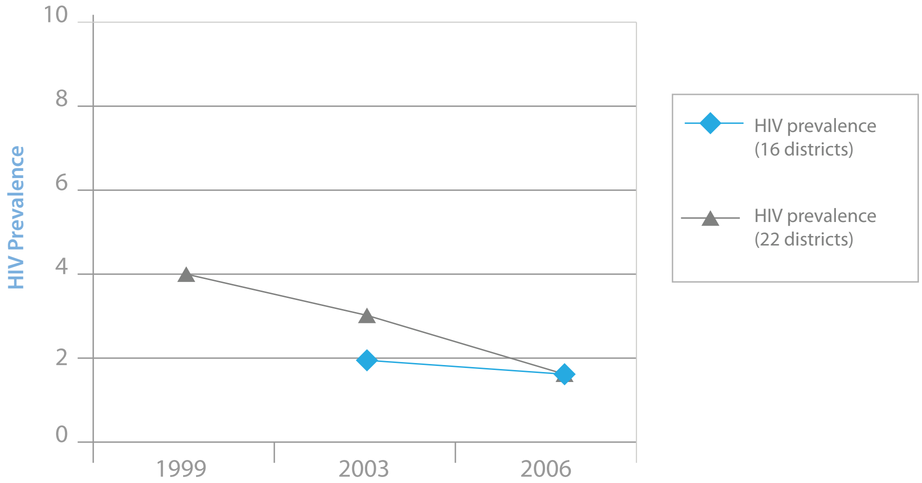
Source: D Prybylski et al, 2009, manuscript in preparation.

What is the evidence that changes in outcomes contributed to declines in incidence?

HIV surveillance trends

In Nepal, bio-behavioral surveys have been conducted among FSWs three times since 1999. The first round was conducted only in 16 districts but the second and third rounds were conducted in 22 districts (the original 16 plus 6 additional districts). It is evident from Figure 16 that HIV prevalence has not increased further among sex workers, and in fact it appears to have declined steadily between 1999 and 2006. This gives a good indication that HIV incidence rates have been held in check or declined during the same period (assuming that the declines in prevalence are not due to increasing mortality or out-migration), although further analysis of the data would be needed to rule out various sources of bias or explore factors that could account for the trend.

Figure 16: HIV prevalence trends among female sex workers along the highway in the Terai, Nepal between 1999-2006, New Era, SACTs, FHI



Are the data sources Uniform?

The IBBS surveys were conducted in 1999, 2003, and 2006. The first of these was conducted in the 16 districts corresponding to the 1997-2002 project districts, and the next two rounds corresponded to the expanded 22 districts of the project. So there was good correspondence between the surveys and the territory covered by the project.

Putting the pieces together: Summary of the evidence

Table 3: FSWs in Nepal: Did the program make a difference? What is the evidence?

Evidence For	Evidence Against
• Prevalence of HIV, syphilis and gonorrhea among sex workers declined between 1999 and 2006 • Condom use increased significantly • Dose response analyses indicated that there were differences in condom use as a function of intensity (# of times) of being exposed to the program	• Prevalence of chlamydia increased • Coverage and exposure data indicate that many women are still not being treated for STIs, which is consistent with the findings that condom use has not reached high enough levels, and some STIs are increasing • Condom use during commercial sex has increased significantly but there is still considerable unprotected high risk sex

What would make the evidence stronger?

This case study provides strong evidence of the plausibility of the Safe Highway project contribution to changes in outcomes that limited the spread of HIV among those who are reached by the program. At the same time, it uncovers some weaknesses for the program to address, including the need for increasing their coverage and looking more closely at the effect of STI treatment.

Additional evidence that would make the case study stronger would be an analysis of turnover among the sex worker population and an assessment of the extent of selection bias in the survey sampling. Corroborating evidence from the clients of sex workers would also be helpful.

Case Study # 3: Female sex workers in Bangladesh

Background information

This case study describes a program targeting street-based female sex workers (SBFSWs) in the city of Chittagong in Bangladesh. The program was implemented by an NGO known as Young Power in Social Action (YPSA). This program began targeting SBFSW in Chittagong in 2003. At that time an HIV epidemic had not yet started in this population, so the main outcomes of interest were the reduction of risk behavior and STIs.

Examining the evidence

Data available for this case study

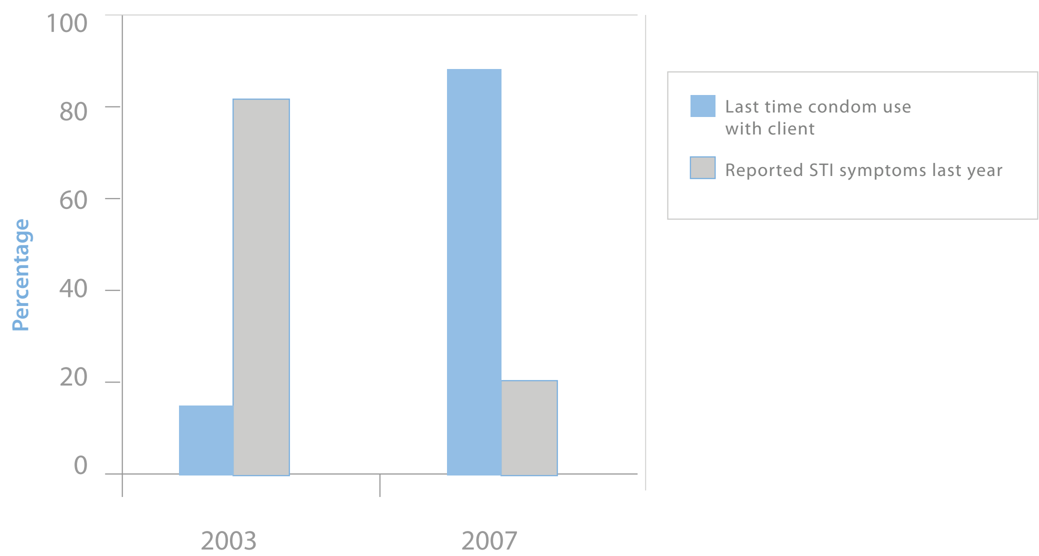
- **Outcome:**
  - » Two rounds of BSS in 2003 and 2007 (condom use, STI symptoms, exposure to program)
  - » Annual sentinel surveillance of syphilis, 2003-2007
- **Size:**
  - » Program mapping data in 2002
- **Output:**
  - » Annual condom distribution from 2003 to 2007
  - » # of sex workers screened at the clinic 2003, 2004 and 2005
  - » # of sex workers treated at the clinic in 2003, 2004 and 2005

What is the evidence that outcomes changed?

Sexual risk behavior

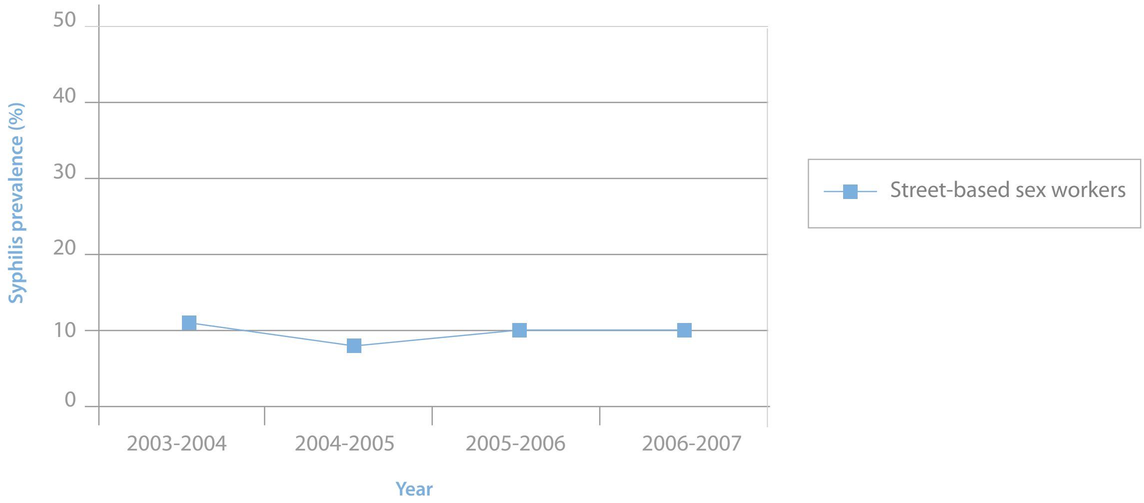
Data from the BSS in Figure 17 shows that the proportion of sex workers reporting condom use during last sex with a client rose from about 15% in 2003 to more than 90% in 2007. If these data are accurate, this is a very significant increase. The proportion of women reporting STI symptoms in the past year also declined significantly between 2003 and 2007, although at least 20% still reported STI symptoms in 2007.

Figure 17: Reported condom use and STI symptoms among sex workers in Bangladesh



The data from the national sero-surveillance system in Figure 17 indicate that active syphilis levels were relatively high and did not decline as expected, given the high reported condom use. In fact the syphilis rates in Chittagong were higher than in other surveillance sites in the country. The reasons for this are not clear, however, one possible explanation is that STIs were largely over-reported in 2003 when the program was still new; once women were informed and better able to recognize STIs, the reporting became more realistic. STI symptoms, particularly in women, are easily confused with other general non STI-related symptoms; therefore this type of over-reporting is not uncommon.

Figure 18: Active syphilis trends among street-based sex workers in Chittagong, 2005-2007 (National sero-surveillance)



What is the evidence that programs contributed to changes in outcomes?

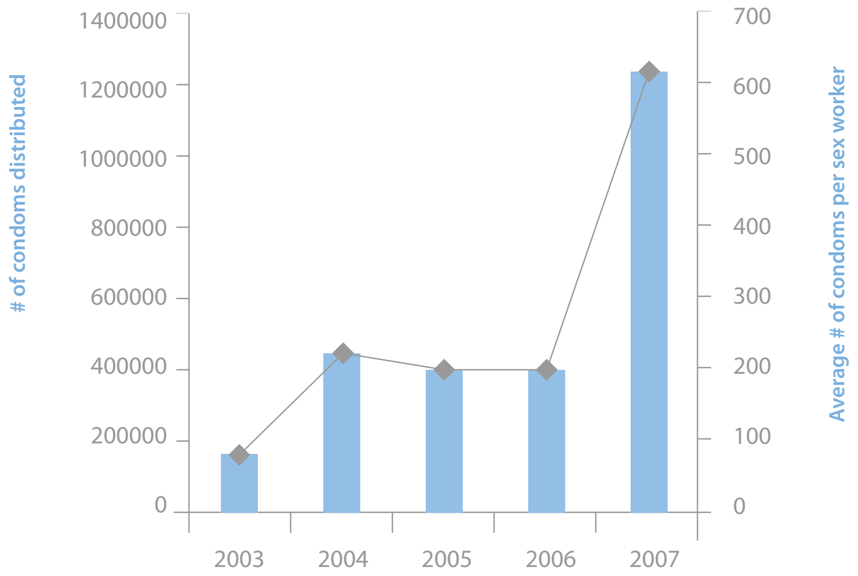
Size

The size of the target population of street-based sex workers in Chittagong estimated through program mapping in 2002 was approximately 2000.

Coverage

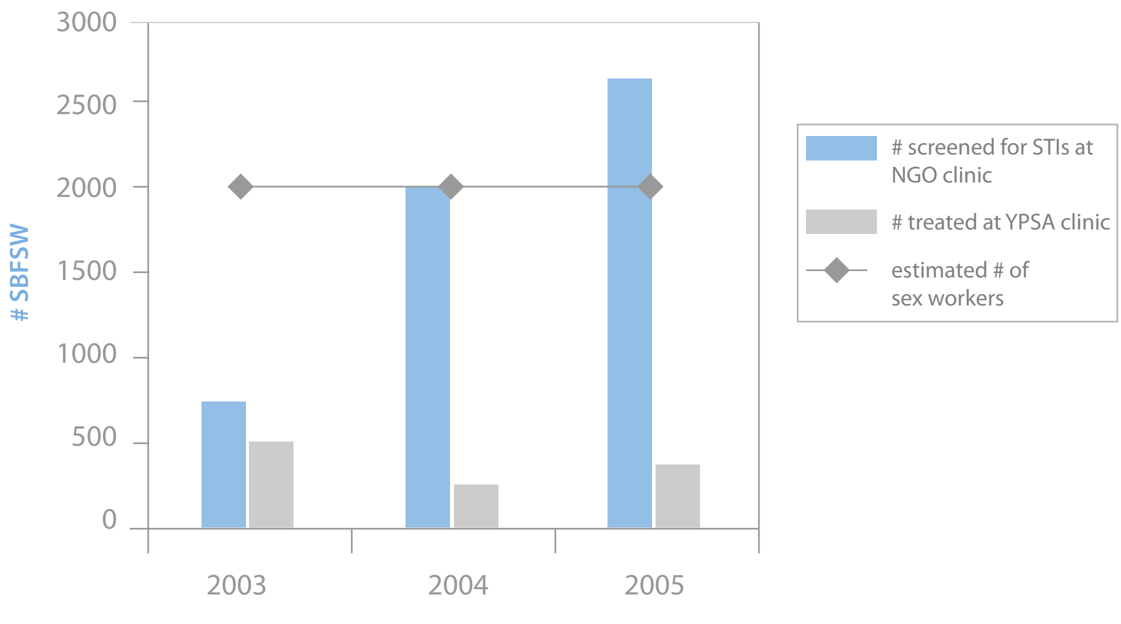
Program data generated by YPSA showed an increase in condom distribution between 2003 and 2007, with the biggest jump in 2007, when the number of condoms distributed more than doubled. Amongst approximately 2000 known sex workers, the average number of condoms distributed increased from an average of 90 per woman per year in 2003 to over 600 per woman per year between 2006-2007; this supply was enough to cover two sex acts per day 300 days a year for 2000 woman if that were the only source of condoms (see Figure 19).

Figure 19: Condom distribution among sex workers in Chittagong, Bangladesh (YPSA program data



YPSA also generated program data showing an increase in the number of individual sex workers screened at their clinics from 700 in 2003 to more than 2,500 in 2005 (which was 100% of the known sex workers) (Figure 20). As the number of women being screened increased, the proportion being treated declined, however, it is not known whether this is a sign that new STIs were occurring at a slower rate or whether the decline was merely a function of the broader group of sex workers being screened.

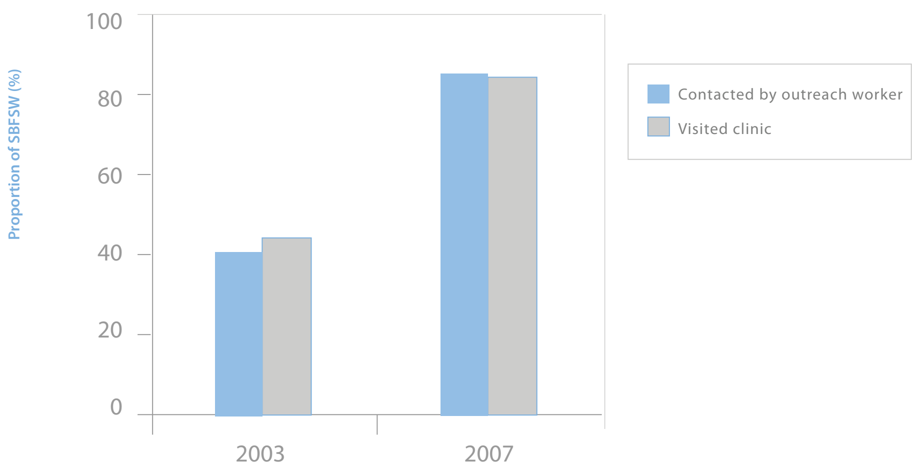
Figure 20: STI treatment of FSWs in Chittagong, Bangladesh (YPSA program data)



Exposure

Data from the BSS surveys conducted in 2003 and 2007 indicate that a substantial proportion of street-based sex workers reported being reached by the program. Both the proportion contacted by an outreach worker and the proportion visiting the clinic in the past year increased from less than 50% to over 80% of the population between 2003 and 2007 (see Figure 21). Although this suggests that a large proportion of the population was accessing the clinic and being screened for STIs, the data are not corroborated by the program coverage data, implying a much lower coverage level than the survey data, at least in 2003. This may indicate that the survey was biased toward the program. Unfortunately, the program data for 2006 and 2007 were not available so the 2007 exposure data could not be compared to the coverage trends.

Figure 21: FSWs in Chittagong reporting contact with YPSA outreach worker and visits to the clinic (BSS data)



Uniformity

The catchment area for the program included all cruising spots in the city of Chittagong. Spots were defined as sites with three or more sex workers during hours of operation. The BSS covered the same catchment area and defined SBFSW as women soliciting clients on the street, and having sex in public spaces or other venues. The YPSA program included all SBFSW spots, and the BSS sampled respondents from those spots as well. The BSS was conducted in 2003 and 2007 and there was considerable overlap between the coverage area of the program and the coverage area of the survey. It should be noted that the program did not track numbers of new patients screened and treated after 2005, so the evidence that the rate of clinic visits was consistently high in 2006 and 2007 was not available.

Putting the pieces together - Summary of the evidence

Table 4: FSWs in Bangladesh: Did the program make a difference? What is the evidence?

Evidence For	Evidence Against	Inconclusive
• Reported condom use rose significantly to nearly 100% last time use with clients (BSS) • Nearly all sex workers reported being contacted by an outreach worker and visiting the clinic (BSS) • According to program data, the majority of sex workers visited the clinic and underwent STI screening in 2005	• The prevalence of active syphilis after 2005 is higher than would be expected if condom use was high as was reported • Reported drops in STI symptoms do not corroborate the active syphilis trend data • Although condom distribution increased, the biggest jump was in 2007, which was after the time period when the majority of the change was supposed to have happened. Prior to 2007, condom distribution was relatively low	• There is insufficient program data from 2007 to correspond to the BSS data so it is unclear what might have happened with regard to screening and treatment of STIs

What would make the evidence stronger?

- Clinic data from 2006 and 2007 are needed to confirm levels of STI screening and treatment.
- Given the "mismatch" between reported condom use and levels of active syphilis, other data sources are required to validate the information.
- BSS data and program data suggest that this program should be having a positive impact, but those data are not reinforced by the one biological marker, which calls into question all the other data. Several components need to be checked, including the uniformity of the populations tested for syphilis with the clinic population, as well as with the BSS and the program catchment area. If all appear to match, then the validity of the self-reported data must be questioned, as well as the STI management guidelines, and treatment records kept by the program (which might be double-counting). The potential for bias toward the program in the BSS survey should also be investigated.

Triangulation exercise: injecting drug users in a hypothetical Asian country

The case studies presented in this chapter have provided examples of the process of triangulating various types of data to help with program evaluation using surveillance data for outcome and impact monitoring.

The exercise below is designed to give the reader a chance to practice using multiple data sources, including surveillance data, to assess the evidence for program contribution to HIV prevention efforts. This exercise pertains to a fictitious country in Asia and several sources of fictitious data are provided.

Suggested questions to address in the analysis include the following:

- What is the evidence that the program has succeeded? Consider both evidence for and evidence against.
- What sources of bias were you concerned about in this analysis?
- What are the programmatic implications of this analysis?
- What additional information would you want to have to make your evidence stronger?

Background information

Rohat is a large city in a fictitious country in Asia where HIV has been spreading among IDUs since the early 2000s. In 2004, the country conducted a large scale Rapid Situation Assessment (RSA) in major cities and found that there were seven cities with more than 1000 IDUs. A large NGO started implementing harm reduction activities in the city of Rohat in 2006. As part of their project preparation, the NGO mapped IDUs and estimated that there were about 5000 in the city, although some were intermittent injectors, who preferred smoking heroin but would inject when heroin supplies were low. By comparison, the earlier RSA in 2004 estimated that there were 3350 IDUs in Rohat. The program had resources to cover only half of the IDU spots in the city. The program targets were set to reach 1000 IDUs in the first year, 1500 in the second year, and 2500 in the third year of the project. The program components included a comprehensive mix of prevention services including needle distribution, peer education, and five drop-in centers in different parts of the city where abscess management and HIV counseling and testing services were available.

Data Source 1: A mapping exercise of IDUs in 2004

Data from rapid situation assessments suggest that the number of IDUs in the country has been increasing at the rate of 10% per year since the year 2004. A mapping exercise conducted in 2004 in major cities estimated that there were 3350 IDUs in Rohat.

Data Source 2: NGO program data from Rohat

A large NGO in Rohat started implementing HIV prevention activities in 2006. As part of their project preparation, they map IDUs in the city and estimate that there about 5000 IDUs in the city, although some IDU are intermittent injectors who prefer smoking heroin, but will turn to injecting when supplies are low. The program has resources to cover only half of the IDU spots in the city, and program targets are set to reach 1000 IDUs in the first year, 1500 in the second year, and 2500 in the third year. Project services include needle distribution, peer education, and five drop-in centers in different parts of the city where abscess management and HIV counseling and testing services are available.

Monthly # of new IDU contacted by peers from 2006-2008 in Rohat							
Year	Month	South	East	North	West	Central	All centers
2006	Jun	6	56	29	27	7	
2006	Jul	9	86	22	18	12	
2006	Aug	7	37	16	26	10	
2006	Sept	2	57	18	38	6	
2006	Oct	9	37	23	18	9	
2006	Nov	3	26	20	19	2	
2006	Dec	5	83	18	15	6	
Total 06		41	382	146	161	52	782
2007	Jan	8	70	15	34	9	
2007	Feb	11	86	22	18	12	
2007	Mar	9	46	20	33	13	
2007	Apr	3	57	18	38	6	
2007	May	11	46	29	23	11	
2007	Jun	4	26	20	19	2	
2007	Jul	6	66	23	19	8	
2007	Aug	9	70	15	34	9	
2007	Sep	14	69	22	18	12	
2007	Oct	11	23	25	41	16	
2007	Nov	3	57	18	38	6	
2007	Dec	14	46	36	28	14	
Total 07		103	663	261	341	116	1,485
2008	Jan	11	105	22	51	13	
2008	Feb	17	129	33	27	18	
2008	Mar	13	69	30	49	19	
2008	Apr	4	86	27	57	9	
2008	May	17	69	43	34	17	
2008	Jun	6	39	30	29	3	
2008	Jul	9	100	34	28	11	
2008	Aug	14	105	22	51	13	
2008	Sep	21	103	33	27	18	
Total 08		112	805	273	351	121	1,663

Monthly number of needles distributed in Rohat							
Year	Month	South	East	North	West	Central	All centers
2006	Jun	72	896	348	324	84	
2006	Jul	108	1,376	220	216	144	
2006	Aug	84	592	160	312	120	
2006	Sept	24	912	180	456	72	
2006	Oct	108	592	230	216	108	
2006	Nov	36	416	200	228	24	
2006	Dec	60	1,328	180	180	72	
Total 06		492	6,112	1,518	1,932	624	10,678
2007	Jan	120	1,120	232	540	140	
2007	Feb	180	1,376	352	288	192	
2007	Mar	140	740	320	520	200	
2007	Apr	40	912	288	608	96	
2007	May	180	740	460	360	180	
2007	Jun	60	416	320	304	32	
2007	Jul	100	1,062	360	300	120	
2007	Aug	150	1,120	232	540	140	
2007	Sep	225	1,101	352	288	192	
2007	Oct	175	370	400	650	250	
2007	Nov	50	912	288	608	96	
2007	Dec	225	740	575	450	225	
Total 07		1,645	10,609	4,179	5,456	1,863	23,752
2008	Jan	225	225	1,470	87	315	
2008	Feb	338	338	1,806	132	432	
2008	Mar	263	263	971	120	450	
2008	Apr	75	75	1,197	108	216	
2008	May	338	338	971	173	405	
2008	Jun	113	113	546	120	72	
2008	Jul	188	188	1,394	135	270	
2008	Aug	281	281	1,470	87	315	
2008	Sep	422	422	1,445	132	432	
Total 08		2,241	2,241	11,271	1,094	2,907	19,752

Data Source 3: Sentinel Surveillance Data

HIV prevalence data from an IDU sentinel site in Rohat showed HIV prevalence as follows:

2004: 7%
2006: 30%
2008: 32%

Data Source 4: Behavioral Surveillance Surveys among IDUs in Rohat

	2006	2008
% who shared a needle last time	65%	40%
% who have consistently not shared in past 6 months	20%	30%
% who received a needle from a peer in past year	20%	75%
% who have visited a DIC in the past year	5%	20%

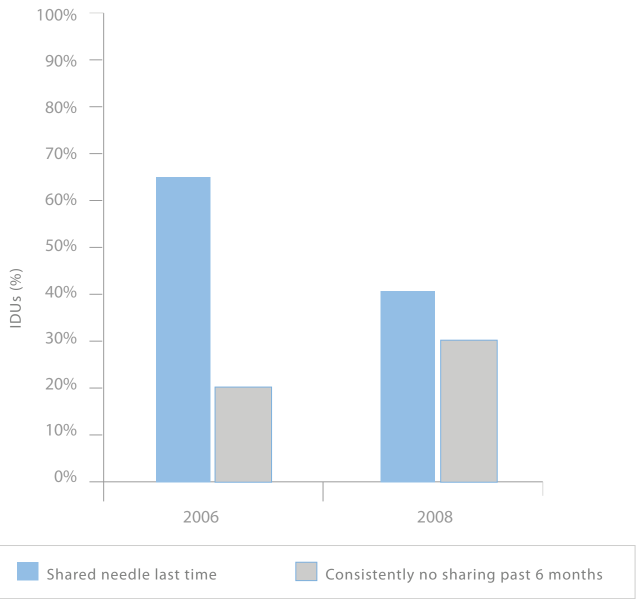
Suggested answers to exercise

What is the evidence that outcomes changed?

Sharing

The main outcome of interest for the program was reduced needle sharing. The data on needle sharing came from two rounds of IBBS for the city of Rohat in 2006 when the program first started, and again in 2008. Needle sharing was measured in two ways: 1) last time sharing; and, 2) consistently not sharing for the past six months. As seen in Figure 22, the proportion of IDUs in Rohat that reported sharing a needle last time they injected declined between 2006 and 2008, while the proportion that reported consistently not sharing increased. This was a positive outcome for the program, although sharing levels were still unacceptably high.

Figure 22: Needle sharing among IDUs in Rohat



Source: IBBS among IDUs in Rohat

What is the evidence that programs caused the changes in outcomes?

The next part of the analysis looked at the evidence that might attribute reduced levels of needle sharing to the program. To do this, both coverage data and exposure data were examined.

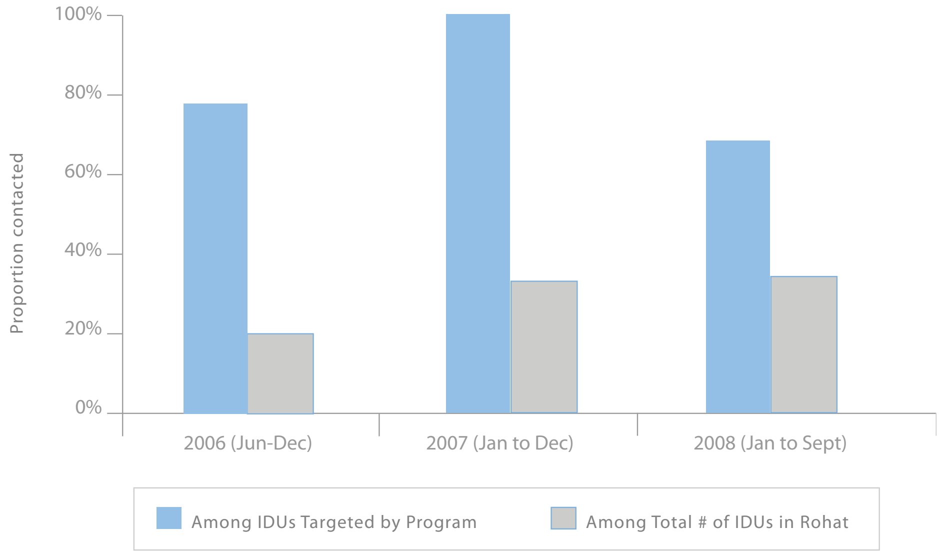
Coverage

Coverage was calculated on the basis of two program indicators: 1) proportion of IDUs contacted by an outreach worker in the past year; and, 2) mean number of needles distributed per IDU in the past year. The numerator for the first indicator was based on monthly data on number of new IDUS contacted by the program, and the numerator for the second indicator was based on the number of needles distributed to IDUs targeted by the program. Two different denominators were used to calculate the indicator. The first based on annual program targets, which were determined on the basis of the program mapping exercise in 2006, and the second based on available population size estimates for IDUs.

Size Estimates

The size estimates were based on RSA data for 2004, which was inflated by 10% annually, in keeping with police data showing that the number of arrests was increasing by 10% a year.

Figure 23: Proportion of IDUs contacted by a peer

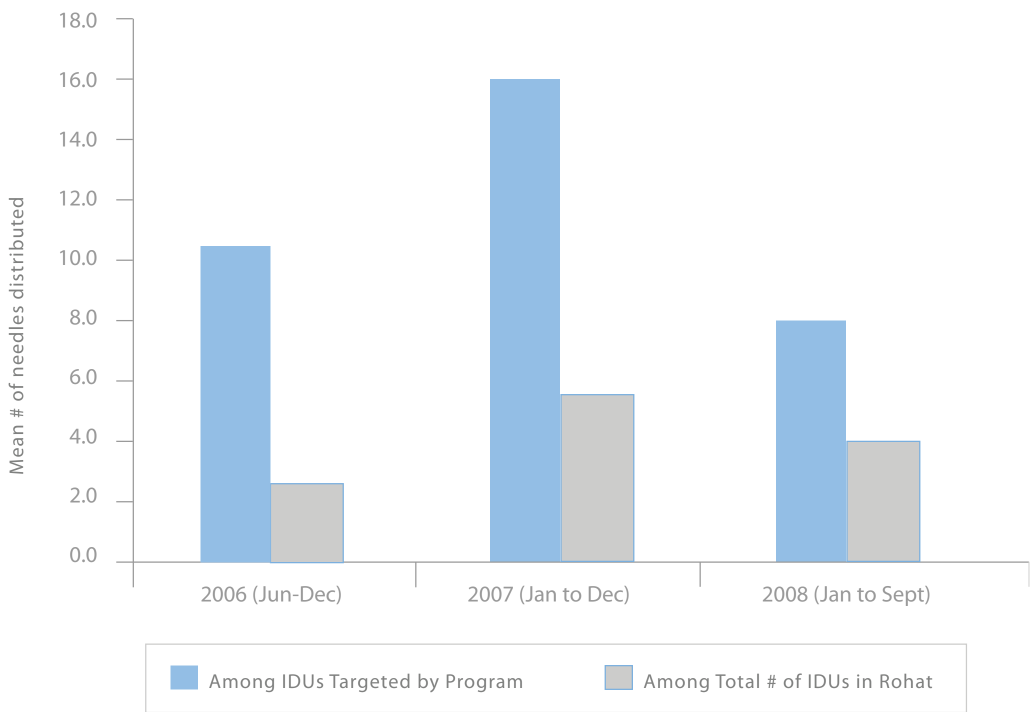


Source: Program counts and RSA data

Figure 23 shows that although the program was successful in making significant progress toward its targets in 2006 and 2007 (using proportion contacted by a peer as evidence), the overall coverage in the city was quite low (below 30%). Based on this analysis, the program may not have been positioned to make a definitive difference in the situation at-large. Coverage also appeared to have dropped in 2008, according to this indicator.

The mean number of needles distributed per year (Figure 24) was clearly inadequate for the total number if IDUs in the city, as well as for the number targeted by the program. The best case scenario was 16 needles per IDU targeted by the program per year in 2007, which dropped to 8 needles per year in 2008 (although the data in 2008 were available only for 9 months).

Figure 24: Mean number of needles distributed per IDU

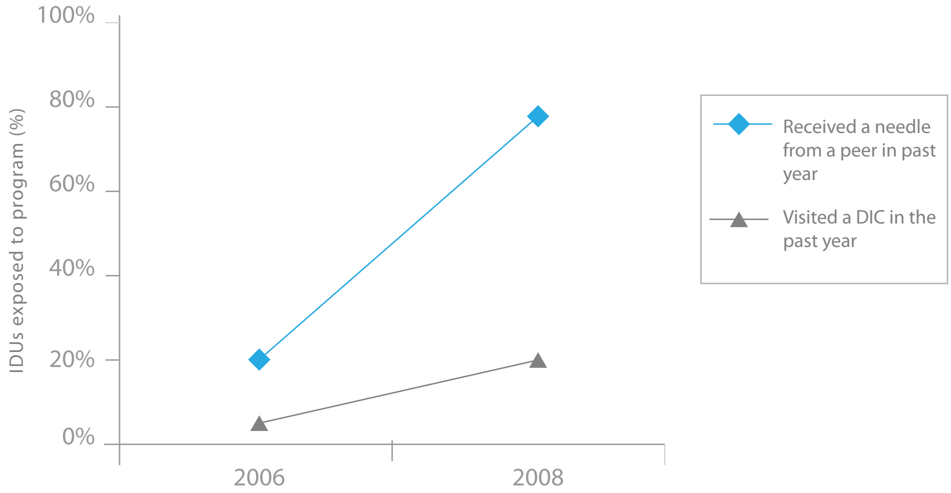


Source: Program counts and RSA data

Exposure

The information presented in Figure 25 on exposure (IBBS survey data) paints a somewhat different picture than the coverage data. The IBBS is ostensibly a probability survey and is therefore representative of the IDU population in Rohat. The exposure data in this survey implies that by 2008, nearly 80% of the IDUs had received needles from a peer sometime in the past year, and that fewer than 20% had visited the DIC. However, this does not seem to fit the expected scenario since the program was only targeting 50% of the total IDU population and their own program data seemed to indicate that needle distribution was limited.

Figure 25: Proportion of IDUs exposed to intervention program



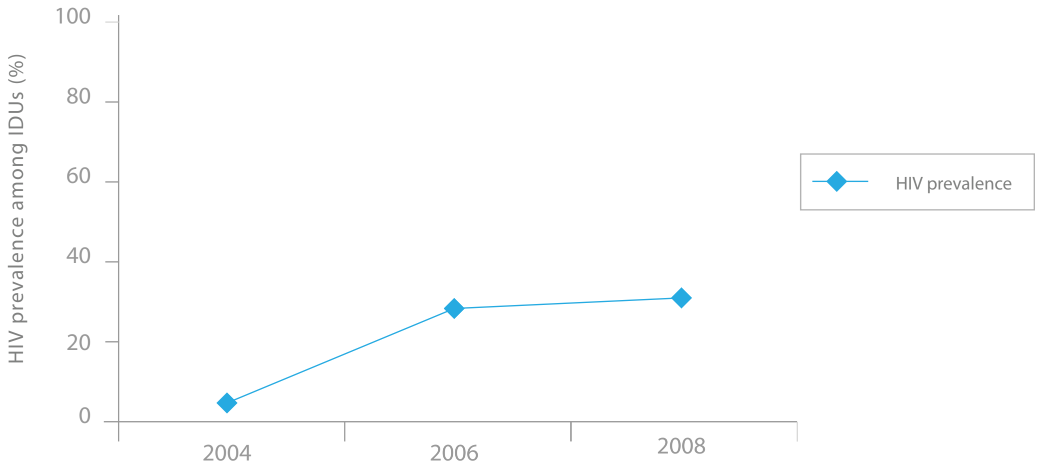
Source: IBBS among IDUs in Rohat

What is the evidence that changes in outcomes contributed to declines in incidence?

HIV surveillance trends

The prevalence of HIV in this population appears to have been on the rise prior to the program start-up in 2006 (see Figure 26). The fact that prevalence did not decline between 2006 and 2008 suggests that HIV incidence among IDUs continues to be high. The stable size of the IDU population reinforces this (Note: if incidence were falling, then over time the prevalence would be expected to fall, although this could take four years or more to be evident). More information about sampling, in- and out-migration, population size over time, treatment and mortality are needed to gain more insight into the trend.

Figure 26: HIV prevalence trends among IDUs



Source: HIV sentinel surveillance

Uniformity

There was a problem of uniformity between the sources of data to really understand what is going on in the program catchment area because the IBBS covered all of Rohat and the program coverage was limited to half of the IDUs in the city. Likewise the population size data was also for the entire city, making it somewhat difficult to measure coverage because the denominator was not specific to the program area.

Table 5: IDUs in Rohat: Did the program make a difference? What is the evidence?

Evidence For	Evidence Against	Inconclusive
• Progress toward program targets (proportion of contacted by a peer using program targets as denominator) • Proportion reporting having received a needle from a peer was nearly 80% • Reported sharing of needles declined during the first two years of the project	• Overall coverage of IDUs in the city is below 30% (using IDU estimates for entire city as denominator) • Mean # of needles distributed per IDU per year is inadequate, even if considering only the program targets • Consistent non-sharing still too low	• The prevalence of HIV has increased since 2004 and is now stable

Conclusions

The program appears to be having some success against its own targets (as evidenced by the coverage data and to some extent the IBSS data, at least in terms of proportion receiving needles). But this is not being translated into successful outcomes since needle sharing is still high. The program does not appear to be distributing enough needles to be making an impact on sharing. If there are any declines in incidence, it is too early to expect them to be reflected as declines in prevalence. In addition, the HIV prevalence trend data do not correspond to catchment area of the program specifically, so the plausible evidence for program contribution is unclear.

There are several possible sources of bias in the data. It seems that the IBBS data might be biased toward including people who are in the program catchment area, since 75% of the population reported receiving needles from a peer. However, this should not be the case because the program is only targeting about one third to one half of the IDUs in the city. The fact that this is similar to what the program targets indicate, 78%, 98% and 55% in 2006, 2007 an 2008 respectively (see Figure 23) is further indication that this might be the case.

We do not know much about the source of the HIV sentinel surveillance data, the extent to which it reflects what is going on in the city, or the program catchment area.

The size estimates of the IDUs may be one reason for the low overall coverage, although using multiple sources of size estimation did not seem to make much of a difference. The program seems to be performing more poorly in 2008, though this could be an "artifact" of inaccurate size estimates (i.e., maybe the assumption that the IDU population is growing at the rate of 10% per year is incorrect), and also the fact that the program data are for only nine months of the year instead of 12 months.

Programmatic implications of the analysis

- All the evidence combined points to the need for scaling up the program because it is not currently large enough and does not have good enough coverage to make the impact that is required.

What additional information could make the evidence stronger?

- For program managers, it would help to have biological and behavioral data correspond to the catchment area of the program (although this is likely not to be feasible, so the findings need to be interpreted with that limitation as a consideration).
- Size estimates more specific to the program coverage area would help the analysis. The program should consider a geographic mapping to obtain more in-depth information about its own targets.

Concluding Remarks

Using data to make a stronger case for and to give a better picture of program contributions to HIV prevention is essential for advocacy purposes to strengthen programs or to mobilize additional resources. Though data related to HIV programs are readily available on many levels in countries, more effort has been placed on collecting the data rather than on routinely analyzing and using data to improve programs and guide responses. The purpose of this guide is to provide the reader a grounded understanding of practical ways to look at data by first describing the types of data, methods for analyzing and understanding data, and ways to think through what the results mean. The case studies[3] in this guide highlight real program successes and also point to program – and data – challenges.

Though this guide is not intended to be a comprehensive manual data analysis, it is hoped that this guide encourages more proactive planning on how to better develop and to improve the monitoring and evaluation data that are collected and the systems for reporting them. More importantly, it is hoped that this guide demonstrates what is possible with existing program information and how influential good data can be when developing strategies for effective HIV prevention, care and treatment programs in any setting.

[3] Though the examples used in this guide address FSWs, IDUs and clients of sex workers, these are not intended to capture all at-risk populations in concentrated HIV epidemics. Because men who have sex with men (MSM) is a key at-risk population, we include here a reference to a document illustrating the use of triangulation data to evaluate MSM program effectiveness across three countries [16].

References

1. Pervilhac C, Stover J, Pisani E, et al. **Using HIV surveillance data: recent experiences and avenues for the future**. *AIDS* 2005,19:S53-S58.
2. Thomas Rehle, Tobi Saidel, Stephen Mills, Robert Magnani, editors. **Evaluating Programs for Prevention and Care in Developing Countries.** Arlington, VA: Family Health International; 2001.
3. UNAIDS. **A framework for monitoring and evaluating HIV prevention programmes for most-at-risk-populations**. In. Geneva, Switzerland: UNAIDS; 2007.
4. The Global Fund to Fight AIDS TB and Malaria. **Monitoring and Evaluation Toolkit**. In. Geneva: GFATM; 2009.
5. Global AIDS Monitoring and Evaluation Team (GAMET). **HIV Monitoring and Evaluation Resource Library**. In; 2009.
6. Last J. **A Dictionary of Epidemiology**. Third ed. New York: Oxford University Press; 1996.
7. UNAIDS. **Clearing the common ground for the "Three Ones": Report of a consultation process**. In. Geneva, Switzerland: UNAIDS; 2004.
8. Boerma JT, Weir S. **Integrating demographic and epidemiological approaches to research on HIV/AIDS**: The proximate determinants framework. *Journal of Infectious Disease* 2005,191:(Supple 1):S61-S67.
9. Brown T, Peerapatsnapokin W. **The Asian Epidemic Model: a process model for exploring HIV policy and programme alternatives in Asia.** *Sexually Transmitted Infections* 2004,80:i19.
10. Mills S, Saidel T, Magnani R, Brown T. **Surveillance and modelling of HIV,STI, and risk behaviours in concentrated HIV epidemics**. *Sexually Transmitted Infections* 2004,80:ii57-ii62.
11. Saidel T. **Potential impact of HIV among IDUs on heterosexual transmission in Asian settings: scenarios from the Asian Epidemic Model**. *International Journal of Drug Policy* 2003,14:63-74.
12. West GR, Cornel AL, Best K, Kurkjian KM, Cates W. **Focusing HIV prevention on those most likely to transmit the virus**. *AIDS Education & Prevention* 2007,19:275-288.
13. Thomas Rehle, Susan Hassig. **Conceptual Approach and Framework for Monitoring and Evaluation**. In: *Evaluating Programs for Prevention and Care in Developing Countries*. Edited by Thomas Rehle TS, Stephen Mills and Robert Magnani. Arlington; 2001.
14. UNAIDS/WHO Working Group on HIV/AIDS/STI Surveillance. **Estimating the size of populations at risk for HIV: issues and methods**. In: UNAIDS, WHO, FHI, UNODC; 2002.
15. Magnani R, Sabin K, Saidel T, Heckathorn D. **Review of sampling hard-to-reach and hidden populations for HIV surveillance**. *AIDS* 2005,19:S67-S72.
16. Ana Coghlan, Philippe Girault, Mike Merrigan, al. e. **Summary Report of Key Findings and Program Recommendations from FHI MSM Program Evaluations (Bangladesh, Indonesia and Nepal)**. In. Bangkok: Family Health International; United States Agency for International Development; 2006.

Data Triangulation

for HIV Prevention Program Evaluation in

Low and Concentrated Epidemics

Data Triangulation

for HIV Prevention Program Evaluation in

Low and Concentrated Epidemics