Training Manual on Methods for Size Estimation of Key At-Risk Populations in the Asia-Pacific Region

UNAIDS
Table of Contents

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Background</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Goals and Objectives</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Learning Objectives</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Intended audience</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Structure of the workshop</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Suggested data sources to support estimation of MARPs</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>Facilitator profiles</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>Agenda</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>Session 1 - Introduction to size estimation</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>Session 2 - Introduction to mapping-based size estimation methods</td>
<td>34</td>
</tr>
<tr>
<td>11</td>
<td>Session 3 - Introduction to survey-based methods with a focus on multipliers</td>
<td>52</td>
</tr>
<tr>
<td>12</td>
<td>Session 4 - Implementing the multiplier method - programmatic multipliers</td>
<td>64</td>
</tr>
<tr>
<td>13</td>
<td>Session 5 - Implementing the multiplier method – unique object multiplier</td>
<td>82</td>
</tr>
<tr>
<td>14</td>
<td>Session 6 - Implementing the Capture – Recapture Method</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>Session 7 - Network Scale-up – Review of Methods and Procedures</td>
<td>112</td>
</tr>
<tr>
<td>16</td>
<td>Session 8 - Network Scale-up – Doing it in practice</td>
<td>140</td>
</tr>
<tr>
<td>17</td>
<td>Session 9 - Introduction to using multiple sources for size estimates</td>
<td>162</td>
</tr>
<tr>
<td>18</td>
<td>Group Exercise 1 - Estimating the size of the female sex worker population in Guntur: Assessing Bias</td>
<td>170</td>
</tr>
<tr>
<td>19</td>
<td>Group Exercise 2 - Triangulation of multiple size estimate methods: How many MSW are in Madurai?</td>
<td>176</td>
</tr>
<tr>
<td>20</td>
<td>Flowchart -</td>
<td>188</td>
</tr>
<tr>
<td>21</td>
<td>Session 10 Presentation - Techniques for extrapolation to national level estimates</td>
<td>190</td>
</tr>
<tr>
<td>22</td>
<td>Country Exercise for National Estimates</td>
<td>219</td>
</tr>
</tbody>
</table>
Background

Reliable information about the number of people at high risk of HIV infection is increasingly recognized as one of the most critical pieces of information required by public health planners to understand and respond to the spread of HIV. A focus on most-at-risk populations (MARPs), including quantifying the size of these populations, is particularly relevant in the context of concentrated or low level HIV epidemics.

A key strategic information gap identified in countries in the Asia-Pacific region in 2008 was the lack of reliable size estimates of populations at risk of HIV infection. This data gap has limited the ability of national AIDS authorities to assess populations to prioritize for HIV prevention intervention, to make resource needs estimates for HIV prevention, and estimate the likely progression of the HIV epidemic. To assist countries and build their capacity to address this gap, the UNAIDS Regional Support Team (RST) for Asia and the Pacific, held a “Regional Training on Methods for Size Estimation of MARPs in the Asia Pacific Region” from 14-17 July 2009 in Bangkok, Thailand.

Goals and Objectives

This workshop curriculum is designed to develop the knowledge, skills and capacity of individuals and agencies responsible for helping produce both local and national estimates of size of MARPs, specifically sex workers, men who have sex with men, and injection drug users. Because the use of data influences the methods for data collection and analysis, it is important to note that the sessions in this training focus on three types of uses of size estimate data:

- Local level program planning, resource allocation, and monitoring of program coverage
- National or regional level prioritization of geographic areas on the basis of epidemic potential
- National or regional level estimates and projections of the burden of disease.

A main thrust of the curriculum is to help strategize on how to selectively gather new data and combine it with existing, to produce national level estimates.

Learning Objectives

Participants should leave the workshop with:

- Awareness of the methods available for collecting size estimation data and an understanding of which may be appropriate for their setting;
- Skills for analyzing and reconciling size estimation data which are already available;
- A working understanding of what, where, and among whom additional data need to be collected in their country;
- A step-by-step plan for coming up with national level size estimates for any of several high risk groups.
Intended Audience

This curriculum is intended for a wide range of participants, but is most useful for persons who are responsible for planning and implementing national size estimation exercises. Such individuals should have some familiarity with size estimation efforts that have occurred in their country to date and the related data which are available.

Structure of the Workshop

The agenda of the workshop follows a general approach for producing local or regional/national level size estimates. This process is laid out in the introductory session of the workshop and is outlined below, with references for what sessions address each topic.

Session 1: Introductory Session

- Articulate the purpose or uses of local and national size estimates
- Steps in the process of obtaining local and national size estimates
  - Getting started
    - Define population to be estimated
    - Define geographic areas requiring local size estimates
  - Assess each area for availability of local size estimates
    - Inventory data in each geographic area for size related data for the selected group to be estimated
    - In areas with data
      - Assess reliability and completeness
      - Reconcile multiple estimates
  - In areas without data, or where data are not reliable, prioritize areas for further data collection
  - Collect data using mapping-based or survey based methods
  - Assess reliability and completeness
  - Reconcile multiple estimates
  - In areas not selected for further data collection
    - Prepare for extrapolation

Session 2: Introduction to mapping-based size estimation methods

Session 3: Introduction to the multiplier method of survey based size estimation
Session 4: Implementing program-based multiplier method - Review of assumptions and types of bias encountered

Session 5: Implementing unique-object multiplier method
- Review of assumptions and types of bias encountered

Session 6: Implementing capture-recapture
- Review of assumptions types of bias encountered

Session 7: Using multiple sources for size estimation
  - Exercise 1: Assessing the bias with multiple size estimates – Followed by review of answers and key messages
  - Exercise 2: Reconciling multiple estimates - Followed by review of answers and key messages

Session 8: Extrapolation techniques for developing national size estimates

Pre-workshop Preparations by Participants

Participants may get more out of the curriculum if they have basic familiarity with analytical terms and concepts (such as precision, error, bias) and basic understanding of sampling for surveys, including types of probability sampling commonly used for high risk groups (e.g. time-location cluster sampling and respondent driven sampling).

Participants are given a list of the types of data that should be collated and brought to the workshop. The greater the effort to gather these data and become familiar with these sources ahead of time, the greater the ability to develop a workable plan for obtaining national size estimates during the workshop.
Suggested Data Sources to Support Size Estimation of MARPs:

- The groups usually considered for size estimation are any or all of the following:
  - Female Sex Workers (FSWs) by subtype (e.g. brothel, bar, street, hotel, residence, etc)
  - Men who have Sex with Men (MSM)
  - Male Sex Workers (MSW)
  - Transgenders at risk (Hijras)
  - Injecting Drug Users (IDU)
  - Clients of FSWs
  - External migrants

It will be helpful to think in advance about which are the relevant Most-at-Risk Populations (MARPs) in your country, and where concentrations of these populations exist. Be prepared to discuss how geographic locations with high concentrations of MARPs are identified in your country.

- Try to compile all available data sources with potential relevance for size estimation process:
  - For as many years as possible
  - For specific time periods if possible
  - For populations fulfilling specific definition criteria if possible
  - Organized by the smallest geographic (or administrative) area possible (e.g. district, province, etc).
  - Broken down by the smallest available sub-groups (e.g. among FSWs: entertainment based sex workers, street-based sex workers)

- Data sources may include (but not be limited to):
  - Size estimates from situation assessments and mapping exercises by all methods (GIS, social mapping, Rapid Situation Assessments (RSA), census, enumeration, nomination, etc)
    - These studies may have been conducted by the programmes or “external” research groups. Both are valid.
    - Be sure to track down whatever documentation is available on how the exercise was conducted, including definition of the population “counted”, whether the exercise was done for the whole geographic area or only a portion (e.g. biggest towns, urban areas, where interventions planned/existing, etc). This type of information can be critical for interpreting the data

- Monitoring data from programs conducting interventions among MARPs. The most helpful data will have the following characteristics:
  - Will pertain to individuals who meet clear definitional criteria
  - Will be unique for individuals (as opposed to contacts)
  - Be specific by time-period (e.g. available month-wise)
  - Will be specific to individual intervention programs
    - Examples of useful program data (best if specified month-wise)
- Numbers of individuals contacted by peers/ORW
- Number of individuals receiving condoms from peers/ORW
- Number of individuals receiving needles/syringes
- Number of individuals screened for STIs
- Number of individuals treated for STIs
- Number of individuals visiting the DIC
- Number of individuals treated for abscesses
- Number of individuals referred for counselling and testing services
- Number of individuals referred to STI centers
- Number of individuals registered with specific NGO programs
- Number of individuals enrolled in long or short-term treatment (e.g. short or long-term detoxification for IDUs) by treatment center

It can also be useful to note whether the programme uses unique identity numbers to track individual beneficiaries, or if the programme has a method for tracking drop outs, i.e. those people that have moved or died or no longer participate in the programme. If yes, note the definitions and protocols used to maintain this tracking system.

- Survey data with following characteristics:
  - Uses probability sampling methods (e.g. cluster sampling, time-location sampling, or respondent driven sampling)
  - Measures exposure to program indicators (e.g. as outlined above), and can be linked to specific programs (i.e. to be used in conjunction with program data of specific programs)
    - It will be helpful to bring copies of questionnaires, results, and information about how the survey as done.

- Drug related arrests (pertaining to specific time periods):
  - Number of individuals arrested for using drugs
  - Number of individuals arrested for selling or trafficking drugs

- Drug seizure data
- Country specific data on the proportion of drug users who are injectors
- Information from official government departments, such as Overseas Workers, Manpower, Emigration, etc., and from unofficial agencies, that pertains to numbers of individuals (segregated by gender) leaving the country for work in foreign countries, lengths of contracts, reasons for deportation, etc.

- General population size data by geographic division, segregated by age and gender

Also bring:

- Any information from specific size estimates exercises that have been done in your country using any method
- District level maps of the country (digitized if possible so that you will be able to show visual representations of population sizes by geographies
Facilitator Profiles

Tobi Saidel

Tobi Saidel has worked as an epidemiologist in Africa and Asia for the past two decades. Her early work with HIV incidence studies in the early 1990s in east Africa led to her involvement in HIV surveillance, and eventual specialization in the development and initiation of Second Generation Surveillance systems throughout Africa and Asia. She was an early promoter of behavioural surveillance as an important counterpart to biological surveillance, particularly among MARPS. This gave rise to her efforts to develop and improve sampling methods for tracking hidden populations in systematic and reliable ways.

More recently she has been working on improving capacity for using data to understand HIV epidemics with the goal of better informing the response, evaluating its impact, and influencing the surrounding policy dialogue. As part of this, she has been working on improving methods for estimating sizes of MARPS, which are important for virtually every aspect of planning and assessing HIV related responses, both locally and nationally.

Throughout her career, she has been involved in developing guidelines, manuals and training materials on all of these topics, and has conducted numerous workshops and one-on-one trainings in Asia, Africa, Latin America and most recently Eastern Europe and the Middle East and North Africa.

Virginia Loo

Virginia Loo is an epidemiologist specializing in HIV surveillance, monitoring & evaluation, and strategic planning.

She received her Ph.D. in Epidemiology from the School of Public Health, University of California - Berkeley. After completing her degree, she joined the U.S. Centers for Disease Control and Prevention in Atlanta, where she was an Epidemiologic Intelligence Service Officer in the Global AIDS Programme, Care and Treatment Branch.

From 2004-2006, she was based in India, managing the impact evaluation activities and routine monitoring system for Avahan India AIDS Initiative, a large scale targeted intervention programme in six high prevalence states in India, sponsored by the Bill & Melinda Gates Foundation.

She has been an independent consultant since 2007, working with multiple UN agencies and the World Bank, in addition to other development sector organizations and private foundations.

As a consultant, a majority of her work has been in training and technical support to National AIDS programmes on surveillance topics such as mapping and size estimation and conducting and analyzing probability surveys of MARPS.

She has also developed guidelines, toolkits, and training curricula on various surveillance, monitoring & evaluation, and strategic planning topics. For this work, Virginia has worked with colleagues in Afghanistan, Bolivia, India, Indonesia, Lesotho, Maldives, Middle East/North Africa region, Myanmar, Nepal, Pakistan, Paraguay, Sri Lanka, Uruguay, and Vietnam. She is currently based in Hawaii.
Donna Stroup
Agenda
Regional Training on Size Estimation of MARPS in Asia-Pacific
July 14-17, 2009, Bangkok

Training objectives:

- Understand the importance and uses of mapping and size estimation MARPS in the context of HIV epidemics
- Review classical and new methods and field procedures for mapping and size estimation
- Using multiple sources to develop size estimates
- Develop strategies for using data for national size estimation of MARPS

DAY 1

8:30-9:00 Registration and Coffee

9:00-9:30 Welcome Remarks

Workshop Opening Address:
Setting the Context for Size Estimations of MARPS in Asia-Pacific

9:30-9:45 Introductions of Trainers, Participants and Facilitators

Trainer: Tobi Saidel, PEMA

9:30-10:30 Session 1. Size Estimations
Background and Introduction to Size Estimation

Key Issues
- Purposes of size estimations
- Which populations need to be counted
- Overview of methods and applications of size estimates

10:30-11:00 COFFEE BREAK

11:00- Session 2. Classical Size Estimation Methods – Trainer: Virginia Loo
### Session 3. Introduction to Multiplier Methods

**Trainer:** Virginia Loo

**Key Issues**
- General concepts related to multipliers
- Survey and sampling issues that affect multipliers

### Session 4. Implementation of Multiplier Methods: Programmatic Multiplier

**Trainer:** Tobi Saidel

### Session 5. Implementation of Multiplier Methods – Unique Object

**Trainer:** Virginia Loo, PEMA

### Session 6. Capture-Recapture – a variation of the multiplier

**Trainer:** Tobi Saidel

### Summary of Multipliers

### Day 2

**Session 7. Network Scale-up: Review of Methods and Procedures**

**Trainer:** Rob Lyerla, UNAIDS Epidemic and Impact Monitoring Team, Geneva

**COFFEE BREAK**

**Session 8. Network Scale-up: Doing it in Practice**

**Trainer:** Donna F. Stroup, Data for Solutions, Inc.

**COFFEE BREAK**

**Session 9. Introduction to using multiple sources for size estimates**

**Trainer:** Tobi Saidel

**Group work – Exercises based on India IBBS data**

**COFFEE BREAK**

**Review key messages from the exercise**

**Trainers:** Tobi Saidel and Virginia Loo
### DAY 3

<table>
<thead>
<tr>
<th>Time</th>
<th>Session/Activity</th>
<th>Trainer/Facilitators</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00-10:00</td>
<td>Session 10. Developing strategies for using data for national size estimation of MARPS</td>
<td>Virginia Loo</td>
</tr>
<tr>
<td></td>
<td><strong>Key Issues</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Options for making national estimates using sub-national data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Review case studies: methods of using and interpreting existing data and planning</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Case studies: Bangladesh and Indonesia</td>
<td></td>
</tr>
<tr>
<td>10:00-10:30</td>
<td>Open Discussion – Questions and Answers</td>
<td>Tobi Saidel and Virginia Loo</td>
</tr>
<tr>
<td>10:30-11:00</td>
<td>COFFEE BREAK</td>
<td></td>
</tr>
<tr>
<td>Rest of Day</td>
<td>Participants will work with data from their own countries (which they have been asked to bring) to make provisional size estimates and to plan for implementing size estimation activities in their countries</td>
<td>Country Facilitators:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trainers: Tobi Saidel, Virginia Loo, Rob Lyerla, Donna Stroup</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WHO SEARO: Renu Garg;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNAIDS: Amala Reddy, Bob Verbruggen, NweNwe Aye, Nalyn Siripong;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>USAID RDMA: Patchara Rumakom;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FHI: Shanthi Noriega;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EWC: Wiwat Peerapatanapokin;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CDC: Evelyn Kim.</td>
</tr>
</tbody>
</table>

### DAY 4

<table>
<thead>
<tr>
<th>Time</th>
<th>Session/Activity</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 – 12:00</td>
<td>Country presentations of any preliminary size estimates produced and country plans for group review and suggestions</td>
<td>JVR Prasada Rao, Director UNAIDS RST AP, WHO SEARO</td>
</tr>
<tr>
<td>12:00 – 12:30</td>
<td>Closing Remarks and Ceremony</td>
<td></td>
</tr>
</tbody>
</table>
Session 1

Introduction to Size Estimation
Session 1 Summary: Introduction to Size Estimation

Objectives

- Review the multiple purposes and users of size estimates for MARPS
- Understand the importance of defining populations and geographic areas for size estimation
- Review the relationship between local and national size estimates, and the steps involved in producing each

Contents

This session introduces the contents of the workshop, and lays out an approach for obtaining national level estimates of MARPS. It covers some of the key design issues for designing local size estimation activities and emphasizes the relationship between local and national estimates.

Specifics of what is covered in the session/ tips for facilitators

This session describes the multiple purposes and users of size estimates, and lays out some key design features, including a brief description of methods for local size estimation, a discussion of how to define the major MARPS (sex workers, men who have sex with men, and injecting drug users) for size estimation purposes, a discussion of how to address timeframes and changing population sizes, and a description of how to define geographic areas for size estimation. An approach for obtaining national size estimates is proposed, which involves assessing each geographic area for the availability of local size estimates, prioritizing areas for further data collection, reconciling multiple estimates through a process of assessing biases, and preparing for extrapolation in areas with no data.

Facilitators should ask participants to think about whether they identify more with “A”, “B” or “C” type users of size estimates (which are described in the early part of the session), and encourage them to think about the perspectives of each of these different “users” when thinking through the various design issues for size estimation.
Session 1: Presentation

Facilitator notes:

Welcome to participants

To achieve increased understanding of the importance of size estimation of high risk groups. Explain the impetus of the workshop:

- Countries gearing up for new round of UNGASS reporting
- Some countries trying to finalize estimates and projections
- Some countries doing strategic planning
- Some countries applying for global funds
Facilitator notes:

Facilitator: Read objective

- Reliable information about the number of people at high risk of HIV infection is one of the most critical pieces of information required by public health officials/planners to understand and respond to the spread of HIV.
- Because not all people are equally exposed to HIV, it is critical to identify and quantify groups of people who are exposed to HIV more frequently or at an earlier time than others, and who are therefore more likely to acquire and transmit HIV.
- This involves people who have a lot of sexual partners with whom they have unprotected vaginal or anal sex, and people who engage in unsafe injecting. Such people can be difficult to identify, let alone categorize and count. Yet it is incumbent upon health planners to do so, in order to respond to the problem of HIV spread in a focused and targeted way.
Facilitator notes:

The agenda for this training follows a general approach for producing local, regional and national size estimates. The process is an iterative one that involves using existing data sources, prioritizing new data collection, reconciling the different sources, extrapolating to fill in gaps, and periodically updating existing estimates with newer or more complete information.

The training objectives over the next several days are:

- To familiarize participants with the process of developing local and national level size estimates for specific high risk groups in their countries
- To develop an awareness of specific data collection methods for size estimation and an understanding of which are appropriate for different circumstances
- To share best practices for analyzing and reconciling data that are already available in country
- To discuss step-by-step plans for producing national level size estimates for different high risk groups

The training objectives over the next several days are:

- To familiarize participants with the process of developing local and national level size estimates for specific high risk groups in their countries
- To develop an awareness of specific data collection methods for size estimation and an understanding of which are appropriate for the settings they work in
- To share best practices for analyzing and reconciling the data that are already available in country, which may have been collected for different purposes, but which can serve as a relevant sources of information for developing size estimates
- To identify what new data need to be collected in specific settings, to fill existing gaps or be used for extrapolation purposes
- To discuss with country team members, step-by-step plans, that can be used to produce national level size estimates for different high risk groups in their countries
Facilitator notes:

In this session, we will start by reviewing the multiple purposes of size estimation, and also talk about who uses size estimates, what their interests are, what they do with the data, and ways to resolve conflicting requirements.

Throughout this workshop, we will see repeatedly how local and national size estimates are interrelated, despite being used for different purposes. More often than not, national estimates are derived from local size estimates, but the process has many challenges, some of which we will have a chance to explore more closely by working with real size estimation data from some countries in the region.

- We will briefly introduce the different methods for local size estimation
- We will talk about which high risk populations are important for size estimation, and the different ways in which these populations may be defined in the context of HIV programmes.
- We will address the issue of timeframes and changing population sizes over time
- We will discuss the importance of defining geographic and prioritizing where to collect additional data
- Finally, we will introduce the concept of local level and national level size estimates.
Facilitator notes:

Size estimates are fundamental to understanding the spread of HIV, especially in concentrated or low level epidemics.

Programs depend on size estimates for several purposes:

- At the local level - size estimates are used for planning programmes and developing service delivery strategies. They are also the basis for measuring program coverage.
- At the national level- size estimates feed into estimates of current and future burden of infection and are also used for strategic planning and geographic prioritization of resources within the country.
- The next group of slides suggests t3 categories of “consumers” of size estimates, each with a different orientation. Recognizing these different perspectives is important when developing size estimates that can be understood, endorsed, and used by many types of stakeholders. During this workshop we will refer to these groupings as the “ABCs” of size estimation.
A is for Analysts

- These are the “epi-types” who make estimates and projections. They use models and spreadsheets and may seem to have their heads in the clouds.

Facilitator notes:

A is for analyst. Analysts are the people responsible for estimating the number of people infected with HIV.

The methods used for making these estimates rely heavily on size estimates of high-risk populations, especially in countries with low and concentrated epidemics, because the majority of HIV infections are found in members of these populations, and their lower risk sexual partners.

Countries need national estimates for:

- Contributing to global estimates of burden of disease
- Advocating for resources at the global level
- Increasing their understanding of the epidemic potential nationally, and in different parts of the country
Facilitator notes:

- Big picture people need to allocate resources and plan responses at the global, regional, national, or sub-national level.
- Whatever the level, they are interested in knowing sizes of high risk groups so that they can prioritize where resources are needed for prevention, and for treatment and care.
- When it comes to prevention, big picture people want to get the biggest bang for their buck. So they need to know which geographic areas have the biggest epidemic potential, and which populations are likely to contribute the most new infections. Both of these are determined in large part by the size of the risk populations.
- When it comes to treatment and care, it is also critical to know how many people are infected, so that adequate resources can be made available for treatment facilities, drugs, and psycho-social support. In low and concentrated epidemics, high risk populations constitute a large proportion of the people needing these services, so it is critical to know how many of them there are.
Facilitator notes:

- C is for community and it represents people who are affected by HIV
- And the organizations and agencies who provide services to them.
- Their immediate need for size estimates is to allow them to raise sufficient funds to hire sufficient staff and procure sufficient equipment supplies and facilities to meet the needs of the communities they serve.
- Since many programs for high-risk groups use outreach and peer education models size estimates are an integral part of planning activities and targets.
- Size estimates are also used to measure programme coverage. When programs provide services they keep track of the number of people they reach and measure that against their program targets.
- Sometimes local programs are unable to expand their services to meet local demand so additional programs may be needed. But without population size estimates it is difficult to quantify unmet need.

*Facilitator: Ask for a show of hands of people in the room who are As, Bs, and Cs.*
Facilitator notes: In this workshop, we will be learning about two main types of methods for size estimation, mapping-based methods, and survey-based methods. Mapping-based methods includes census (geographic or social mapping), and enumeration. We will be discussing mapping methods in-depth in the next session:

- These methods are generally used to estimate sizes of populations whose members are frequently present at certain types of locations or venues that are associated with high risk behaviour. Examples of such sites would include sex solicitation points.
- Brothel areas MSM cruising sites where drugs are bought and sold or destination points for migrant workers such as plantations factories mines construction sites etc.
- This type of mapping will “miss” high risk groups who do not congregate at these types of locations. However it will capture a lot of “location-based” risky behaviour (Survey-based methods will be discussed later in the session):
- Include multiplier methods which involve combining information from a probability survey data from other data sources that overlap with the survey in a known way.
- Capture recapture is a variation of a multiplier involving two or more surveys that “tag” and “retag” individuals, and then it uses the information about the numbers tagged and retagged in successive surveys to estimate population sizes. The unique object multiplier method and the recapture methods use the same mathematical properties.
Facilitator notes:

- The most common high risk populations that need to be estimated in low and concentrated epidemics are sex workers and their clients, men who have sex with men and men and women who use injecting drugs and share injecting equipment.
- There may be other populations with multiple concurrent sexual partners outside of the groups just mentioned who are also at high risk for HIV. But these populations tend to play a greater role in generalized epidemics and are not the highest priority in most Asian countries up to this point.
How should sex workers be defined for size estimation?

- Different “users” will have different needs
  - The “As” will most likely want a count of sex workers of all types, in all locations
  - The “Bs” will want to prioritize resources in areas with the greatest number of sex workers with the riskiest behaviours
  - The “Cs” may be more interested in the people they can reach with interventions
    - More visible subset, accessible at venues (brothels, street)
    - Covering only certain geographic areas that they can reach
- Other “definitional” issues
  - Male and female sex workers should be defined and estimated separately because of their different behavioural patterns
  - Sex workers should also be separated by typology (e.g. brothel-based, street-based, bar-based, residence-based)

Facilitator notes: Depending on which of these hats you wear, (A, B, or C), you may be looking for different types of data, require different levels of precision, or define populations differently:

- The “As” will most likely want a count of sex workers of all types, in all locations
- The “Bs” will want to prioritize resources to the areas with the greatest number of sex workers with the riskiest behaviours, i.e. have the highest number of clients in a week), so maybe limited to certain types of sex workers who are known to have more clients in a particular area (e.g. brothel sex workers). This type of prioritization would require some information on typology (see below)
- The “Cs” may be more interested in the people they can reach with interventions
  - More visible subset, accessible at venues (brothels, street)
  - Covering only certain geographic areas that they can reach or that are within their assigned catchment area

(Different methods and implementation of the protocol for size estimation may be appropriate depending on the way the population is defined)

Other “definitional” issues:

- Sex workers may be male or female, but these should be treated as separate populations for the purpose of size estimation. This is because their behavioural patterns, the profile of their partners, and ultimately the epidemic potential for them and their sexual partners is different
- Sex workers should also be separated by typology, because different types of sex workers may have different numbers of partners and different levels of risk. Sometimes it is necessary to conduct rapid assessments to collect this type of information.
Facilitator notes: How should MSM be defined? Men who have sex with men present a special challenge, because they can be defined in many different ways for many different purposes.

- In the human rights arena those working to reduce stigma and discrimination against MSM might favour counting all men who define themselves in terms of their sexual preference for men as part of a national size estimate.
- However when it comes to risk of HIV identity is not the main issue. Identifying as an MSM does not automatically mean increased HIV risk. However being a man with multiple male sexual partners does increase HIV risk regardless of how that man identifies himself. Unprotected anal sex also carries a higher risk of HIV transmission than some other sexual activities.

In the ABC framework:

- The “As” would likely be interested in all men who have sex with multiple male partners, regardless of how they identify themselves. And they would want to differentiate men who have a high volume of partners and/or who tend to be the receptive partner, from those who have fewer partners. They might consider male sex workers or transgenders as a proxy for this subtype of MSM. They would also want to know how many men have both male and female sex partners.
- The “Bs” would likely want to know where the highest numbers of MSM with the highest epidemic potential are located.
- And the “Cs” might have multiple interests, including both identity and behaviour, that would help them customize interventions and target effectively.
How should injecting drug users be defined for size estimation?

- As with the other groups, the definition will vary with the purpose
  - The “As” will be most interested in high frequency injectors who share needles
  - The “Bs” will be interested in areas with the highest number of IDUs with the riskiest behaviours
  - The “Cs” might want to count both non-injecting drug users, as well as IDUs, for program purposes

Facilitator notes:

How should IDUs be defined? As with the other high-risk groups, the definition of IDUs will vary depending on the purpose

- The “As” will most likely be interested in the highest frequency injectors who also share needles
- The “Bs” will be interested in epidemic potential in different parts of the country, so they will want to know the areas with the highest number of IDUs with the riskiest behaviours
- The “Cs” might be interested in reaching all drug users (DUs), not just injecting drug users (IDUs), so they might argue for estimating “non injecting DUs” as well as “IDUs”
  - This is related to programming interests, and the fact that “DUs” need programs sometimes as much or more than IDUs, and the types of programs they need may differ
  - Drug users who have not yet started to inject are an important program target group in their own right, because of the potential to prevent switching to inject, and thereby also reducing HIV epidemic potential
What timeframe should be used for defining high risk populations

- An issue of huge importance for size estimation is the dynamic nature of high risk populations
- Understanding epidemic dynamics means knowing sizes not only of people who currently engage in high-risk behaviours, but of those who did so in the past
  - The “As” will require data on both current and former sex workers, IDUs and MSM
  - The “Cs” will be more interested in current behaviours for programme targeting
- These differences in orientation have to be resolved when planning size estimation activities

Facilitator notes:

Also of huge importance for size estimation is the dynamic nature of high risk populations, and the fact that people “migrate” in and out of risk groups. For the purpose of understanding epidemic dynamics, and epidemic potential, it is important to know not only about people who currently sell sex, inject drugs, or have sex with many partners, but also those who did so in the past, and perhaps no longer do.

- Former sex workers, drug users, and high risk MSM who are infected with HIV, still play an important role in HIV epidemics, and also need to be counted as part of national estimates of numbers of infected people.
- This poses a problem of whether to define risk groups for size estimation in terms current risky behaviour, or “ever” risky behaviour.
  - The “As” will require data on both current and former sex workers, IDU, and MSM, to understand transmission dynamics and epidemic potential, and also to calculate treatment need
  - The “Cs” will mostly likely be most interested in current behaviours for programme targeting, however, some programmes consider people who are already infected with HIV as a primary targets for both prevention and treatment, even when those individuals are no longer involved in high risk behaviour
- The simplest way to approach this problem is to estimate the sizes of people who are “currently” involved in risk behaviours, but also gather qualitative information on turnover (i.e. average duration of involvement in sex work, drug use, etc.)
How to define “current” high risk

• A debate that frequently surfaces during the design phase of size estimation activities is how to define “current” risk
  – Exchanging sex for money at least once in the past one month?
  – Injecting drugs at least once in the past 6 months?
  – Men who had sex with other men ever?

• There is no one “right” answer to this question, but it should be discussed, agreed upon and carefully documented while planning size estimation activities.

Facilitator notes:

- Another important aspect of the “timeframe” debate is how to define “current” involvement in risk behaviour. Should eligibility criteria for size estimation activities include only people who are currently involved in risk behaviours? Should it involve people who have been involved in risk behaviour any time in the past week, month or year? Should the definition be based on frequency of the behaviour? Keeping in mind that those who engage more frequently are at higher risk?
- There is no one right answer to this question, but it is something that should be discussed, agreed upon, and carefully documented, so that those who are using the information can interpret it properly.
Facilitator notes:

- For the purpose of producing national size estimates, and undertaking geographic prioritization within countries, it is important to have size estimates for all geographic areas within a country.
- Given that it is not possible to directly measure the sizes of all risk groups in all locations, one of the challenges countries face is deciding where to invest time, effort and resources on size estimation activities.
- One approach to prioritizing areas for direct size estimation activities involves collating available data to determine which areas warrant more in-depth assessment.
  - This approach requires dividing the country into smaller geographic units, (possibly according to administrative divisions), and categorizing these units in terms of HIV transmission dynamics.
  - Obviously some “units” will have more information than others, and some units will have little or no information on which to base such a categorization. However, those units can be grouped together with other units that share characteristics that might be associated with HIV risk (e.g. high population density, presence of industrial enclave or transportation hub, proximity to international border, substantial in- and out-migration, high number of drug arrests, etc).
  - Sometimes it is necessary to conduct rapid assessments to collect this type of information.
- The advantage of this type of categorization is that algorithms can then be developed to extrapolate from areas with size estimation data to areas in the same category that have no data.
- Identifying categories with large information gaps can be used to help prioritize particular areas for direct size-estimation. This process is illustrated in the diagram on the next slide.
Facilitator notes: This diagram illustrates the process of prioritizing areas for conducting local size estimation using the mapping-based or survey-based methods referred to earlier.

Creating national estimates requires size estimates from all geographic areas of the country. Since size cannot be measured everywhere, the approach is to:

- Divide the country into geographic units (can be administrative units)
- Categorize the units according to common characteristics related to HIV risk, using what is known about transmission dynamics and epidemic maturity
- Identify units/categories with large information gaps
- Where there are no gaps, assess the biases in the existing data, reconcile multiple estimates, and feed into the national estimate. Processes for doing this will be discussed in detail in later sessions of this workshop
- Where there are gaps, prioritize areas for additional data collection, collect the data using mapping or survey-based methods, assess biases, reconcile multiple estimates, and feed into the national estimate
- In places where there are gaps, but where it is not possible to collect additional data, develop provisional estimates using an extrapolation process based on data from areas with similar characteristics…until such time as better information can be collected

*Use these extrapolated estimates to feed into the national estimates*
Page intentionally left blank
Session 2

Introduction to Mapping Based Size Estimation Methods
Session 2 Summary: Introduction to Mapping Based Size Estimation Methods

Objectives

- To know the options in mapping-based size estimation methods
- To understand the key challenges and share country experiences in addressing issues with implementation of mapping-based methods
- To plan to incorporate best practices in future mapping-based size estimation exercises

Contents

This session reviews the mapping-based methods for local size estimation, which are primarily census or geographic mapping and enumeration. The value of mapping for planning and designing effective local interventions is highlighted, as a reason to do mapping beyond obtaining size estimates.

Specifics of what is covered in the session/ tips for facilitators

This session describes the process for geographic mapping, using a census-type approach (i.e. going to all sites and collecting a standard set of information for characterizing and estimating sizes at all sites). Options for adding elements of social mapping or other types of non-size data are discussed. A concrete example of the efforts and the result of mapping MSM in one part of the state of Andhra Pradesh in India is provided to give participants a flavour for the level of resource intensity called for by mapping. An example of the enumeration method is provided, in which a sample of sites is selected strategically to collect information which can later be extrapolated to all such sites.

Many countries have had some experience with mapping. These exercises may have various levels of rigor and may have been conducted for different purposes, but some of the key challenges for mapping, such as developing a comprehensive list of sites, distinguishing MARPS from the general population, and accounting for duplication because of population mobility would likely have been encountered. Sharing of past field experiences is a useful part of this session. The facilitator should guide participants in thinking about their past experience and how it could be enhanced by adopting best practices as described in the session or shared by other countries. Specific attention should be given to optimizing mapping methods for size estimation purposes.
Regional Training on Size Estimation of At-Risk Populations in Asia-Pacific

Session 2 – Introduction to mapping based size estimation methods
Objectives of this Session

• To know the options in mapping-based size estimation methods

• To understand the key challenges and share country experience in addressing issues with implementation of mapping-based methods

• To plan to incorporate best practices in future mapping-based size estimation exercises

Facilitator notes:

• In this session we will review the various approaches used for the category of local size estimation based on mapping of sites or venues where MARPS can be found.

• As the options are reviewed we will discuss challenges faced in implementing these methods and learn from real country experience on how to optimize the implementation of these types of methods.

• Most of the challenges faced can be addressed or minimized by incorporating some key best practices during the planning and implementation phases of these methods.
Facilitator notes:

As the name of the category suggests, all mapping based methods required some type of mapping or listing of the spots in a given geographic area where MARPs can be found. Obtaining as complete a listing of these types of spots as possible is the initial step of all types of mapping.

Within the category of mapping based methods, there are two approaches:

- The census approach involves field teams going to each site in the listing to collect specific information to characterize the site and estimate the size of the MARPS population found at the site.
- When there are not sufficient resources to visit all spots, enumeration offers a variation on doing a complete census of site. With this approach a stratified sample of sites is made and only those sites are visited for specific data collection. The results of the size estimation from this sample of sites is then extrapolated to obtain an overall size estimate. The key to reliable estimates from enumeration are based on how the sample of sites is taken.
Facilitator notes:

Now we will go into more detail about the basic census method of mapping, often referred to as geographic mapping, because the method associates size estimates and other data with specific geographic locations or sites in a particular area.

- The first step is to comprehensively list all the sites where a majority of MARPs are found. For example, for FSW this may be those places for soliciting clients; for MSM this may be sites where people cruise for sex partners. And for IDU, this may be places where injection happens, such as shooting galleries, or where the dealers are found.
- Once all the sites are listed, the field teams should visit each site to obtain a count or estimate of the number of MARPs found at that site. The size estimate for each site may come from a combination of direct observation, in which the field team tries to visually identify all MARPS and counts individuals; and interviews with key informants.
- Appropriate key informants are those who are familiar with the specific site and are observant or knowledgeable about the presence of MARPS at the site. The opinion of several key informants are taken and if there is wide disagreement a process of obtaining consensus among key informants is followed.
- When an estimate of size is obtained for each spot, the results are summed up to develop an overall estimate for the geographic area.
Social mapping – variation on census methods

- In addition to rough size, other characteristics of spots and HRGs in area are captured
  - Types of sexual/drug using practices
  - Patterns of movement between spots
  - Power structures operating in the site that influence risk
  - Social networking at the spot (i.e., how people know each other and interact each other)
  - Other issues/concerns of importance to HRG
- Methods for collecting data include:
  - Focus group discussion with HRG
  - In-depth interviews with HRGs
- Involvement of HRG community in collection of data, drawing maps, and identifying spots with social meaning gives context to geographic mapping
- Conducted as a add-on to geographic mapping, or as substitution in style of a rapid situation assessment for intervention planning

Facilitator notes:

Geographic mapping can be very resource intensive, but the advantage of the method is that it provides very specific information about the distribution and type of MARPS in different spots. In addition to collecting data on size, this kind of specific information can be extremely useful for local program planning and design of services.

- Adding an additional layer of qualitative data about the social structure of the site can be a valuable modification to the method of geographic mapping. The types of data about spots and the MARPS found in particular spots can include:
  - Focus groups or in-depth interviews. The engagement of members of the MARPS community as interviewers, or to draw maps of the geographic area and describe places with social meaning can provide a richer understanding to the patterns of movement and issues of importance to MARPS that can be crucial for adjusting interventions to be appropriate for a local context.
  - In some cases this qualitative layer is added on to a more quantitative geographic mapping of a town or city. In other situations, these types of more qualitative data are used as an initial phase of mapping or intervention planning, sometimes called a rapid situation assessment.
Facilitator notes:

This slide gives an example of how the community can be engaged in drawing maps of their local areas, in a technique called broad mapping.

- The process involves gathering a group of individuals from the MARPS community and provide a large piece of paper.
- Then through discussion and work as a group, a map of the local areas is drawn, and participants mark areas of social significance to their community on the map. More detail is added, so for the example of sex workers, the solicitation points are drawn, the location of the police station may be noted, the location of key landmarks, or the home of an important community members may be indicated.
- These types of maps can also be useful for indicating patterns of movement of MARPS between spots, which are important for reducing duplication in the size estimates.
Facilitator notes:

This diagram shows the general flow of a mapping exercise. It involves a series of steps:

- First is the planning and preparation of the exercise. Here the definition of the population being estimated, the types of venues to be visited, and the boundary of the geographic area being mapped are decided.
- Data collection tools specific to the population, the local context, and the protocol adopted are developed.
- The initial listing of known spots are listed and the field work involves estimating the size, gathering additional information, including potentially social mapping.
- The standardized data collection tool allows for a more uniform approach to conducting the field work spread across large geographic areas and in potentially numerous sites. Visiting sites and talking to key informants will result in identification of additional sites to add to the list, which also must be visited and assessed.
- After data has been collected and analyzed, the last step is to present the findings and discuss the implications of the results among various partners.
Facilitator notes:

- This slide gives an example of the amount of effort that was required for doing mapping of MSM in nine districts of a state in India Andhra Pradesh. This mapping was done in the context of launching a new intervention program and the exercise was used both to gather critical information for planning and the design of services as well as to engage and empower the community to take part in the intervention.

- The mapping involved large numbers of field team to sweep through large geographic areas and complete the mapping in a short, intensive period of time. First a one day planning and training of trainers session was held with a smaller group of team leaders and managers. Then members of the community were gathered (a large number – 120) to be the field teams and trained for 2 days in the purpose and method of mapping. Using structured tools designed specifically for this exercise the field teams went out and in 3 days completed the mapping of their areas. The teams came back together to collate and discuss the results as a group and decide on next steps for planning the intervention.

- You can see that in the 3 days, a large number of towns and hotspots were covered. Notice that the number of towns and hotspots varies a lot by district. In planning a mapping exercise, estimating these types of numbers will be helpful for planning the amount of time and number of people needed to complete the mapping.
Facilitator notes:

These data come from the same mapping exercise featured in the earlier slide.

- The table shows the number of MSM estimated for each district covered by the mapping.
- We can see that the size estimation portion involved quantifying several different sub-types of MSM. This is one of the strengths of mapping as the information collected can be quite specific about each site and will depend on what the specific needs of each program may be.
- To ensure standardized data are collected in each site, these types of sub-groups must be specified on the data collection tools and the training must ensure that the field teams are clear about the definitions used for each sub-group and how the size should be estimated by observation and talking to key informants.
Sources of census method data & implications of use for size estimates

- Geographic mapping for intervention planning or monitoring
  - A critical step for effective interventions for MARP.
  - Population definitions and geographic scope may be limited to intervention target
  - Completeness of spot listing depends on field strength of programme
- Sampling frame development for survey preparation
  - Clusters identified may reflect definition of group targeted by and scope of survey
  - Completeness of spot listing depends on field experience of survey teams and resources (e.g. time) to complete sampling frame
  - Size of cluster/spot may not resolve issues of duplication or movement across sites.
- Mapping primarily for size estimation purposes
  - Completeness of spot listing depends on field experience of mapping teams, cooperation of interventions in the area and independence of teams.

Facilitator notes:

For the census-style of mapping there are several ways this type of data may be gathered. The completeness of the mapping will also depend on how well the program has been able to establish rapport with the community and how well they know the area.

- Sometimes geographic mapping is conducted as part of planning for an intervention at the launch of a program or as a periodic update of the field situation as part of regular program monitoring. Many countries and donor agencies have built in mapping as an initial step for rolling out focused interventions for MARPS, but it is important to recognize that when mapping is done for the purpose of planning interventions, this may naturally limit the geographic boundary and definition of the MARPS used, to be specific to who is being covered by the intervention.
- Another context for doing detailed geographic mapping is in the preparation of a survey of MARPS. Some types of probability surveys require sampling frames that contain all sites or spots where MARPS are found. Again the completeness of this type of mapping depends on the geographic boundary and eligibility criteria of the MARPS included in the survey. The survey team may also face similar problems of access to the community and this may limit how comprehensive their listing of sites may be. It is also possible that the sampling frame development process does not address the issue of duplication of individuals who may frequent several spots.
- The final situation is when mapping is done for the specific purpose of size estimation. In this situation the exercise can be tailored to address the needs specific to the use of the size estimation data, in terms of geographic boundary and definition of the MARPS estimated.
Facilitator notes:

Let's look at the enumeration method of mapping. This is a modification of census mapping in which the field teams visit only a sample of sites and extrapolates their findings to the remaining sites.

- This requires a comprehensive listing of sites and is useful when that list is possible to gather through another method. Eg, if bar based sex workers are being estimated and the city is large and there are several hundred bars, it may not be feasible to go to every bar. There may be a listing of bars from the government agency which regulates alcohol sales. Using this listing it may be able to construct a stratified sample of bars. The bars could be grouped according to their geographic location within the city or by the size of the bar (if that information is available), or any other characteristic which may make it likely to have a different number of FSW.

- This slide gives another example of stratifying sites, there are 2 types of entertainment establishments where FSW work: karaoke bars & massage parlors & a complete listing of both types of establishments are available. When the sites are grouped by type we see we have 350 karaoke bars and 276 massage parlors. So a random sample of 50 bars and 50 massage parlours is selected from the list, and only those establishments are visited. The results are that on average each karaoke bar had 6.7 waitresses that also took clients and massage parlors had 2.4 women on average who were FSW.

- How could you calculate the size of this group of FSW from these data? [Give group a few minutes and then walk through calculation][350 * 6.7] + [276 * 2.4] = 3007 or about ~3000 [How could you randomly select the sample: Listing sites in some order (e.g. by geography, by registration number, by duration of operation) and using a random start and regular interval (i.e. systematic sampling)]
Enumeration method: When to use it?

- This approach is more efficient for getting an approximate size of a large, dense area, where there are many sites to cover.
- It will not be helpful for planning interventions at the site level.
- The reliability of the size estimate will depend on:
  - the completeness of the site listing and
  - the ability to characterize the sites adequately to draw an appropriate sample.

Facilitator notes:

So when is enumeration a good choice for mapping?

- This method is more efficient if the number of sites is very large and dense area to cover. But the disadvantage is that if specific information about each site is needed (e.g. the contact information of the owner of the establishment or the hours of operation etc.) especially other data besides the size of the MARPS population this will not be available at all sites.
- The reliability of the size estimate depends on two key factors: First the ability to list the sites comprehensively. This may be a problem for sites that are just street location where street-based FSW solicit as these sites are not likely to exist as a complete listing already. In most cases sites change frequently and new sites may be identified when visiting other sites and talking to key informants.
- The second limitation on enumeration is that the ability to generalize the findings from sites visited to sites not visited will depend on how well the sites are characterized and put into strata. If there is a lot of variation even within the strata used there may be a large variability in the final size estimated. For example in the previous example if some karaoke bars were very large e.g. 20 or 30 women who sell sex and some were very small e.g. only 1 or 2 women sold sex. The average of 6.7 waitresses per bar who sold sex would be associated with a wide confidence interval. And if for some reason the sampling did not achieve a representative sample this could greatly skew the size estimate.
Facilitator notes:

- Now we will talk about some of the key challenges or limitations when using mapping methods.
- Does anyone want to share some of the challenges they have faced in trying to use mapping in their country?
- [Give participants some chance to share]
Challenges to all mapping methods

- What if a subset of the HRG is not found at public sites?
  - Is mapping an effective approach for all types of IDU? Migrants?
- How to ensure a comprehensive listing of sites?
  - What if the sites are continually changing?
- How to count a mobile population?
  - How to avoid double counting/overlap of sites
  - What is the size of a site which has different #s of HRG at different times of day or days of the week
- How to distinguish people who fit the operational definition?
  - e.g. FSW soliciting at bus stands, MSM cruising in parks
- When exact counts are not possible, how do you reconcile different estimates from a range of key informants?

Facilitator notes: Challenges that are common to all mapping methods.

- 1st, mapping methods will only be useful for estimating the size of the MARPS who can be found at public sites. There is a portion of the population which prefers to remain hidden or does not go to public sites (e.g. some MSM only find partners through the internet or newspaper ads. So it is important to think about whether a segment of the community will not be included.
- 2nd, obtaining a comprehensive listing of sites is challenging, especially when the sites change frequently due to law enforcement or other local authorities which may be trying to disrupt MARPS from gathering in fixed places. Some sites may only be known to members of the community, and field teams must be able to gain access or learn about these sites.
- 3rd, most MARPS populations are not found in a fixed location, the population moves frequently & may go to several sites/spots on a regular basis. This leads to the difficulty of avoiding double counting of MARPS in different spots. Gathering information about these patterns of movement is very challenging.
- 4th estimating the size of the population requires distinguishing those people who are in the group of interest from other people who may also be at the site. For example, it may be difficult to distinguish MSM at a cinema from other men at a cinema.
- Finally, most mapping exercises rely on key informants to give an estimate of the MARPS at a particular site, but different key informants may give different responses and developing a consensus estimate may be difficult and require time to discuss and reconcile these responses.
- Many of you have already faced these challenges in the field. It would be great to have people share some of the techniques they have used to address some of these issues. [Give participants some time to share experiences or ask questions]
Facilitator notes: Things to do to strengthen how the size estimates from mapping based methods are used:

- It's extremely valuable to assess & document the limitations that are encountered when using the mapping methods. Eg., characterizing the MARPS who were included in the mapping, and those who may have been missed or not captured fully. Eg., that a mapping of MSM cruising sites was done and it was easier for key informants to estimate the size of male sex workers and transvestites because they were more “visible” but more masculine or clients of sex workers were more difficult to distinguish from non MSM also at the sites.
- In addition to this, it can be helpful to assess the completeness of the listings by triangulating sites identified by different groups (e.g. can the list of sites in a sample of town be cross checked by another field team; does the NGO have listings of sites they use for outreach & can these be compared to what the mapping field teams found; are there other key informants who can review the listing of sites & identify other sites which are not listed).
- An important technique for enhancing the results of mapping exercises is to engage members of the MARPS community to be members of the field teams. Being part of the team may increase the number of sites identified & to help build rapport with the larger community. This type of diversity tends to strengthen the ability of the team to do effective mapping.
- Ensuring explicit training for team members on how to establish rapport and distinguish MARPS in the field. Both classroom sessions and role plays as well as field experience are extremely valuable components to build into training.
- It can be helpful to visit complex sites more than once. This may be built into the protocol, especially if sites are very busy, large, or have very different operational times. Going back to talk to key informants to gain consensus on a final estimate for the site is also an important approach.
- Finally, some countries have begun to incorporate an aspect of validation in their mapping protocols. This involves selecting a random sample of sites where the mapping is repeated by an independent team and the results are compared to determine whether the number of sites and the size of each site are similar, or if there is wide, unexplained variation in the results.
Facilitator notes:

- In summary, we can say that mapping methods are an important set of tools for doing size estimation, but they are also a fundamental part of implementing MARPS interventions, that should be considered in all geographic areas where programs are implemented.
- The key to using the results of mapping methods is to be clear how the mapping data matches the intended definition of the population being estimated. In many cases, there is some degree of non-overlap between what is needed and the limitations of the method. By understanding more specifically what the difference is, it is possible to adjust the size estimate result appropriately, to get an answer closer to what is needed.
- Finally, mapping based size estimates have great potential to be very detailed and precise, especially when used in an application such as calculating program coverage. However, as in all methods, the reliability of the data depends heavily on the quality of the implementation of the method.
Session 3

Introduction to Survey-based Methods with a Focus on Multipliers
Session 3 Summary: Introduction to Survey-based Methods with a Focus on Multipliers

Objectives

- To describe the general principle and mechanics of using survey-based multiplier methods of size estimates
- To review key concepts in representative sampling of hard to reach populations, such as FSW, IDU, and MSM
- To compare characteristics of mapping based methods to survey-based multiplier methods for size estimation

Content

This session lays out the general approach for the multiplier method and discusses key assumptions, and general implementation issues, including the importance of probability sampling methods for MARPS.

Specifics of what is covered in the session/tips for facilitators

Multiplier methods are an important type of survey-based method, based on the comparison of overlap in population of two independent data sources. The session walks through the basic formula and uses a simple example to illustrate the calculation. The basis for the requirement of survey representativeness and accuracy of the second data source in relation to the multiplier formula is emphasized. A combination of math and diagrams to illustrate the “proof” of the multiplier is provided in this session, to appeal to different types of training participants. Similar diagrams are used to illustrate finer points of implementation in subsequent presentations on the multiplier.

Information about options in probability sampling for MARPS are briefly presented. Some participants may want more explanation about the methods. References can be provided, or additional time on the subject can go into more detail on sampling methods.

Finally, a general comparison of the populations and resource constraints of mapping compared to the multiplier are presented.
Session 3: Presentation

Regional Training on Size Estimation of At-Risk Populations in Asia-Pacific

Session 3: Introduction to survey-based methods with a focus on multipliers
Facilitator notes:

- In this session, we will review the basic concept behind the multiplier method, a technique that obtains size estimates from representative surveys of MARPS. Because this method depends heavily on how representative the samples from such surveys are, we will spend a little time discussing the options for sampling MARPS to obtain representative or probability samples. And then at the end of the session we will compare the two types of size estimation approaches: mapping and multiplier methods.
- Subsequent sessions in the training will address the specific implementation of different styles of multipliers, especially addressing the types of bias that may occur through the implementation of the multiplier.
Facilitator notes:

- To step back for a minute, it is useful to place the multiplier methods in a large category of size estimation methods which rely on surveys to obtain part or all of the information needed to collect the size of the data.
- In this grouping we have the multiplier method (either using program data or what is described as the unique object) and capture recapture. These two methods are quite related and are based on the same mathematical formula, but the implementation in the field looks different.
Facilitator notes:

The general approach for the multiplier method is as follows:

- It is a comparison of two independent data sources from the MARPS. The first source of data is a count or listing of individual MARPS who have accessed some type of service or may be members of a group or have received a “unique object” that is distributed as part of applying this method.
- Source 2 comes from a survey of a representative sample of the MARPS of interest. It gives the % of respondents who say they have accessed the service, or belong to the group, or received the object, corresponding to the definition used for Source #1.
- We can see from the formula that the idea is that the # of people in Source 1 over the total size of the group being estimated should be the same as the % of respondents who say yes in Source 2. This is because if the sample is representative than the people included in the count of Source 1 will be proportionally represented in the survey.
- With a little algebra we can transform the formula so that the total size of the group of interest is equal to the # or count of people in Source 1 over the % of people in Source 2.
Simple Multiplier example:

- **Source 1**: # of MSM who have visited a specific VCT in the last 3 months (N=150)
- **Source 2**: A probability survey of MSM that asks respondents whether they have visited a specific VCT in the last 3 months (25%)

Calculation: $0.25 = 150 / \text{total MSM population}$ or $\text{total MSM population} = 150 / 0.25 = 600 \text{ MSM in the population}$

Facilitator notes:

- Let’s look at a concrete example of how this works. Let us say we want to estimate the number of MSM in a city. And that we have a source of data from a specific place for counseling and testing that has a count of the number of MSM who came for testing in the last 3 months. Let us say that the number is 150.
- And then a probability survey of MSM in the city is conducted and one of the questions included in the questionnaire asks respondents if they have gone to the specific site for VCT in the last 3 months. Let’s say that 25% of the respondents in the survey said yes, they had visited the VCT in the last 3 months.
- Then if the people in the survey are representative of all the MSM in the city, then 25% should be equal to the actual number of people who got VCT (150) over the size of the total population. With this formula we estimate there are 600 MSM in this population.
- Note that it is very important that the list from the VCT is able to distinguish individuals and will not double count people who come several times for testing during the specified time period. It also requires that the listing in Source 1 distinguish MSM from non MSM who come for testing. And then finally the data from Source 2 must match Source 1 in the sense that all the people who are included in Source 1 must be eligible for the survey (e.g. live in the city where the survey is being done) and have a non-zero chance of being included in the survey.
Facilitator notes:

- One of the critical assumptions of the multiplier method is that the sources of data used must be independent. For example, the likelihood being in Source 2 (the survey) is not related to the likelihood of being included in Source 1. So MSM who are included in the survey should be just as likely to have gone to VCT as MSM who did not participate in the survey.
- This can be assured if the survey population is truly representative of the total MSM population.
- Achieving independence can be difficult due to the field realities of conducting a probability survey and the likelihood that some amount of selection bias is introduced.
Methods for obtaining probability samples of high risk groups

- Special probability sampling methods have been developed for high risk groups
  - Addresses issues of mobility and the hidden nature of high risk groups

- Time-location cluster sampling (TLC)
  - Clusters are defined in terms of both time and place
  - Clusters = accessible sites where HRG gather (e.g. hotspots)
  - Clusters are selected probabilistically and the survey team randomly approaches potentially eligible individuals at the cluster.

- Respondent driven sampling (RDS)
  - Seeds are selected to start the sample
  - Each participant is asked to recruit a limited number of eligible friends for the survey

Facilitator notes: To implement the multiplier method it is helpful to understand the options for obtaining probability samples of high risk groups. Conventional methods of sampling do not work well for most MARPS because these groups tend to be highly mobile and often hidden.

Two methods of sampling are currently used to obtain probability samples of MARPS such as FSW, MSM, IDU, and to some types of clients of FSW.

- 1st-style of sampling is called time location cluster sampling and is a variation of traditional cluster sampling. Here the cluster, or the unit of sampling is a hotspot or a venue where the MARPS can be found. Each cluster is defined both by a physical location (e.g. street corner, bar, cinema, etc.) and a time (e.g. time of day and day of the week). The sampling frame consists of full listings of all sites and times of day/days of the week which have different numbers and types of MARPS. Eg, a weekend for a street cruising spot may have a different group of MARPS than at the same location on a Wed night. Or the afternoon may have a different composition of MARPS than in the evenings. Clusters are selected randomly as in other types of cluster sampling and field teams go to the cluster at the appointed time and select respondents randomly.

- 2nd-style of sampling is called respondent driven sampling and is a chain recruitment method here, participants are asked to recruit their friends for participation in the survey. This method is a highly controlled & uses special statistical methods to analyze the data to obtain a probability sample that can be weighted to account for the method of selection of respondents. The advantages to the RDS approach is that it does not require a sampling frame and can reach segments of the population that do not appear at sites or venues. The key assumption of the method is that all members of the MARPS are connected socially in a single large network.
Facilitator notes: Each of the methods has strengths and imitations. This slide describes some of the key situations in which selection bias may occur with each method.

- For time location cluster sampling, the listing of clusters may be incomplete and biased in some way, e.g., it may be easier to list larger solicitation points for FSW and the smaller sites may get left out. However, the FSW who solicit at smaller sites may be different and under-represented in the sample.
- The survey team must also approach selected respondents at the spots, which may be inconvenient or uncomfortable for the MARPS members. This may cause some MARPS to hide or avoid members of the survey team and this group would be under-represented.
- Finally, the survey team may not be able to distinguish all the MARPS at the cluster at the time of the survey and selectively approach only those who are more visible.
- With respondent driven sampling, there is the challenge of how respondents may selectively choose which friends they try to recruit for the survey, e.g. those who may need the incentive more than others.
- It is also possible that the whole MARPS community is not connected socially and less social pockets of the population are missed because they are only found in smaller networks that the sampling method does not reach. For this reason, RDS does not usually work as well with clients of FSW who may not disclose to their friends that they buy sex and may not be part of a larger social network.
Facilitator notes:

Now, with that brief review of the multiplier method, we can compare the uses of mapping and multiplier methods in different contexts. (read slide)
Session 4

Implementing the Multiplier Method - Programmatic Multipliers
Session 4 Summary: Implementing the Multiplier Method - Programmatic Multipliers

Objectives

- To describe the mechanics of implementing programme-based multipliers
- To identify key types of potential bias and their effect on resulting size estimates
- To summarize best practices in implementation, and to reduce potential bias related to the implementation of size estimation using programme-based multipliers

Contents

This session provides detailed examples of how the programme-based multiplier works and key issues in implementation that can create a bias in the resulting size estimate.

Specifics of what is covered in the session/tips for facilitators

The session begins with a review of the programme-based multiplier method and describes a concrete example of implementing the method with IDUs. The session then reviews common problems experienced when implementing program-based multipliers, and then goes through a series of scenarios where assumptions are violated or there are problems with the data. Participants are given the opportunity throughout the session to consider the types of bias that will be introduced into the size estimate for each these scenarios, and to think through the implementation problem that caused it, and how it might be avoided with better planning.
Session 4: Presentation

Regional Training on Size Estimation of At-Risk Populations in Asia-Pacific

Session 4: Implementing the multiplier method: programmatic multipliers
Objectives of this session

• To describe the mechanics of implementing programme-based multipliers
• To identify key types of potential bias and their effect on resulting size estimates
• To summarize best practices in implementation, and to reduce potential bias related to the implementation of size estimation using programme-based multipliers

Facilitator notes:

• We have just seen how the basic “mechanics” of the multiplier method works.
• In this session we will look into the details of implementation of this method for programme-based multipliers
• We will identify some of the common problems that can occur, and summarize best practices for minimizing these problems.
Facilitator notes:

As we saw in the previous presentation, the multiplier method relies on two sources of data:

- the first one being a count or listing of MARPS members who access a particular programme or service
- and the second being a probability survey with questions designed to measure the overlap between the two data sources.
Examples of programme data that can be used as a multiplier

- # of individuals contacted by an outreach worker or peer from a specific NGO
- # of individuals who were given condoms by outreach workers or peers from a specific NGO
- # of individuals who were given needles and syringes from outreach workers or peers from a specific NGO
- # of individuals who attended a drop-in centre of a specific NGO
- # of individuals who had an STI check-up at a specific NGO clinic
- # of individuals newly registered with a specific NGO during the past three months

For size estimation multipliers, it is CRITICAL to have a count of individuals as opposed to contacts

Facilitator notes:

First let’s explore the kinds of data that might be gathered by a programme for a MARPS, that could be used as a multiplier.

- Most HIV prevention programmes have systems for keeping track of the quantity of services they provide. These systems sometimes track numbers of services provided (e.g. # of condom demonstrations performed by outreach workers, number of STI exams performed in the clinic each month, number of needles and syringes distributed by peers).
- Sometimes they track contacts (e.g. # of contacts made by peer educators with street-based sex workers), and sometimes they track individuals (e.g. # of individuals contact by and outreach worker, and all the other examples listed on this slide).
- Some programmes have more sophisticated systems for tracking data than others, and some have better quality control than others. But the reality is that it can be very difficult to count individual people contacted, as opposed to the number of contacts made.
- For size estimation multipliers, it is critical to have a count of individuals, as opposed to contacts. And the count must be fairly accurate, otherwise there will be problems with the size estimate.
Other important characteristics of programme-based multipliers

- Programme based multipliers must be specific in terms a time-period e.g.
  - # of street-based sex workers who had an STI exam in April, May and June
  - # of IDUs who were registered in the programme in the month of June

- The programme-multiplier must include only those people who meet the definition of the group you are trying to estimate
  
  e.g.
  - If you are trying to estimate the number of street-based sex workers, then the programme multiplier must include only street-based sex workers, not street-based plus home-based and brothel-based

Facilitator notes:

- Programme-based multipliers must be specific in terms of the time-period they reference. So it must be clearly articulated exactly what data the programme is capable of generating, so that the survey questions can be formulated to exactly match.
- Programme-based multipliers must be very specific to the group you are trying to estimate. So, for example, if you are trying to estimate the number of street-based female sex workers, then you multiplier must include only street-based sex workers and not street-based sex workers and brothel-based sex workers.
Facilitator notes:

- Just as there were specific requirements for the programme-multiplier data, there are also requirements for the survey data. In addition to coming from a survey that is representative of the group being estimated, the survey must also encompass the area that is covered by the program. In other words, the people who will be counted on the programme listing must also have a chance to be selected in the survey.
- So, for example, if the multiplier being used was # of individual street-based sex workers who received condoms from a peer, if the count is for both city A and city B, and the survey covers only city A, there will be a “mismatch” between the two data sources.
- However, the reverse is not true. The survey area can be broader than the listing area. This is not technically considered a “mismatch” in the sense that it will not cause the estimate to be biased. However, the less overlap there is between the two data sources, the less precise the estimate will be. In other words there will be a wider confidence interval around the estimate so there will be less certainty as to the actual size.
Facilitator notes:

This slide shows a straightforward example of how the programme multiplier should work.

- We want to estimate the size of the IDU population in a specific city.
- The “Recovery Alliance”, a major NGO working with IDUs in the city, reports that 240 IDUs visited their drop-in centre in the month of May.
- A survey conducted with IDUs in the area during the first week of June, asked respondents whether they had visited the “Alliance” DIC in the past month. They showed a placard to the respondents with the Recovery Alliance logo printed on it, and gave the address of the DIC.
- When the results of the survey were analyzed they find that 40% of the respondents reported going to the Alliance the drop-in center last month.
- Using the standard multiplier formula we calculate that there are 600 IDUs in the city.
- The idea is that the sample (people in the orange rectangle) is representative of all the IDUs (people in the green rectangle). And similarly there are a certain number of IDUs who visit the Recovery DIC (see overlap of yellow circle and orange rectangle).
Facilitator notes:

The group should be given a chance to brainstorm about the type of problems that can occur with the programme multipliers, that could potentially cause the resulting size estimates to be biased.

Suggested answers:

- Programme fails to record all the beneficiaries they provide services to
- Programme unable to track individual beneficiaries (as opposed to contacts)
- Programme counts beneficiaries who are not part of the group being estimated (e.g. unable to distinguish street-based sex workers from other female beneficiaries)
Inaccurate programme counts

- Field examples:
  - NGO unable to track unique beneficiaries (counts duplicates)
  - NGO fails to remove inactive beneficiaries from its register
  - Program counts beneficiaries who are not part of the population of interest (e.g. DUs and IDUs when survey is only IDUs)

- What is the effect?
  - pop size = count/ % in survey
  - If count is overestimated
  - pop size will be overestimated

Facilitator notes:

- The first type of problem we will explore is inaccurate programme counts. This is something that might happen if the NGO were unable to track unique beneficiaries. For example, if the program tracks contacts instead of individuals (e.g. number of visits to the DIC, as opposed to number of individuals who visited the DIC) or if there is double-counting (e.g. same IDUs is contacted by more than one outreach worker and is counted by both).

- There can also be a problem if the programme definition does not match the definition of the population whose size is being estimated. For example, if the estimate is for IDUs, and the survey is among IDUs, but the program counts both injecting drug users and non-injecting drug users.

- What does this do to the estimate? [give the participants some time to work out the answer and then walk through the calculation]

- In this situation, the list or count from the program appears to be larger than it actually is in reality. So this means that the numerator in the size estimation formula is inflated, resulting in a larger size estimate that is larger than it would have been if we had used the real (lower) number.
Facilitator notes:

In this example, the NGO fails to count all the beneficiaries they provide services to. This can happen if the systems for tracking beneficiaries or services, etc. are weak. So the programme may count only a portion of beneficiaries, which will cause the count to be lower than what it actually is.

What does this do to the estimate?[give the participants some time to work out the answer and then walk through the calculation]

- In the diagram, the yellow circle represents the group who are counted by the NGO. The darker green box represents the group of people who received services, and when asked in the survey whether they received services, they said that they did, but in fact they were not counted by the program as having received services.
- In this case the numerator in the size estimation formula will be lower than it should be, causing the overall size estimate also to be lower. (Assuming of course that respondents who did receive services are able to report accurately that they did).
Facilitator notes:

Now let us turn to measurement problems that can happen with the survey instead of the program counts.

In this example, it is the survey respondents who report having received services when actually they did not. This can happen if the questionnaire is not clear, or the NGO is not well branded (e.g. if there are multiple NGOs in the area providing services, and respondents are not able to accurately distinguish having received services from the NGO in question, as opposed to some other NGO).

What does this do to the estimate?[give the participants some time to work out the answer and then walk through the calculation]

- The result is that the NGO will report the actual count, but the proportion reporting receiving the services will be higher than it should be. For example, perhaps 30% will report receiving services, when actually it was only 20%. This will result in the denominator in the formula being too high, which will cause the overall estimate to be too low.
- This problem can possibly be reduced if the NGO has very strong branding, and if the interviewers are able to show respondents something that is easily recognizable as being specific to the NGO in question (such as their logo).
Facilitator notes:

In this next example, we have the opposite situation, i.e. the respondent received services, but fails to report it when asked as part of the survey. Again this can happen because of confusion over the question, or failure to remember on the part of the respondents.

What does this do to the estimate? [give the participants some time to work out the answer and then walk through the calculation]

- In the diagram, the lower-right quadrant of the yellow circle represents the group of people who received services from the NGO and were selected as respondents in the survey, but who reported not having received services. This makes the proportion measured in the survey lower than the real proportion would have been if people had reported accurately.
- This in turn makes the denominator in the formula too low, with the result that the size estimate is too high.
Facilitator notes:

What happens when there are multiple measurement problems?

- In this case there are two problems. The first is that some of the survey respondents who actually received services said that they did not, causing the denominator of the formula to be lower than it normally should be, with a resulting overestimate of the population size.
- However, it turns out that the NGO has also underreported the number of beneficiaries they served, which causes the numerator of the formula to be lower than it otherwise would have been. This has the effect of lowering the size estimate.
- In this case the two effects are working in opposite directions, so the effects may in fact cancel each other out. The problem is that you will never know that this is happening unless you have an accurate assessment of the biases, and are able to take them into account. Bias like this cannot actually be measured, and it is best avoided.
Survey not representative

- Field examples:
  - Respondents who are in contact with intervention more likely to be identified and sampled, and/or less likely to refuse participation.

- What is the effect?
  - pop size = count/ % in survey
  - % of respondents who report participation in the intervention is larger than in the actual population
  - Population size is likely to be underestimated for programme-based multipliers

Facilitator notes:

- In this final example, we have a situation where the survey was supposed to be representative of the population but it wasn’t. As you can see in the diagram, in reality, around half of the population was part of the intervention & the other half was not. Sample, represented by the yellow box, shows around 75% of the sample was part of the intervention population, so the survey was not actually representative of the whole population. It was biased toward intervention participants. This can definitely cause problems for the size estimate.

- In the case of the programme-based multiplier, there is a greater likelihood of the sample including people who have received services. Let’s say that 75% of the sample report receiving services, where as in the actual population (as visible in the diagram), only 50% of the population really received services.

- This will have the effect of making the denominator of the size estimation formula higher than it would have been if the survey had been more representative. The high denominator will cause the size estimate to be too low.

- Surveys being biased toward interventions are a common problem. It can be caused by many different factors, including sampling frames being developed with the help of NGOs, and even survey respondents being identified with the help of the NGO.

- Even if the sampling process itself does not have these types of problems, respondents themselves may be more likely to participate in surveys if they are already in contact with an NGO, just because of the familiarity and increased comfort level of a survey team, especially one that identifies itself as working with the program.
Facilitator notes:

Programme-based multipliers are conceptually very easy to implement, but actually difficult to implement well. They are subject to a number of potential implementation problems which can cause the size estimates to be significantly over or underestimated, and you will never know it.

These types of biases cannot be measured, and you will never know about them unless you actively assess them. And even then, they will be difficult to quantify. Even worse, in the likely event that multiple problems occur simultaneously, it becomes even more difficult to disentangle the effects.
The best way to deal with bias is to avoid it

- Two data sources must be well-matched in terms of population definition and geographic coverage area (survey should encompass the entire catchment area for the count)
- Survey questions should be well matched to the count/list data
- The potential quality of the multiplier data should be assessed ahead of time
- The survey should be representative of the population being estimated

Facilitator notes:

- The best way to deal with these kinds of biases is to avoid them. This is best done at the planning stages of the exercise, when you will need to make certain that:
  - The two data sources are well matched in terms of population definition, and geographic coverage area.
  - The survey questionnaire is well matched to the count data. If NGO data are being used, it is incumbent upon the researchers planning the survey to be very familiar with the way the multiplier data are measured and recorded. For example, if an NGO is able to generate quarterly numbers only (e.g. number of individuals registered April, May, June), but is not able to break that down into monthly figure), then it would not work ask respondents whether they had been registered in the past one month, because this could not be matched to the multiplier. You would need to ask respondents whether they were registered in the past three months, and that would only work if the survey were done in the month after the quarter had ended (in this case, early July). That is a rather narrow restriction, so this might not be a good choice of multipliers, or it would have to be tweaked in some other way in order to be usable.
- The potential quality of the multiplier data should be assessed before designing a size estimation exercise around it. We sometimes have very wishful thinking that data quality is better than it is. On the other hand, the numbers do not have to be exact. It may be better to have a good approximation than no information at all.
Session 5

Implementing the Multiplier Method – Unique Object Multiplier
Session 5 Summary: Implementing the Multiplier Method – Unique Object Multiplier

Objectives

- To describe the mechanics of implementing the unique object multiplier
- To identify key types of potential bias and their effect on resulting size estimates
- To summarize best practices in implementation to reduce potential bias in unique object size estimation

Contents

This session provides detailed examples of how the unique object multiplier works and key issues in implementation that can create a bias in the resulting size estimate.

Specifics of what is covered in the session/tips for facilitators

The session begins with a review of the unique object variation of the multiplier method and describes a concrete example of implementing the unique object with street based sex workers. The session reviews common problems experienced in the field and encourages participants to consider whether the resulting bias will underestimate or overestimate the true size of the population. With each example, the participants should be encouraged that this implementation problem is avoidable through careful planning, training, and implementation of the method.
Session 5: Implementing the multiplier methods: unique object multiplier
Objectives of this Session

- To describe the mechanics of implementing the unique object multiplier
- To identify key types of potential bias and their effect on resulting size estimates
- To summarize best practices in implementation to reduce potential bias in unique object size estimation

Facilitator notes:

- In this session we will get into the details of implementing the unique object style of multiplier method.
- We will also discuss some of the key problems that may occur with the implementation of this method and summarize the best practices for minimizing these problems and potential sources of bias.
Facilitator notes:

- With the unique object multiplier instead of relying on counts or listings of MARPS members from a source of programmatic data, An object is chosen that is unique (i.e. cannot be obtained from a source other than this activity) and that can be distributed to individuals in the MARPS community who meet the eligibility criteria of the survey. This distribution of unique objects is started and completed 1-2 weeks before the survey starts.

- The primary advantage of the unique object is that there is greater control by the survey team in ensuring the reliability of the count used in Source 1.

But the choice and distribution of the unique object requires some thought and attention to implementation to ensure that the increased control results in a more reliable estimate.

As always, the method works only if the survey data used in Source 2 comes from a representative sample of the MARPS being estimated.
Facilitator notes: Let us look at a concrete example of how the unique object works:

- If we want to estimate the size of the street-based sex worker population in a specific city. And if the survey team is able to produce or obtain a special key chain that is unique and distinctive and asks peer educators to distribute 150 of these key chains to street based sex workers in the city, 2 weeks before the survey is going to start.
- In the questionnaire, respondents are asked whether they received the key chain from a peer educator in the last two weeks. Each interviewer has an example of the key chain to show to respondents to help them understand what unique object is being referred to.
- When the results of the survey are analyzed they find that 10% of the respondents say they had received the key chain.
- And using the multiplier formula we calculate there are 1500 street based sex workers.
- Again, the idea is that the sample (people in the orange rectangle) is representative of all the street based sex workers (people in the blue rectangle). And similarly there are a certain number of FSW who receive the unique object (people in the yellow circle) and a proportional sample of them are selected for the survey (see overlap of yellow circle and orange rectangle).
Facilitator notes:

Let us do some brainstorming about what could be some problems for implementing this method. [Give participants a chance to volunteer ideas. If participants are slow to respond, give them some ideas as described in the bullets.]

These are all good ideas, let us look at some of the common problems and discuss how they can be addressed in implementation.
Facilitator notes:

- The first type of problem is not knowing an accurate count of people who received the unique object. Sometimes distribution happens such that not all the people who receive the object are eligible for the survey. (Here represented by the yellow circle extending beyond the blue rectangle). A concrete example might be when unique objects are given to both street based and bar based sex workers, but the survey and the size estimate is supposed to be for only street based sex workers.

- What could be the effect on the estimate when this problem occurs? - If we look at the formula, the inaccuracy lies in the # of objects used in the numerator. We may think 150 objects were given to street based sex workers, but in reality only 100 objects went to street based sex workers. So the estimate we calculated based on the numerator being 150 objects comes out much higher, than if we plugged in 100 (the actual number) in the formula.

- This problem with implementation leads to an overestimation in the size of the population.
Facilitator notes:

What happens when the count of objects distributed is larger than what was actually given out.

Sometimes this happens when the people distributing the object are not able to access so many people and end up keeping or discarding some of the unique objects. Or in one case, peer educators liked the object and didn’t understand the purpose of the exercise, so kept some of the objects for themselves.

What does this do the estimate?[give the participants some time to work out the example, then walk through the calculation]
Facilitator notes:

What if we have the situation, when respondents say that they have received the object when they really have not.

This might occur when the object is not unique or resembles something that has been given to many people or is something that people can buy in a store. Or if the object is something that friends want to share with each other, so the object may be passed on to more than one person and each of these people will say they have received it. Or if the interviewer does not have an example of the object to show respondents, they may misunderstand the object that is being described and confuse it with something else that they received recently.

What effect does this have on the size estimate? [Give respondents a chance to decide and then walk through the answer with them.]
Facilitator notes:

Sometimes the opposite situation occurs, where respondents did receive the object but don’t remember. And if these people are in the survey they will say no, they did not receive the object.

This can occur if the object is distributed too far in advance of the survey and the object is not very memorable, so the respondents don’t remember accurately. Or if the respondents are not shown the object during the interview and misunderstand the object the interviewer is describing.

What effect does this have on the estimate? [give the participants time to work out the example then walk through the answer with them]
Objects are given to people who are more likely to participate in the survey (dependence of Source 1 and Source 2)

Field examples
- Peer educators hand out objects to people they know, and listing of hotspots used for survey come from NGO
- Unique objects are given to people as they are recruited for the survey
- Police raids are very common in an area, IDU who are willing to be approached to be given the unique object are more likely to be approached and accept an invitation to participate in the survey

• What is the effect?
• pop size = # of objects/ % in survey
  • If real % in pop is smaller
  • than % in survey, then pop size is UNDERestimated

Facilitator notes:

Sometimes the assumption of independence is violated, and there is a situation in which the people who receive the object are also more likely to participate in the survey. This may happen if the peer educators are distributing objects and they are more likely to give the objects to people they see regularly and are part of the intervention.

• This would be ok, EXCEPT IF the way the sampling is done for the survey there is a greater likelihood for people in the intervention to be selected. This happens when the listing of spots used in the sampling frame comes from the NGO and over-represents the sites they are familiar with and have the strongest intervention coverage.

• Or in one case, there was a situation where distribution of the object occurred during the survey and the unique objects were given to people as they were recruited for the survey. This was a situation in which insufficient training was given to the survey teams about how to manage the unique object.

• Another more common occurrence is when MARPS are very vulnerable and subject to raids or harassment by local authorities such as the police. In this situation the person who is open to receiving the unique object from a survey team worker and the person open to being approached and invited to participate in the survey are skewed in the same direction.

What effect will this have on the size estimate?[Give the participants some time to work through the example and then walk through the answer.]
Facilitator notes:

Then there is the situation where the unique objects are not distributed randomly.

For eg, what if only people who know the peer educators receive the unique object, or what if unique objects are distributed only to sex workers who come to the STI clinic for a checkup. What if the unique objects are only given out at a small number of the hotspots in the city, or what if only one sub-type of all the sub-types included in the survey are given the object. For example, if only brothel based sex workers receive the object, but the survey is for both streets based and brothel based sex workers.

We can visualize this situation by looking at this diagram. Let’s say that the people on the left half of the box are reached by peer educators, while people in the right half of the box are not. And the unique object is given out by peer educators, so only people on the left hand of the box receive unique objects. How will this affect the estimate? [Give the participants time to work out the example and then walk through the answer.]
Optimizing the unique object method

- Improve the precision of the survey estimate, i.e. narrow 95% CI for point estimate of % of respondents who received the unique object
  - Increase the number of unique objects to increase the % of respondents who received it
  - Distribute the object widely across the HRG population

- Maintain independence of the distribution of the unique object & participation in the survey

Facilitator notes:

- Although it does not matter that the unique object is distributed randomly, we can optimize the unique object by planning to distribute the object widely in the eligible population. Doing this will help with the precision of the estimate that is obtained through the multiplier method formula.
- How does distributing the object widely help with the precision of the estimate we get? We can think about this using what we may already know about how sampling and statistics work. For any variable, if there is a high degree of clustering, or clumping of a characteristic of people in the population, when we take a sample, we have a higher chance of completely missing the pocket of people in the population with that characteristic. The same is true for the unique object. If we distribute the object in only one part of the population, there is a higher chance that due to sampling error we will miss the pocket of people who received the object.
- So the object should be distributed as widely as possible, even if this is not done in a random way. Or for example, the object can be distributed only to people who the peer educators know, but ideally the peer educators would cover a wide area of many of the hotspots across the area where AMRP are being estimated.
- Based on a similar idea, we can also distribute a larger number of objects which will increase the percentage of people in the survey who will say they received the object, which will also likely result in more precision around the estimated percentage.
Facilitator notes:

So let’s take a step back and summarize what matters and does not matter in how the unique object is implemented.

It is important that there is an accurate count of the number of objects given out to people who are eligible for the survey. There should be independence between who gets the object and who participates in the survey, and the object should be distributed as widely as possible, not only to people in a small number of clusters/hotspots.

It does not matter that the object is distributed randomly, and there is no strict rule on the number of objects that are distributed. However, it does help that the number of objects is a high enough of percentage of the MARPS to obtain a reasonably precise estimate of size.

Summary of what matters and what does not matter

- Does matter
  - Accurate counts of the object given out to those eligible for the survey
  - Independence between who gets the object and who participates in the survey
  - Wide distribution of the object, i.e. not clustered by “spots”

- Does not matter
  - Random distribution of the object
  - Number of objects given out relative to the size of the population or the sample size
Essential implementation issues for the unique object method

- Only those eligible for the survey should receive the object
- Object should be easily recognizable by respondents
- Object should not be available through another source
- Only one object should be given to the same individual
- HRG members should not redistribute/give their unique object to another HRG member
- HRG members should not be motivated to collect more than one object
- Accurate counts of how many objects are distributed should be maintained
- Object should not be highly visible or ‘mark’ a person as a HRG member

Facilitator notes:

The following tips for planning the unique object implementation should be considered and incorporated in the training of people involved in the field work. [Read the bullets]
Key messages about multipliers

- Multipliers are a relatively low cost option for size estimates, if a survey is already planned.
- Every time a survey of HRG is done, a multiplier should be included.
- But, the reverse is not true, i.e. Surveys should not be planned primarily to use the multiplier method.
- Unique object multipliers afford the greatest control and potentially most accurate estimates of size.
- If the estimate is biased, it’s not the method that’s the problem, but the implementation of the method,
- so
- Invest on the front end, to avoid trouble on the back end

Facilitator notes:

So to summarize the overall use of the multiplier method, whether using programmatic data or the unique object. [read bullets]
Page intentionally left blank
Session 6

Implementing the Capture – Recapture Method
Session 6 Summary: Implementing the Capture – Recapture Method

Objectives

- To define the capture-recapture method of size estimation
- To explore the similarities and differences between the capture-recapture and multiplier methods
- To highlight reasons for choosing one or the other

Contents

This session defines the capture-recapture method, describes how it works, and shows how it is similar conceptually and mathematically to the multiplier method.

Specifics of what is covered in the session/tips for facilitators

This session begins with a description of the capture-recapture method as a variation on the multiplier method. Participants are walked through the logic and sequence of conducting a classical capture-recapture data collection exercise, and provided with the formulas for using the data collected to calculate a size estimate. Then the session goes on to describe variations on how capture-recapture can be implemented, including one design that is almost identical to the unique object multiplier, the main difference being that the multiplier survey collects other data in addition to the size estimation data. The critical assumptions for CRC are reviewed, and some conclusions are presented about the comparative advantages of the CRC and multiplier methods.
Regional Training on Size Estimation of At-Risk Populations in Asia-Pacific

Session 6 – Implementing the capture-recapture method
Objectives of this session

• To define the capture-recapture method of size estimation
• To explore the similarities and differences between the capture-recapture and multiplier methods
• To highlight reasons for choosing one or the other

Facilitator notes:

In this session we will briefly describe the capture-recapture (CRC) method of size estimation, compare it to the multiplier method, and highlight reasons for choosing one method or the other.
Facilitator notes:

- Capture-recapture is a variation of the multiplier method, equivalent to it mathematically, but different in its implementation.
- Like the multiplier method, CRC measures the overlap between two data sources. As a “classical” size estimation method, it has typically been used to count animal populations that were hidden and in-motion (hence the famous fish-pond reference).
- We have categorized CRC under “survey-based” methods, because it requires at least one probability survey (even if it is a “take-all” survey). This differentiates it from mapping, in terms of the types of resources and technical capacity needed to implement it.
Multiplier: Capture Recapture (CRC)

- 3 phases
- Map all the sites/hotspots where the population to be estimated can be accessed
- Capture 1 - “Tag” all and keep track of # tagged
- Capture 2 – “Tag” all and keep track of
  - # “retagged”
  - # tagged for the first time

Facilitator notes:

- Typical implementation of capture-recapture involves identifying (i.e. mapping) all sites where members of the population in question can be accessed and then conducting two (or sometimes more) successive field visits where all members of the population are “captured” (i.e. tagged).
- During the first visit, the number of people tagged is recorded
- During the second visit, in addition to the number of people tagged, the number of people retagged is also recorded
Facilitator notes:

This flow chart illustrates the process for capture-recapture and also shows the formula for calculating the size estimate.
Operational example of classic capture-recapture

- Identify (i.e. “map”) all sites where MSM are found.
- First capture
  - Go to every site
  - “Tag” all the MSM met at the site by giving each one an memorable object (e.g. a yellow card with an invitation to an event, or a small memorable gift, etc).
  - Keep track of the number “tagged”
- Second capture
  - Go back to all the same sites a or so later
  - Again “tag” all the MSM met at each site, this time using a slightly different object (e.g. a red card with an invitation to an event, or a different gift).
  - Ask every MSM you “tag” whether he received the object given at the first tag.
  - Keep track of the total number tagged, and also of the number who were retagged (i.e. also tagged the first time)

Facilitator notes:

This is a list of the steps that one might follow for conducting a classical capture-recapture exercise with MSM

The fieldwork involves:

- Identifying or mapping all high risk MSM sites (i.e. sites where MSM tend to look for sexual partners), and then sending a field team at a particular time to each of the sites to “tag” all MSM met. Tagging at all times should be done within a short period of time to avoid double-tagging (of people who move between sites), and care should be taken not to tag people twice
- Tagging means giving something to the person identified, e.g. a yellow card with an invitation to an event, or information about services, etc., or a small memorable gift. The object should be unique (i.e. something that the person would not be likely to obtain from anywhere else). The number of MSM tagged is recorded
- After the initial tag, another tag is done a week or so later, this time using slightly different gift (e.g. a red card with an invitation to an event, or a different small gift). Each MSM who is tagged is asked whether he was also tagged the first time, and a tally is kept of how many MSM are tagged in total, and how many were retagged. With this information, a size estimate can be calculated
Facilitator notes:

There are many variations on how a CRC can be conducted

- If resources are limited, it may be possible to go to only a subset of sites or and tag only a subset of eligible individuals. However, if such a scheme is followed, certain “rules” should be observed
  - The first capture need not be random, but the second one must be random.
  - It will also be important to ensure that a large and geographically diverse proportion of all sites are visited, otherwise there may be implications for the precision of the results
- Operationally and conceptually, this is very similar to the unique object multiplier design

Mathematically it is the same as well, as shown on the next slide
CRC and Multiplier Mathematically Equivalent

<table>
<thead>
<tr>
<th>Capture Recapture Formula</th>
<th>Multiplier Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S = C_1 \times \frac{C_2}{R}$</td>
<td>$S = \frac{M}{P}$</td>
</tr>
<tr>
<td>Where:</td>
<td>Where:</td>
</tr>
<tr>
<td>$S=$Size</td>
<td>$S=$Size</td>
</tr>
<tr>
<td>$C_1 =$ # captured 1st time</td>
<td>$M =$ # of objects distributed</td>
</tr>
<tr>
<td>$C_2 =$ # captured 2nd time</td>
<td>$P =$ Proportion of those surveyed</td>
</tr>
<tr>
<td>$R =$ # captured both time</td>
<td>who received objects</td>
</tr>
</tbody>
</table>

C1 is equal to M
P is equal to R divided by C2
Therefore by replacement the two formulas are equal

Facilitator notes:

This slide walks through the justification for why the capture recapture and the multiplier methods are mathematically equivalent.

Each involves “tagging” a group of individuals and then retagging them from a probability sample.

By calculating the proportion that was tagged both times, it is possible to deduce the total population size, provided that the proportion comes from a random sample.
Assumptions and implementation requirements of Capture Recapture

- Requires two independent data sources
  - Note: The easiest way to ensure independence is to make at least one of the sources a random survey

- Ability to correctly identify “captured” and “recaptured” individuals

- No in- and out-migration between the two data sources

- Important to avoid small samples or clustering in both sources because of sampling error/precision

Facilitator notes:

The capture-recapture method has the same very important assumptions as the multiplier, which, if violated, would pose serious threats to the validity of the estimates.

- The first is the requirement of independence. What this means is that people who are selected in the first data source should not be more or less likely to be selected in the second data source. In other words, being selected the first time should not be related in any way to being selected the second time. The easiest way to ensure that this happens is to make sure that the second capture is random. That way, it really doesn’t matter how people are selected the first time, because as long as they are selected randomly the second time, they will not be more or less likely to be selected than people who were selected the first time.

- The second requirement is the ability to correctly identify the “tagged” and “retagged” individuals. That means that the individuals in the second survey need to answer correctly whether they were “tagged” in the first survey.

- The third requirement is that there should be no in- and out-migration of the population being estimated. In other words, the group should remain relatively stable between the two surveys.

- And finally, it is important to avoid very small samples or clustering when tagging, because the less overlap there is between the captured and the recaptured populations, the greater will be the uncertainty around the estimate (i.e. precision will be low).
Facilitator notes:

- Capture-recapture has the advantage of being simpler, since it is done only for the purpose of size estimation, and only requires data on whether the person is eligible, and whether they were tagged in the previous survey.
- However, implementation of CRC can be resource intensive, requiring similar inputs as a more complete survey or mapping exercise.
- The multiplier method on the other hand, if added to an already planned survey, such as a BSS or IBBS, offers a low-cost option for obtaining size estimates, which, if done with careful planning and implementation, in collaboration with the survey team, should be as good as a CRC, but with a much lower price-tag.
- Such surveys also have the advantage of providing a lot of other data, and if done with other sampling methods (e.g. RDS), have the possibility to count populations beyond only venue based.
Session 7

Network Scale-Up
Counting the Uncountable
Session 7: Presentation

The Network Scale-Up Method

Counting the Uncountable
Bangkok 2009
Finding the distribution of the number of people whom people know

- The objective is to understand the basic components of social structure.
- One quantity that seems important is the number of people whom people know.
  - We call this $c$

Facilitator Notes: The network scale-up method was developed in conjunction with our team’s research on the rules governing who people know and how they know them – that is, the rules of social structure. The particular list of people whom people come to know in a lifetime may appear random, but the rules governing whom we come to know are surely not random. One basic component of social structure is the number of people whom people know. We call this number $c$.

This number has a distribution, of course, and it probably changes across societies.

This number has a distribution, of course, and it probably changes across societies.


### A primitive model

- We can derive this number from an assumption.
  - Let \( t \) be the size of a population (e.g. the U.S.), and let \( e \) be the size of some subpopulation within it.
  - We assume that the fractional size \( p = e/t \) of that subpopulation also applies to any individual’s network, other things being equal.
  - That is, everyone’s network in a society reflects the distribution of subpopulations in that society.

Facilitator Notes - original network scale-up model was a four-part equation: (1) the event population (called \( e \)); (2) the total population (called \( t \)) within which \( e \) is embedded; (3) the probability, \( p \), that anyone in \( t \) knows someone in \( e \); and (4) the number of people whom people know, \( c \). History: Bernard was in Mexico City, after the earthquake in the 1985. No one knew how many people had died in that earthquake, but one person told Bernard that “there must be thousands dead, because everyone knows someone who died.” We did a random, representative street-intercept survey and found the percentage of people who reported knowing someone who died in the quake. That gave us two parts of the equation. We knew Mexico City had around 18 million people at the time) and we knew \( p \). We reasoned that if we knew \( c \), then we could solve for \( e \). This set up our research program on finding, not just the average \( c \) in a population, but its distribution.
The scale-up method to estimate $c$

To test this, we ask a representative sample of people to tell us how many people they know in many sub-populations whose sizes are known:

- e.g., diabetics, gun dealers, postal workers, women named Nicole, men named Michael


Facilitator Notes: To find $c$ and its distribution, we ask a representative sample of people in $t$ to tell us how many people they know in many sub-populations whose sizes are tracked reliably in public statistics. If our model is working, then we should be able to estimate accurately the size of those same populations. To the extent that we can do that, we have more confidence in our estimates of sub-populations whose sizes are unknown.

Do people answer accurately?

- This works only if people can and do answer our questions accurately and researchers recognized early on that this was a problem.

- Researchers expect that continued research on this problem will improve the estimates of hard-to-count populations.


Facilitator Notes: We had earlier studied the problem of informant accuracy in network data—that is, the extent to which people could report accurately with whom they interacted over various lengths of time. We knew from this research that informant accuracy would threaten the validity of the estimates from the network scale-up model. Just as with $c$, we expect that research on informant accuracy will result in incremental improvements in the estimates of $e$. 
A maximum likelihood estimate of an individual’s network size:

\[
c_i = 1 \cdot \frac{\sum_{j=1}^{L} m_{ij}}{\sum_{j=1}^{L} e_j}
\]

where there are \( L \) known sub-populations. (Here \( i \) is the individual, who knows \( m_{ij} \) in subpopulation \( j \).)

Network size is (the sum of all the people you say you know in some subpopulations of known size, divided by the total size of those subpopulations) times the population within which the subpopulations are embedded.

Facilitator Notes: To move beyond the basic model, we need to estimate the size of the network of each respondent in our scale-up surveys. We use a maximum-likelihood method to estimate the size of an individual respondent’s network. This method for estimating \( c \) is due to Killworth.

Estimates of \( c \) are reliable

- This doesn’t deal with the problem of informant accuracy, but the estimates of \( c \) for the U.S. are very stable.
- Across seven surveys, the original researchers consistently found an average network size of 290 (sd 232, median 231).
- And 290 is not an average of averages. It’s a repeated finding.
Is 290 is an artifact of the method?

- They tested this in three ways:
  1. Make the estimates using a different method.
  2. Experiment with parameters and see if the outcome varies in expected ways.
  3. Compare values of $c$ across populations of known relative sizes.

Facilitator Notes: We were surprised to find that the number 290 was so stable and we tried three different ways to disprove it. (1) We estimated the number using a method of counting that was different from the one we developed initially; (2) We introduced error into the data to see how it affected the outcome; and (3) We compared our results for $c$ across populations of known relative sizes.

Reliability I

(1) In one survey, they estimated $c$ by asking people how many people they know in each of 17 relation categories – people who are in their immediate family, people who are co-workers, people who provide a service – and summing.

This summation method once again produced a mean for $c$ of 290.


Facilitator Notes: In this test, we estimated $c$ by asking respondents to tell us how many people they knew in each of 17 categories: people in their immediate family, people they know from work, people who provide a service, and so on. Once again, our estimate of the average $c$ for the U.S. was 290.
Reliability II: Change the data

(2) They then changed reported values at or above 5 to a value of 5 precisely. The mean **dropped** to 206, a change of 29%.

They then set values of at least 5 to a uniformly distributed random value between 5 and 15. They repeated the random change (5 – 15), but only for large subpopulations (with >1 million).

The mean **increased** to 402, a change of 38% **-- in the opposite direction.**

Facilitator Notes: In the second reliability check, we introduced errors into the data. We tested whether the changes in our estimates of c conformed to the changes we introduced to the data.

Reliability III: Survey clergy

(3) They then surveyed a national sample of 159 members of the clergy – people who are widely thought to have large networks.

Mean c = 598 for the scale-up method

Mean c = 948 for the summation method

Facilitator Notes: And in our third test of reliability, we surveyed a sample of clergy and estimated the size of their networks. In this survey, we used our original scale-up method and the summation method. Here, the two methods produced quite different estimates of c, but both estimates are, as expected, larger than the 290 for the general population.
So, 290 is not a coincidence

1. Two different methods of counting produce the same result.

2. Changing the data produces large changes in the results, and in the expected directions.

3. People who are widely thought to have large networks do have large networks.

Facilitator Notes: In each case, then, the result of our reliability tests were expected, giving us some confidence in the estimate of c for the U.S.

Something is going on

- This next slide shows the probability, for two of their surveys, of knowing no one in each of 29 populations of known size, by the actual size of those populations.

- The two distributions track, except for the expected offset.
Reliability vs. validity

- The researchers were measuring something, and were doing so reliably, but if the model works, then they ought to be able to use it to estimate the populations whose sizes are not known.

- The researchers can create a maximum likelihood estimate for the size of an unknown subpopulation based on what all respondents reported and estimates of their network sizes – roughly speaking, inverting the previous formula.
Can we predict what we know?

- We can test this by seeing how well we do on the 29 populations of known size.
- In fact, the overall result is encouraging, but we don’t estimate some known-size populations well:

Facilitator Notes: The sequence for applying the method is as follows: (1) Ask a representative sample of people how many people they know from a set of a lists of populations whose size is tracked in public records. (2) Apply the maximum likelihood measure (in slide 8) to estimate c (the size of the network) of each person in the survey. (3) Estimate the size of each subpopulation. The word “estimate” here means checking the accuracy of the method by comparing what the method predicts against the known-sizes of the populations. The original scale-up method used 29 populations of known size and 3 populations of unknown size. With 20 known-size populations, there should be sufficient data to check the accuracy of the method for any given test. (4) If the estimates of the known-size populations are accurate, then the estimates of the unknown-size populations may be reasonable. Of course, if alternative estimates are available for the unknown-size populations, then these should be compared to the estimates from the network scale-up method.

\[ r = .79 \] \ldots rises to .94 without the outliers

Facilitator Notes: Note the two outliers in this graph (way above the diagonal). These are twins and diabetics. When these two data points are removed, the correlation rises from 0.79 to 0.94. The problem, of course, is that the outliers represent something that needs to be accounted for in order to improve the accuracy of the method. Just removing the outliers, therefore, improves the correlation but not the method itself. See next slide for more about this.
Over- and under-estimation

- There is a tendency for people to overestimate small populations (<2 million) and to underestimate large ones (>3 million).
- The two largest populations are people who have a twin brother or sister and diabetics.
- These are highly overestimated.
- Without these two outliers, the correlation rises from $r = .79$ to $r = .94$

Facilitator Notes: From the data in the previous slide, we see that people tend to overestimate small populations (people named Nicole, people who are undergoing kidney dialysis) and underestimate large ones (people who have a twin sibling, diabetics). In other words, when we ask people “how many people do you know who have a twin brother or sister”, the answers tend to be overestimates. And conversely for small populations: people tend to underestimate the number of people they know in small populations. A valuable piece of information comes from out of this: In building a network scale-up survey, researchers should shoot for known-size populations that range in size and type. For example, the known-size populations should not all be based on health conditions or types of jobs.

Stigma vs. not newsworthy

- Being a twin or a diabetic is neither stigmatizing, nor newsworthy.
  - From ethnographic evidence, personal information about close co-workers or business associates can take a decade or more to be transmitted ... and in the case of being a twin or a diabetic, may never be transmitted.
  

Facilitator Notes: What causes these tendencies to over- and under-estimation? Is it just the size of the subpopulation? Or is it the fact that some things (like being HIV-positive) are stigmatizing. In that case, we may know people who are HIV-positive but not know that we know them because they haven’t told us about their HIV status. But being a twin or a diabetic is neither stigmatizing, nor (as in violent crime) newsworthy. The fact that someone has a twin may simply never come up in conversation, even after decades of knowing someone. This finding comes from Shelley’s indepth studies of network processes.
Another encouraging result

- Charles Kadushin ran a national survey to estimate the prevalence of crimes in 14 cities, large and small, in the U.S.

- He asked 17,000 people to report the number of people they knew who had been victims of six kinds of crime and the number of people they knew who used heroin regularly.


Facilitator Notes: The UCR is the Uniform Crime Report system in the U.S. In this slide, the UCR estimates for heroin use in the 14 cities are the blue squares. The estimates from the network scale-up survey in the same cities are the red dots. For the most part, these estimates are very similar.
It’s less expensive, but …

- The fact that we track well with official estimates means only that we have a much, much less expensive way to get at these estimates – not that the estimates are correct.

- And estimates of other crimes in those 14 cities did not track so well.

Reliability, validity, and accuracy

- So, while definitely reliable and perhaps valid, our estimate of network size (and its distribution) is not sufficiently accurate.

- There are at least three sources of inaccuracy: transmission effects, barrier effects, and informant reporting.

Facilitator Notes: Which raises the question: How can we improve the accuracy of the method? There are at least three sources of inaccuracy in the method.
Compromising assumptions

1. Transmission effects: Everyone knows everything about everyone they know.
2. Barrier effects: Everyone in $t$ has an equal chance of knowing someone in $e$.
3. Inaccurate recall. People don’t recall accurately the number of people they know in the subpopulations we ask them about.
   - The accuracy problem is discussed earlier.

Facilitator Notes: Transmission effects show up when people do not know that they know something about their friends, family members, and acquaintances. For example, you may know someone you work with every day and not know that she is a member of a particular church or that she is suffering from some chronic illness, and so on. To learn these facts about someone requires that they tell you. Not knowing things about people you know because those people don’t tell you is an information transmission problem. A lot of information about people, however, is blocked by social and physical barriers (more about this in the next slides). And even when people know things about their network members, they may not dredge up the information when a survey researcher asks about it. Thus, asking “How many people do you know named Michael?” may result in an underestimate because of poor recall or an overestimate because of rounding (that is, people saying to themselves, “well, I can count three Michaels who I know and I must know more than that, so I’ll just say five in answer to the question.”

Network physical barriers

- Oklahomans know more Native Americans than Floridians do.

Facilitator Notes: There are physical and social barriers to knowing people in various populations. There are more American Indians in Oklahoma than there are in Florida. We expect that people in Oklahoma will, therefore, know more American Indians, on average, than people in Florida do. The next graph shows this.
Correlation between the mean number of Native Americans known and the percent of the state population that is Native American is 0.58, \( p = 0.0001 \).

Facilitator Notes: From our surveys, across the U.S., the correlation between the average number of Native Americans known and the percent of the population that is Native American is 0.58. This is a statistically very significant correlation.

Network social barriers

- Race (Black may know more diabetics than White people do.)
- Gender (Men may know more gun dealers than women do.)
- Even first names are associated with the barrier effect.
- We address the barrier effect by using a random, nationally representative sample of respondents.
- However, using the method on specific populations may still lead to incorrect estimates.

Facilitator Notes: Besides physical barriers, there are social barriers to knowing people in various populations. Black people are more likely to be diabetic than are white people and so we expect that black people are more likely to know a diabetic than are white people. Even first names are subject to barrier effects. People in California are more likely to know someone named Carmen than are people in states that have a low fraction of people with Hispanic names. Over time, with many applications of the network scale-up method, we should learn more about barrier effects and continue to improve the accuracy of the network scale-up method.
The transmission effect

- Researchers study transmission bias by asking people why they do or do not tell their network members various things about themselves.
- The researchers recruited 30 people who were members of one of the known populations used in the network scale-up method.

Facilitator Notes: Transmission bias comes from the fact that some things more difficult than other to know about our acquaintances. Stigmatizing information (for example, a suicide in the family's history) is not shared with everyone one knows. Some things, like having a twin sibling or one's blood type may simply never come up in conversation.

Interview egos and alters

- In the interview, they randomly selected male and female first names proportionate to their representation in the 1990 US census.
- For each of 25 hits, the respondent provided some information about the alter, including the alter's phone number. Total=30x25=750.
  - They contacted 220 of 750 named alters and asked them things about themselves and about ego.

Facilitator Notes: To test the transmission effect, we asked 30 people if they knew someone named Michael, someone named Nicole, and so on. We chose the names so we'd have a good distribution—some common names, some less common, and some uncommon. We kept asking people about first names until we got a hit—that is, the respondent said that he or she knew someone with that first name. We did that until we got 25 people in each respondent's network. Then we asked the respondent for some information about each of the 25 network alters. With 30 respondents and 25 alters for each respondent, there were 750 alters. We were able to contact 220 of the 750 and ask them things about themselves—things that we had asked the original respondent about them. The next slide shows a summary of results.
Facilitator Notes: Everyone knows that their network member is an American Indian, but just over half know that a given network member is diabetic. Similarly, about 58 percent know that a network member is in an organization called the JCs. These transmission errors may have different causes. Diabetics may not want to discuss their illness with friends and co-workers, while friends and co-workers may be uninterested in the organizational memberships of the people they know.

Findings from the alter study

- It is much easier to know that someone is a kidney dialysis patient than it is to know that they are a diabetic

- Diabetes is much less visible.
Some things are easy to get right

- 99% know their alters’ marital status.
- Egos know how many children 89% of their alters have.
- 98% know the employment status of their alters.

Facilitator Notes: 99% of respondents in this little study knew the marital status of their network members. Marital status is something that comes up in conversation in most social situations.

Some things are harder to know

- When asked about the number of siblings the alter has, egos say they don’t know 52% of the time.
- Egos say they know the state in which 70% of their alters were born, but only 57% of the reports (ego’s and alter’s) agree on this.

Facilitator Notes: These traits, like being in the JCs, are things that don’t often come up in conversation with casual acquaintances. My co-workers may not know I have siblings, or the state I am from, but as relationships progress, these pieces of information come up as people get to one another.
Some people withdraw

- Gene Shelley found that people who are HIV+ withdraw from their network in order to limit the number of people who know their HIV status.
- Eugene Johnsen confirmed this by showing that HIV+ people have, on average, networks that are one-third the global average.


Facilitator Notes: One source of transmission bias is the fact that people in stigmatized populations withdraw from interacting with people in their networks. You may know someone who is HIV-positive who has ceased interacting with you entirely because he or she wants to limit their circle of friends and acquaintances. This was found first in ethnographic research, by Shelley, and was later confirmed mathematically, by Johnsen.

Can we account for these errors?

- Can we use this kind of information to tweak the model?
- We tried to develop weightings for classes of characteristics about subpopulations … classes like “things that carry a strong stigma” and “things that carry a moderate stigma” and “things that just don’t come up in conversation.”
- While we found some signals like these, we don’t know how to know whether two populations require the same weighting.

Facilitator Notes: So far, we have not been able to improve the model using the information we’ve collected about transmission and barrier effects. We turned to modeling the errors, but with more empirical tests of the model we expect further improvements.
A theory of transmission bias

- Assume that people report correctly what they know.

- The comparison of the data from clergy and others shows that whatever the errors are, they are consistent.

Facilitator Notes: Shortly before his untimely death, Peter Killworth proposed the following: (1) Instead of assuming that people are inaccurate, assume that they report correctly what they know. After all, whatever the errors are in the model’s predictions, we see from slide 16 – which we repeat in the next slide – that those errors are consistent.

- They tried correcting inaccuracy empirically by changing the way they collect and adjust data, but this did not produce good results.

- Can we make adjustments using a model?

Facilitator Notes: To do this, we asked: Can we adjust the model itself so that its predictions are more accurate?
A theory of transmission bias

- Assume that people report correctly what they know.

- The comparison of the data from clergy and others shows that whatever the errors are, they are consistent.

Facilitator Notes: Shortly before his untimely death, Peter Killworth proposed the following: (1) Instead of assuming that people are inaccurate, assume that they report correctly what they know. After all, whatever the errors are in the model’s predictions, we see from slide 16 – which we repeat in the next slide – that those errors are consistent.

Facilitator Notes: Whatever the errors are in the model’s predictions, we see from this slide that those errors are consistent.
Facilitator Notes: Whatever the errors are in the model’s predictions, we see from this slide that those errors are consistent.

Most Americans know a Christopher

- It’s likely that you know at least one Christopher
  - That is, the probability of knowing NO Christophers is close to zero.
- Twins are likely to be underreported.
- But what’s the truth? How can we draw the curve on that jagged diagram so that the true values are represented?

Facilitator Notes: From this graph, we see that Americans are very likely to know at least one person named Christopher. We also see that twins are probably underreported. The population of twins is very large (about 1 in 125 births), but about 30% of Americans reported in our surveys that they did not know anyone who has a twin. The problem is, we don’t know what the truth is. We’d like to be able to re-draw the graph in slide 43 so that the true values were represented, not just what people report.
Suppose people report accurately

- In other words, given the structure of that diagram, they decided to trust their informants and assume that they are reporting correctly what they know.
- It's just that what they know is incorrect.
- That jaggedy curve doesn't tell us where the curve would be if people responded honestly to correct information instead of honestly to incorrect information.

To do this, instead of assuming inaccurate informants, suppose we assume that people are accurate in their reporting. It's just that what they know is incorrect.

---

This means adjusting the $x$-axis rather than the $y$-axis

- Suppose that widows don’t tell half the people they know about their being a widow.
- The 0.13 on the $x$-axis would remain the same but the number that people would be responding to would be $0.13/2$.
- To make the $x$-axis the effective size of that population, we would slide it to the left while the $y$-axis would remain the same.

Facilitator Notes: Widows are 0.13 of the population in the U.S. … Suppose that widows only tell half the people they know that they are widows. Then, some people who report that they don’t know any widows would be incorrect, but would still be reporting correctly what they know. To adjust for this, we would slide the $x$-axis in the graph to the left while keeping the $y$-axis the same.
- Of course, we have no idea what the transmission error might be – that’s what we tried in vain to get with weightings.

- We only know that if the numbers remain the same on the y-axis and we make up the effective sizes on the x-axis, the jaggedy line would go.

Facilitator Notes: How big an adjustment should we make to the x-axis? We don’t know – that’s what we could not find out with any of the weightings.

- Killworth did this analytically by satisfying certain mathematical properties

- We know the probability of knowing none and also of knowing just one person in a subpopulation. These have to be related mathematically, which leads to a well-defined set of values for the effective subpopulation

- We can then compute the predicted distribution of c. This next diagram shows that we may be on the right track.
Returning to using this to scale up

- Seroprevalence: 800,000 ± 43,000;
- Homeless: 526,000 ± 35,000;
- Women raped in the last 12 months: 194,000 ± 21,000.
  - These are all close to other estimates made with various enumeration or surveillance methods.

Facilitator Notes: Finally, let’s return to the application of the network scale-up model. In the mid-1990s, when we first tested the model, we estimated the population of people who are HIV-positive, people who are homeless, and women in the U.S. who had been raped in the previous 12 months. This slide shows our estimates for those populations. These estimates were close to those made with other methods.
Next steps

- Each improvement of the model produces incremental improvement in estimating population sizes.
- The key is collaborative effort among modelers, survey researchers, and ethnographers.

Facilitator Notes: We can’t claim that the network scale-up method produces the most accurate estimates of hard-to-count and uncountable populations. However, as our knowledge improves about transmission and barrier effects, the estimates improve – and by a known amount for the known-size populations. As the model’s predictions get better for known-size populations, we can have more confidence in its predictions of hard-to-count and uncountable populations.
Page intentionally left blank
Session 8

Network Scale-Up
Doing it in Practice
Session 8: Presentation

Estimating Sizes of Populations at High Risk for HIV: Implementing Network Scale-Up Methods

Donna F. Stroup, PhD, MSc

Partial support for this work provided by UNAIDS

Bangkok, July 2009
Learning Objectives

- Understand basic concepts of network-scale up methods in estimating population size
- Understand two methods for calculating personal network size (known population and summation)
- Critically evaluate NSU for use in population size estimation
- Identify areas for research or practical evaluation

Borrowing from Sociologists

- Socially useful reasons for knowing about personal networks
- P Killworth, E Johnson, C McCarty, G Shelley, R Bernard: see http://tinyurl.com/mf5uwa
- Requires access only to general population
- Intuitive assumption: people’s social networks are, on average, representative of the general population.
Designing a Network Scale-up Survey

- Define the respondent population – not the same as MARP

- Choose a sampling frame

- Choose a survey mode

- Decide on sample size

- Design questionnaire – here we deviate from usual surveys

Choosing the Respondent Population

- Respondent population should contain the MARP of interest
  - US Population: to estimate homeless
  - Urban Population: to estimate IDUs

- Size of respondent population must be known from other sources

- Transmission effects

- Barrier effects
Defining the Sampling Frame

- Sampling frame = list of respondent population
- US: random digit dial telephone numbers
- Census or voter registration records

Choosing a Survey Mode

- Your choices:
  - face-to-face
  - telephone
  - mail
  - drop and collect
  - web
- Effect of mode on responses
- For HIV/AIDS, face-to-face makes sense; or ballot-box
Determining Sample Size

- As always, related to expected margin of error
- Most applications have used network size
- Estimates of network size are stable (in a population) but large (or unknown) standard error.
- In US,
  
<table>
<thead>
<tr>
<th>Survey of</th>
<th>Margin of error (alters):</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>± 26</td>
</tr>
<tr>
<td>1,000</td>
<td>± 16</td>
</tr>
</tbody>
</table>

Designing the Questionnaire

- Definition: who is a network alter
- Three Sections
  - Demographics (used to estimate bias)
  - Question(s) to estimate number of alters respondent knows in target population
  - Question(s) to estimate respondent's network size (c)
Alter Boundaries

- Who is an alter?

- Boundary of 12 months will produce very different estimate from boundary of 2 years

- One approach:
  - you know them and they know you
  - by sight or sound
  - you have had contact within the past 2 years
  - could contact them if you wanted to

A Bit of Notation

- $t = \text{size of general population (T)}$

- $e = \text{size of target (at risk) population (E) (?)}$

- $P_r = \text{the proportion of respondents reporting that they know a member of E}$
Network Scale-up

- Suppose 1/100th of the population have some characteristic
- On average, assume 1/100th of any network to possess that characteristic also
- If $e$ represents the size of $E$, $t$ is the size of $T$, $c$ is the number of persons known by each respondent, and $m$ is the mean (average) number of persons of $E$ known by each respondent, then $m/c = e/t$.

Estimating c: Personal Network Size

- Scale up from known populations
- Summation method
Using Known Populations

- Select several population of known size (the more the better)
- Known populations should have approximately same demographics as respondent population
- Populations should vary in size/type:
  - all health conditions => barrier effect
  - only large populations => large estimation error
  - only small populations => bias from few “hits”
  - current thinking: 0.1% to 4% of total population

US Known Populations

- Existing data:
  - Census
  - Crime Statistics
  - Government Surveys
- Recurring collection in subsequent years
- All data should reflect same time period; be aware of population lags
- Susceptible to transmission and barrier effect
Demographics and Number Known

<table>
<thead>
<tr>
<th>Population</th>
<th>State</th>
<th>Sex</th>
<th>Race</th>
<th>Age</th>
<th>Education</th>
<th>Marital status</th>
<th>Work status</th>
<th>Religion</th>
<th>Political Party</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native Americans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gave birth in past 12 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adopted a child in past year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widow(er) under 65 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On kidney dialysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postal worker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commercial pilot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member of Jaycees</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opened a business in year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have a twin brother or sister</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Licensed gun dealer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Came down with AIDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males in prison</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homicide victim in past year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suicide in past year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Died in vehicle in past year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women raped in past year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homeless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Some Experience with Names

- Generally collected in Census
- Variation in size of groups
- Typically ascribed
- Countries and cultures vary in use of names
- Prone to barrier effect
Names and Demographics

<table>
<thead>
<tr>
<th>Population</th>
<th>State</th>
<th>Sex</th>
<th>Race</th>
<th>Age</th>
<th>Education</th>
<th>Marital status</th>
<th>Work status</th>
<th>Religion</th>
<th>Political Party</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christina</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christopher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jacqueline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>James</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jennifer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anthony</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kimberly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stephanie</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>David</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicole</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summation Method for Estimating c

- Ask people directly how many people they know – try it!
- Manageable sub-tasks
- Culturally relevant categories, mutually exclusive and exhaustive
- Categories should be small enough so that estimation is reliable.
Example: Summation Method

- Immediate family
- other birth family
- Family of spouse/partner
- Co-workers
- Other people at work
- Best friends
- People known through hobbies/recreation
- People known through ... (religious organizations, neighborhood, school)
- People known through others
- Childhood relations
- People who provide a service

Developing Categories for Summation Method

- Culturally relative

- Mutually exclusive, exhaustive, small enough for accurate counting

- Research needed
  - cultural consensus analysis
  - large culturally relevant categories and split
Developing Categories for Summation Method

- Culturally relative
- Mutually exclusive, exhaustive, small enough for accurate counting
- Research needed
  - cultural consensus analysis
  - large culturally relevant categories and split

Distribution of $c$, network size*

$Avg = 290$
$Median = 231$
$sd = 232$

*Source: Killworth, et al. 1998
Advantages of Summation

- quicker than using known populations
- less problem with transmission or barrier effect
- Does not require finding the populations

Disadvantages of Summation

- lack of statistical basis
- subject to double-counting (e.g., co-worker and social contact)
- lack of ability to verify: In using known populations, network size can be verified by back calculating
Test for 29 networks of known size


**Test for 29 networks of known size**

Twins, diabetics

$R = .79, .94$

without outliers

- There is a tendency for people to overestimate small populations (<2 million) and underestimate large ones (>3 million).
- The two largest populations are people who have a twin brother or sister and diabetics.
- Without these two outliers, the correlation rises from $r = .79$ to $r = .94$.
Summary Network Scale-up

Two Steps:

1. Estimate network size, c

2. Use c with respondent's estimates of unknown populations to scale-up to e, unknown size of subgroup at risk in population

Step 1a: estimating c, summation method

1. Get $c_i$ for each respondent by adding up their estimates for each relation category

2. Average these $c_i$ estimates over all respondents to get c
Step 1b: estimating $c$, known population method

- $t$: size of population to which you are scaling up; same for each respondent
- $e$: sum of known populations you are using; same for each respondent
- $m_i$: sum of all reported known population sizes for respondent $i$
- $c_i = \frac{m_i \times t}{e}$
- $c =$ average $(c_i)$

Where does the 290 come from?

$$c_i = t \times \frac{\sum m_{ij}}{\sum e_j}$$

- $L$ known subpopulations ($j$). Here $i$ is the respondent, who knows $m_i$ people in subpopulation $j$, of known sizes (e.g., diabetics, postal workers, members named Zaid)

- Network size is (the sum of all the people you say you know in some subpopulations of known sizes, divided by the total size of those subpopulations) times the population within which the subpopulations are embedded.
Step 2: estimating unknown size of MARP

- $t$: size of population to which you are scaling up; same for each respondent
- $c$ (from either 1a or 1b)
- $m$: average of respondents' estimates of the number of people known in subpopulation
- Size of unknown subpopulation $= e = (m/c) \times t$

Network Scale-up: Strengths

- Does not require contact with hidden population
- Can be imbedded into national surveys (short)
- Can produce sizes for several hidden populations at once
- Reality checks are possible (with known populations)
Network Scale-up: Issues

- Estimates of network size are reliable but not absolutely precise
- Underrepresentation of people who know members of the hidden population
- Transmission effects: Everyone knows everything about everyone they know.
- Barrier effects: Everyone in $t$ has an equal chance of knowing someone in $e$.
- Recall effects: People recall accurately the number of people they know in the subpopulations we ask them about.
- Cultural effects: Most experience in U. S.

References: Network Scale-up

http://nersp.nerdc.ufl.edu/~ufruss/scale-up.htm

Papers from the N-SUM project: the Network Scale-Up Method
- Estimating the Size of Event Populations II (Bernard et al., 1989)
- Estimating the Size of Personal Networks (Killworth et al., 1990)
- Estimating the Size of Event Populations I (Bernard et al., 1991)
- Who Knows Your HIV status? I (Shelley et al., 1995)
- ...

Powerpoints:
- Background and theory of the network scale-up project
- How to conduct a network scale-up survey

Other papers related to the network scale-up method:
- How many people do you know in prison? (Zheng, Salganik, and Gelman 2006)
- Fighting a Hydra: A Note on the Network Embeddedness of the War on Terror (Moody 2006)

Excel sheets for calculating network scale-ups:
- 1. Estimating the size of an uncountable population using the back estimation method
- 2. Estimating the size of an uncountable population using the summation method


Questions?

Data for Solutions, Inc
Analytic solutions for practical problems
www.datafor solutions.com
Session 9

Assessing Bias and Reconciling Multiple Size Estimates
Session 9 Summary - Assessing Bias and Reconciling Multiple Size Estimates

Content:

This session introduces the concepts of assessing bias and reconciling multiple estimates and it prepares participants for the exercises covering these two topics.

Objectives:

- Introduce the concept of how to move from local to national size estimates, and the importance of assessing bias and reconciling multiple estimates.
- Prepare participants for the two exercises covering these two topics.

Specifics of what is covered in the session/tips for facilitators:

This session starts by recapping what has been covered in the workshop thus far and reviews the various methods that can be used to generate local size estimates. It then introduces the participants to the process of using the local size estimates for more macro-level purposes, such as program planning (at the local level), or estimates and projections and geographic prioritization at the national level. Emphasis is placed on the need to assess biases, generate hypotheses about what could account for differences in local level estimates (for example, if mapping exercise and survey with multiplier give widely different estimates), and then making adjustments and developing plausible ranges, using the information at-hand. These concepts are then explored much more in-depth in the two exercises, which allow participants to learn by attempting to go through the process using real data from the region. Facilitators should treat the exercises, and particularly the answers, as additional training materials, since many of the lessons of this session are built-in to the exercise.
Session 9: Presentation

Regional Training on Size Estimation of At-Risk Populations in Asia-Pacific

Assessing bias and reconciling multiple size estimates
Objective of this session

- Introduce the concept of how to move from local to national size estimates, and the importance of assessing bias and reconciling multiple estimates
- Prepare participants for the two exercises covering these two topics

Facilitator notes:

- This session is designed to prepare participants for the exercises on assessing bias and reconciling multiple estimates
- It introduces the concept of the importance of assessing bias reconciling local estimates
Facilitator’s Notes:

- At this point in our workshop, we have reviewed several of the main methods for local size estimation, including mapping, survey-based multipliers, and capture-recapture.
- We have also seen how implementation problems can lead to biased data sources that can cause problems with the estimates. Some common examples of this that we have already seen in the sessions on multiplier methods, relate to what can happen when two data sources being used for a size estimate do not match, or are not independent of one another.
- Next we are going to explore ways of systematically assessing biases and reconciling multiple estimates, using a “hands-on” approach involving real data from the region.
Why do we need to reconcile differences?

- What happens when we are ready to move from local to national level size estimates?
- What happens when we need to use local size estimate to make programmatic decisions?
- What happens when there are multiple estimates available in the same location?

Facilitator notes:

- What happens when we are ready to move from local to national level size estimates?
- What happens when we need to use local size estimate to make programmatic decisions?
- What happens when there are multiple estimates available in the same location?

Local level size estimates are the building blocks of national estimates. Without them it is difficult even to have crude estimates of what is going on in the country. However, as we have seen, when we implement methods for developing local size estimates, things frequently do not go as planned, and we find that the data that has been generated is of questionable validity.

Sometimes there are multiple estimates with widely varying results, so it is difficult to know which ones to use, for any purposes (be they local or national).
Facilitator notes:

In this diagram of the process of moving from local to national level estimates (seen earlier in the workshop), we can see where the process of assessing data and reconciling differences fits into the overall process.

(Facilitator can walk the participants through the flowchart again, and point out that assessing bias and reconciling multiple estimates is a process that should happen before feeding those numbers into a national estimate).
Facilitator notes:

- In places that have local estimates, before using those estimates for any purpose, including for estimates at the national level, it is important to first assess the quality of the data and look for biases.
- Where there are multiple estimates, these need to be reconciled so that they can be used for more macro level purposes, (e.g. developing national size estimates, and undertaking geographic prioritization)
Group Exercise 1

Estimation the Size of the Female Sex Worker Population in Guntur

Assessing Bias
Group Exercise 1 - Estimating the Size of the Female Sex Worker Population in Guntur: Assessing Bias

**Background:** This example comes from the Avahan program in India. Andhra Pradesh is one of the States in India that has been hit the hardest by HIV. There is an active sex trade and the Avahan program implements interventions with female sex workers in several districts of the state. Guntur is one such district in Andhra Pradesh, with more than 4 million people, two major cities, and covering more than 11,000 square kilometers.

**There are two main types of size estimation data available for Guntur:**

1. Multiplier method data through the IBBS survey: The IBBS was conducted among sex workers in Guntur district. The survey population definition was: Women who sold sex for money in the last month and who solicited in sites included in the sampling frame (brothels, street and home). The sample covered the whole district. The IBBA survey had a sample size of 400.

2. Programme mapping data: The Avahan programme mapped sex workers as part of the pre-intervention activities. They estimated that there were 4,271 female sex workers in the district.

**Size Estimate using unique object multiplier**

- A unique object (key chain) was distributed to members of the community in advance of the survey. The team reported that checks were done confirming that key chains had been distributed to the target population members, and that the key chains were distributed widely throughout the district. Five hundred key chains were distributed.
- The IBBA survey found that 42.5% of the respondents reported having received a key chain.

**Size Estimate using registration data multiplier**

- The NGO from the Avahan programme in Guntur reported that there were 1000 sex workers registered in their program
- The IBBA survey found that 84.8% of sex workers in Guntur responded “Yes” when asked “Are you registered with any of these NGOs” (referring to NGOs that had been named in the previous question).

**Part A: Calculating the Size Estimates from different methods**

A1. What is size estimate for the unique object multiplier
A2. What is size estimate for the registration data multiplier

**Part B:** The two multiplier methods produced similar estimates (1176 and 1179). However, both estimates were significantly lower than the number mapped (4271). In addition, the program reported 1000 registered sex workers, and the proportion reporting in the IBBA that they were registered in the program was (84.8%).

What types of bias in the each of the different data sources do you think might be at play?
**B1: Mapping data**

**B2: Program-based multiplier**

**B3: Unique-object multiplier**

For each of the data sources, describe the evidence that the estimate from that source is too high or too low. Explain your answers and discuss what additional information would be helpful in clarifying the situation.

**Bonus Questions**

# 1: Suppose you learn that the key chain was distributed by FSW peer educators (PE) of the programme, and the PE gave the objects only to FSW who were their friends in Guntur. How do you think this might have influenced the size estimate based on the unique object multiplier? Support your answer.

# 2: Suppose you learn that the unique objects were distributed only in the areas close to the NGO, and there was a large portion of the district where no sex workers received unique objects. How do you think this might have influenced the size estimate? Support your answer.
Answers

Part A
A1: $500/0.425 = 1176$
A2: $1000/0.848 = 1179$

Part B
B1: Mapping based estimate = 4271

Evidence of bias

- The mapped number (4271) is higher than the registered number (1000 registered with the program)

- IBBS indicates that the majority of the population (84.8%) is registered with the program. This would suggest that the 1000 registered sex workers represent about 85% of all sex workers in the district, which, according to the program-based multiplier estimate are about 1179. If the IBBS in fact covered the entire district (which is what we are told), and if the IBBS was truly representative, this would support the hypothesis that the mapped number is too high.

Hypothesis: Mapped number is too high

Some reasons why mapped numbers can be overestimated

- Population is very mobile resulting in double counting
- If NGOs are involved in the mapping, numbers can sometimes be inflated to increase funding possibilities

Additional information that would be helpful in clarifying the situation

- What proportion of sex workers in the district are targeted by the NGO? This might help explain the "mismatch" between the 1000 reached and the > 4000 mapped.
- Degree of mobility in the population. If the population was highly mobile, there may have been double-counting during the mapping, causing the mapping estimated to be too high.

B2: Program-based multiplier estimate = 1179

Evidence of bias

- The mapped number (4271) is much higher than the program-based multiplier estimate (1179)
Hypothesis: Program-based multiplier estimate is too low

What can cause the program-based number to be an underestimate?

☐ The NGO may have failed to record all the FSWs who are registered. In this case the program based multiplier of “1000” people registered would be too low, which would cause the overall number also to be too low.

☐ In the flipside, survey questionnaires must be specifically matched to programme-based multipliers in order to produce valid estimates. If there was a “mismatch” between the way the programme defined “registration” and the way the questionnaire elicited information on registration from respondents, this could be problematic. Specifically in this case, if respondents reported being registered, when they actually were not, this could have led to an underestimate.

☐ If people who were registered with the program were more likely to participate in the survey than people who were registered, this could have also caused an underestimate. In this case, the assumption of independence between the two data sources would have been violated. This is a common problem in surveys of high-risk populations.

Additional information that would be helpful in clarifying the situation

☐ It would be important to know how the NGO/programme defined registration, and how the community targeted by the survey defined it. In other words, was there a “common” understanding of what it meant to be registered?

☐ It would also be important to examine the questionnaire closely to make sure it was possible for respondents to distinguish which program was being referred to (e.g. that the “branding” was clear enough).

☐ Finally, it would be important to talk to the survey implementers to understand the likelihood that people who were registered with the NGO had been more likely to participate in the survey. This can happen in a variety of ways, e.g. if the sites frame was incomplete (e.g. included mostly NGO sites, or included only a part of the district), if NGOs were involved in helping to identify eligible respondents or if people who were in contact with the NGOs were more comfortable participating in the survey (i.e. less likely to refuse participation).

B3: Unique object multiplier estimate = 1176

Evidence of bias

The mapped number (4271) is much higher than the program-based multiplier estimate (1176)

Hypothesis: Program-based multiplier estimate is too low

What may have caused the unique object multiplier estimate to be an underestimate?
If the object was not sufficiently unique, so that respondents reported receiving it when actually they did not, this could have caused an overestimate.

Likewise, if people who received the object were more likely to participate in the survey than people who did not receive the object, this could have resulted in an underestimate. This situation could arise if program people were involved in distributing the object and if people in the program were more likely to participate in the survey (for similar reasons as described in the answer to B2). In this case, we can say that the survey was not completely random, and also that the assumption of independence between the two data sources is violated.

Additional information that would be helpful in clarifying the situation

To assess these possible biases, it would be important to know how respondents were selected for the survey, and whether it seems likely that people in contact with the NGO were more likely to participate (e.g. if NGO supplied list of sites for the survey, or if people exposed to the program were less likely to refuse participation in the survey).

It would also be important to know something about how the unique objects were distributed, and if it was done in a way that relates to the possible bias in the survey (i.e. if people who were part of the program were BOTH more likely to have received the unique object, and likely to have participated in the survey, this is a violation of the assumption of independence which leads to an underestimate)

Answers to Bonus Questions

# 1 It does not matter who the unique object is given to, as long as people who received the object are no more or less likely to participate in the survey than people who did not receive the object). But if the same people who receive the object are also more likely to participate in the survey, this can lead to an underestimate, which stems from the violation of the assumption of independence.

# 2 It is not necessary for all respondents to have an equal chance of receiving the unique object. As long as people who receive the object are no more or less likely to participate in the survey, the estimate should not be biased. So, even if objects are distributed only in locations near to the NGO, this should not be a problem, as long as the survey is not similarly limited to areas close to the NGO. In other words, the survey itself should not be biased, by sampling only from areas close to the NGO.

However, this raises another issue, which is that the probability of population members to receive an object prior to the survey should not be too small, otherwise the precision of the estimate may be affected. This situation could potentially arise if objects are distributed only near the NGO, and the number receiving the object is small relative to the total size of the population being estimated. If this situation were to happen, we might say “in statistical terms” that the estimate would not be “biased”, (as long as the survey was random), but the precision might be low (if the ratio of the population size to the number of objects distributed is high). When planning how many objects to distribute, a rule of thumb is that it should not be less than 10% of the estimated population size.
Group Exercise 2

Triangulation of Multiple Size Estimate Methods: How many MSW are in Madurai
Group Exercise 2 - Triangulation of Multiple Size Estimate Methods: How many MSW are in Madurai?

**Background**

Madurai is a large district in Tamil Nadu State in India with a population of more than 2.5 million people. The Avahan India AIDS Initiative runs an intervention for both MSW and FSW in Madurai. The programme was interested in obtaining a size estimate of the MSW population to serve as a denominator for intervention coverage.

Data from a recent analysis of different size estimates from the Avahan programme were published in a recent paper (cite). Three methods of size estimation and their results are presented below:

<table>
<thead>
<tr>
<th>Size estimate method</th>
<th>Estimated Size</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programme-led geographic mapping</td>
<td>1400</td>
<td></td>
</tr>
<tr>
<td>Programme data (registrations) multiplier in IBBS</td>
<td>1763</td>
<td>1527-2098</td>
</tr>
<tr>
<td>Unique object multiplier in IBBS</td>
<td>3676</td>
<td>2747-5102</td>
</tr>
</tbody>
</table>

The two multiplier method estimates came from a probability survey of MSW in Madurai conducted between May-June 2006. The survey was designed to get a representative sample of MSW in the district.

The unique object was a special small bag (where sex workers could carry ID cards and condoms) that was produced for the survey. In each district, 1000 small bags were distributed to the MSW community through the peer educators of the programme.

The programme data multiplier was based on the number of MSW who were actively registered with the programme during the period February-June 2006, which is inclusive of the three month period just prior to the start and the duration of the survey conducted among MSW in Madurai district. The survey instrument asked two questions related to this multiplier:

- Are you registered with the TAI NGO programme (the name of the NGO and the logo was shown to the respondents)?
- When were you registered with the TAI NGO programme?

The registration data came from an electronic data base of routine monitoring data that can pull out various statistics, including the number of people who register during a specific time period. The registration process takes place when the programme makes contact with an MSW in the programme area and is the first step to a person accessing services (either outreach or at the drop in center/clinic) with the programme. At registration, some basic information is gathered about the individual and the person receives an id number.

**Part A** As found in the earlier exercise, one common problem with multiplier methods from survey data is that the survey may not be representative of the entire group for which the size is being estimated. Often times, surveys are biased toward including people who have been in regular contact with the programme.
How would you adjust your interpretation of the size estimate given by the programme registration data multiplier, if you had evidence to suggest that the survey over-represented people who were in the intervention?

**Part B** The total number of MSW registered in the intervention during the period February-June 2006 was 1093 people and the number of MSW identified by the program was 1400 (time period not specified). One interpretation might be that 1093/1400 or 78% of the identified population are currently active in the intervention. However, in the IBBS, only 62% reported being registered with the program.

What could explain the difference between these two figures?

**Part C**

During further exploration of the use of the multipliers for this district, a number of implementation issues were identified.

**Unique Object**

1. It turns out that for the unique object, the distribution process was not closely monitored and undistributed objects were not collected back.
2. Distribution was supposed to have taken place before the survey started, but most of the objects were actually distributed during the first 2 weeks of the survey.

**IBBS Survey**

3. During the survey, respondents were asked if they received the unique object, but the object itself was not shown to the respondent to confirm that they knew what object was being described.

**Program Multiplier**

4. Upon further discussion with the NGO it was learned that Individuals who have not accessed project services in the past 3 months are removed from the current register, but the MSW may not realize that they are no longer “currently registered.” And the survey question asked respondents if they were registered, not clarifying if this meant currently or ever registered.

For each of these issues, indicate whether the resulting estimate would have under or over estimated the actual size. Consider what might be plausible levels of bias that may have occurred and develop a lower or upper limit on the change in the size estimate.

**Part D** In summary, how would you rate what is plausible given the information available and the analysis you have done in this exercise?

<table>
<thead>
<tr>
<th></th>
<th>Plausible on a scale of 1 (unlikely) to 5 (very likely)</th>
</tr>
</thead>
<tbody>
<tr>
<td>That the actual size is &lt;1400</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>
That the actual size is between 1400-1700
That the actual size is between 1700-3600
That the actual size is >3600

D1: And what would you report as the likely size estimate for MSW in Madurai?

D2: What are the key questions you would ask to make a firmer conclusion about the actual size?

**Part E** If you were repeating the mapping exercise, the unique object, or the programme multiplier what would you do differently, or what information would you make sure was collected as part of implementing each size estimation method?

E1: For mapping exercise:

E2: For conducting the survey in general:

E3: For the unique object multiplier:

E4: For the programme multiplier:
Bonus Exercise:

When the programme was in the planning phase for the survey and looking at potential programme based multipliers, they had a number of options to choose from. (See table below)

<table>
<thead>
<tr>
<th>Programme Monitoring Data:</th>
<th>MSW</th>
<th>FSW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number registered between Feb 2006-June 2006</td>
<td>1093</td>
<td>5174</td>
</tr>
<tr>
<td>Number contacted by outreach worker in 1:1 session in the month</td>
<td>3861</td>
<td>16165</td>
</tr>
<tr>
<td>Number contacted by peer educator in 1:1 session in the month</td>
<td>3434</td>
<td>13027</td>
</tr>
<tr>
<td>Number received a project health (id) card in the month</td>
<td>63</td>
<td>502</td>
</tr>
<tr>
<td>Number who came to clinic in the last 3 months</td>
<td>615</td>
<td>2260</td>
</tr>
<tr>
<td>Number who have seen a condom demonstration in the month</td>
<td>5132</td>
<td>15138</td>
</tr>
</tbody>
</table>

Before deciding whether you use any of these monitoring data, what are some of the key questions you would have about the data? And do you think registration was the best choice among these to use? Why or why not?
Answers

Part A
In theory, over-representation of intervened people in the survey would indicate that the estimate given by the programme multiplier is an underestimate of the actual size of the population, i.e. programme multiplier size estimate < actual size of the population

Part B
Any of the three numbers (mapping-based estimates, program counts, or proportion reporting being registered in the IBBS) could be biased, or all three could be biased. Since we don’t know which one is “wrong”, we have to explore all three. Only after assessing the possible biases will it be possible to make a more informed judgement about which figures are more accurate.

B1: Mapping-based estimate=1400

Evidence of bias
- 78% registration (1093/1400) is higher than the 62% registration measured by the survey. So this raises questions about whether mapped estimates are biases.

Hypothesis: Mapped number is too low
Some reasons why mapped numbers could have been too low
- The mapping happened early before the program scaled-up
- Some MSW are hidden and therefore were not counted in the mapping
- Only part of the district was included in the mapping (e.g. areas with interventions or larger urban areas)

Additional information that would be helpful in clarifying the situation
- Contact NGO to find out dates of mapping relative to the dates of programme scale-up
- Contact members of the MSW community to find out how the mapping was done
- Consult NGOs, MSW community and project documents to find out what areas in the district were included and how areas were chosen for mapping

B2: Survey-Based Estimate (1763)

Evidence of bias
- 78% registration (1093/1400) is higher than the 62% registration measured by the survey.
Hypothesis: Proportion reporting registration in the IBBS is too low (ultimately resulting in an inflated size estimate)

Some reasons why the IBBS measure of proportion of respondents registered could have been too low

- The survey sampling frame might have been broader than the area where mapping was done. The survey may have included only intervention areas, whereas the mapping included both intervention and non-intervention areas.
- Some respondents could have “failed” to report being registered, e.g. if registration status was unclear.

Additional information that would be helpful in clarifying the situation

- It would be important to clarify the catchment area of the survey, and see how it compared to the mapping catchment area.
- It would also be important to understand what would trigger respondents to say that they were registered, and whether that matched with what the programme was measuring.

**B3: Program multiplier (1093 people registered in past 3 months)**

Evidence of bias

- 78% registration (1093/1400) is higher than the 62% registration measured by the survey.

Hypothesis: Programme counts of number of people registered too high resulting in a higher proportion of the population appearing to be registered.

Some reasons why program counts could be too high

- People who lose ID cards could be re-registered and therefore counted twice.
- NGO may not have adequate systems for identifying people who are already registered and therefore register some MSW more than once.
- NGO may not have good records of MSW as distinct from MSM and therefore list some MSM as MSWs.
Additional information that would be helpful in clarifying the situation

- It would be important to meet with NGO staff to understand how duplicate registration is avoided, and exactly how MSW are identified and defined. It could be that they are defined differently by the program than by the survey.

Part C

C1: Exact number of objects distributed is not known

1000 bags were supposed to be distributed, however, if fewer were actually distributed, this would have resulted in a lower proportion of respondents receiving the keychain, which would ultimately result in an overestimate of the actual size.

According to the calculation that was used, 1000 bags were distributed, and 27.2% of the population reported receiving a bag in the IBBS. So the resulting size estimate was 3676.

However, supposing that only 75% of the bags had actually been distributed. If that information were known, the calculation would have changed to 750 bags distributed, with 27.2% reporting receiving a bag in the IBBS. So the estimated number would have been 2757.

If you were able to determine the number of bags actually distributed (which would have been known if better monitoring were in place), then a downward adjustment could be made. In this case the adjustment would be by 25%.

Original Calculation: 1000/0.27=3676

Adjusted estimate: 3676 * (1 - 75%) = 3676 * 25% = 2757

C2: Objects distributed during the survey instead of before

If bags were distributed during the survey, that means a proportion of respondents might receive the object after participating in the survey, so the object would be counted as having been distributed, but the respondent reporting not having received the object. The affect would be similar to C1 where the multiplier was too large, resulting in an overestimation of the actual size.

If bags were distributed for the first two weeks of the survey, and the survey had continued for 5 weeks, then a rough calculation would be that up to 40% of respondents received the object after participating in the survey. So this would mean that the estimate would need to be adjusted downward by 40%. The resulting calculation would be.

Original Calculation: 1000/0.27=3676

Adjusted estimate: 3676 * (1- 40%) =2676 * 60% = 2205

C3: Respondents did not accurately report having received the object
In this situation, the proportion who report receiving the bag is too low, so, as with C1 and C2, the result will again be an overestimation of the actual size.

If you are able to ascertain that approximately 10-20% of respondents who received a bag did not report receiving it, then you might decide to lower the estimate by about 15% (following similar logic as in the previous examples).

Original Calculation: 1000/0.27=3676

Adjusted estimate: 3676 * (1- 25%) =2676 * 85% = 3125

C4: Mismatch between program count and question on questionnaire

In this last situation, there is a mismatch between the program multiplier and the survey questionnaire, because the proportion of respondents reporting being registered would be higher than those counted by program, which would be only those who had been active in the past three months. The result of this mismatch would likely cause an underestimate of the actual size.

This would be difficult to adjust because of the difficulty of estimating what proportion of previously registered IBBS respondents might have been dropped from the NGO register for not being active in the past three months. One possibility would be to use the proportion of IBBS respondents who reported both being registered and accessing services in the last 3 months. So instead of using a figure of 62% (who reported being registered), the proportion might drop down to something lower e.g. 50% (who report being registered and accessing services in the last three months). This has the effect of increasing the size estimate as follows:

Original calculation: 1093 / 62% = 1763

New calculation: 1093 / .5 = 2186

Part D

It seems unlikely that the actual size is smaller than the mapped number (1400). We looked at some information in part B that suggested that the mapped number might be too low. We also made some calculations in Part C that suggested that the unique object multiplier may have produced an inflated size estimate, and the program-based multiplier might have produced a deflated estimate. If several biases were occurring at the same time, then it becomes difficult to make an estimate, but establishing a range might be possible.

Given the supporting evidence for the hypotheses that 1400 from mapping is too low, and 1763 from the program-based multiplier is too low, and 3676 from the unique object multiplier is too high, a reasonable range might be something like 2000-3000. But a number of follow-up questions would be important to get answers for, specifically:

- What was the scope of the survey and mapping in terms of geographic coverage of the district?
- What is the likely pattern of distribution of bags?
- What proportion of the bags are likely not to have been distributed?
How were registration numbers cleared of dropouts, and are there data from the survey to assess what proportion of the respondents who reported being registered may have been considered “inactive” according to the NGO records?

Part E

E1: For mapping exercise:

- Ensure the report of the mapping has a clear methods section that describes how hotspots were identified and what geographic areas were covered by the mapping. Clarify how duplicates or double counting are addressed during mapping.

E2: For conducting the survey in general:

- Ensure the survey protocol clearly defines the geographic scope of the survey and the definition of the population included in the sample.
- Document the sampling frame development process including what areas were included and how the sampling frame was constructed.

E3: For the unique object multiplier:

- Provide clear training for persons distributing the object to 1) avoid giving the same individual more than one object, 2) to try to prevent people from passing the object to each other, and 3) to devise a method for counting the actual number of objects distributed and collecting back the undistributed objects.
- Show the object when asking respondents about whether they have received the object, and also make sure the object is sufficiently unique, and memorable.
- Distribute all the object before the survey start.

E4: For the programme multiplier:

- Match the survey question and the programme count by clearly understanding how the programme count is calculated and who is included.
- Include supplemental questions in the survey to assist in testing the reliability of the multiplier question.
Answer to Bonus Exercise

Clarification questions about the programme monitoring data:

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the monitoring system able to count unique individuals, or does it just count contacts?</td>
<td>Yes, the monitoring system can distinguish individuals, through use of the ID number given at registration.</td>
</tr>
<tr>
<td>What is the definition of MSW who are eligible to receive services from the programme?</td>
<td>Aravani (transgender or sari clad men), Kothi (&quot;receptive partner,&quot; dresses as a man, &amp; Double Deckers, (DD) all who sell sex to men.</td>
</tr>
<tr>
<td>Is there any other programme working in the district?</td>
<td>No, the TAI programme is the only intervention for MSW in the district</td>
</tr>
<tr>
<td>Is the TAI logo and name recognized by the MSW?</td>
<td>Yes, the NGO is strongly branded and well recognized by the community</td>
</tr>
</tbody>
</table>

Which multipliers would work well?

<table>
<thead>
<tr>
<th>Multiplier</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration data</td>
<td>May be one of the best choices among these potential multipliers. There are a sizeable number registered during the time period in question. Registration is also a distinct activity, marked by the receipt of an ID number, so it may be easier for survey respondents to recall this event.</td>
</tr>
<tr>
<td>Outreach data</td>
<td>Can be problematic because of difficulties distinguishing between # of individuals contacted, and # of contacts made. Systems for avoiding double-counting are more complicated and not routinely had by many NGOs.</td>
</tr>
<tr>
<td>Condom demonstrations</td>
<td>There may be similar difficulty in counting distinct individuals who have seen a condom demonstration in the month. The numbers observed in the table for outreach, and especially condom demonstrations are also much larger than the mapped size of the population. We would expect the mapping size estimate to reflect the maximum number of people that could be targeted for outreach, so this suggests that there may be a problem with double counting individuals in the numbers reported for outreach/condom demonstrations.</td>
</tr>
<tr>
<td>Clinic visit in the last 3 months</td>
<td>May be a good choice because going to the clinic may be a distinct event for respondents to recall, however, remembering whether their last clinic visit was in the last 3 months may be a bit difficult, as the MSW may be encouraged to get a regular STI check up and so visiting the clinic is not a one time event. It would also be important to confirm whether the clinic can track how many individuals come to the clinic in a three month period, and confirm that this number will not include duplicates.</td>
</tr>
</tbody>
</table>
Getting started (S1)
- Define population to be estimated
- Define geographic areas requiring local size estimate
- Assess each area for availability of local size estimate

Local Estimate Available?

No

Single Estimate-Assess Bias (E1)
Multiple Estimates

Yes

Is estimate valid?

Yes

Local size ready to feed into national estimate

No

Prioritize areas for more data collection based on criteria

Collect size estimation data

Mapping-Based Method (S2)
If done properly, provides robust estimates of "visible" risk populations

Survey-Based Method
- Multiplier (S3-S4-S5)
- Capture-Recapture (S6)

Resource intensive because they require probability sampling but can be added on to planned surveys

Reconcile Estimates (S9 & E2)
- Assess biases
- Select best estimates or develop range

Extrapolation Process Required (S9 & 10)

Asia-Pacific Size Estimation Manual Sessions
(S1) Background and Introduction
(S2) Mapping
(S3) Intro to multiplier
(S4) Program multiplier
(S5) Unique object multiplier
(S6) Capture-Recapture
(S10) Extrapolating for national estimates

Exercises
(E1) – Assessing Bias
(E2) – Reconciling multiple estimates
(E3) – Extrapolation
Session 10

Techniques for Extrapolation to National Level Estimates
Summary – Techniques for Extrapolation to National Level Estimates

Content:

This session describes the process in extrapolating size data from areas where there is local size data to areas where no direct size data is available. The size estimates for local areas can then be summed to obtain a regional or national level estimate of size.

Objectives

- Describe the reasons and types of extrapolation that are commonly used with size estimation of high risk groups
- Make participants aware that extrapolation approaches follow a continuum of simplistic to complex
- Give a real world example of implementation of a complex extrapolation exercise from Indonesia
- Describe one approach for planning and executing an extrapolation exercise for obtaining a national level size estimation process step by step

Specifics of what is covered in the session/tips for facilitators

The session begins with an illustration of what kind of extrapolation may be necessary to go from selected local areas with size estimates to reach an overall national estimate of size. Multiple stages of extrapolation, and options of simple to complex formulas or algorithms are presented. The importance of proxy data is also highlighted, those data which provide information about size, but must be transformed through a formula to result in an actual size estimate. A complex style of extrapolation is presented from Indonesia, which is justified based on the amount of data and the size of the country. The steps to planning and implementing this extrapolation process are laid out.
Session 10: Presentation

Regional Training on Size Estimation of At-Risk Populations in Asia-Pacific

Techniques for extrapolation to national level estimates
Facilitator notes:

The learning objectives of this session are to:

- Describe the reasons and types of extrapolation that are commonly used with size estimation of high risk groups
- Make users aware that extrapolation approaches follow a continuum of simplistic to complex
- Give a real world example of implementation of a complex extrapolation exercise from Indonesia
- 4. Describe a step-by-step approach for planning and executing an extrapolation exercise for obtaining a national level size estimation process

At the end of the session, country teams should be able to review their own data and follow the initial steps of planning what type of extrapolation exercise may be relevant for their situation.
Facilitator notes:

First let’s describe the concept of extrapolation in the context of size estimation, i.e. how data from one area is taken and used to generate estimates in an area without data.

- In this diagram the squares represent small geographic units such as districts that together make up a country. Squares colored in light blue indicate the places where size estimate data have been collected, either using a mapping-based method or a survey-based method. The white squares are districts without this type of data.
- To get a national level estimate, size of the high risk groups in the white square districts must be extrapolated from the size estimate data available in the blue squared districts.
- The diagram also tries to illustrate that even within the districts where size estimation data have been collected, there may be uncovered areas. This occurs because mapping-based methods may be implemented in only the largest towns or urban areas, or selected areas due to limited resources. Similarly for data collected through survey-based methods, because most surveys are limited to selected areas and may not be truly representative of an entire district. For this reason, extrapolation may be necessary from areas with data to uncovered areas in the same district.
Facilitator notes:

In some countries, there may also be a situation in which data are available for the area of interest, but this data may not be the actual size of the HRG. Instead there exists another kind of data which is a proxy for size, i.e. there is a relationship between the proxy for size and the actual size of the HRG of interest.

Some common examples include (Read bullets in slide)…

In each of these examples, the key to using the proxy is to understand the relationship between the data that is available and the size of the HRG of interest. The ideal situation is that this relationship can be expressed as a numerical formula, so the extrapolation from proxy to estimate of size is transparent, rational, and systematic.

Many kinds of data may be available to use as a proxy in this way, the most basic source of data used in this way is the size of the general population. Population data are available in all areas, but the problem is that there are many characteristics about a place which will change the % of the total population who belong to an HRG. So other proxies for size must be used in combination with total population size to create a meaningful extrapolation.
Facilitator notes:

Let’s look at the following commonly used formula for extrapolating the size of different HRG. These examples are quite simple.

(Read bullets out loud)

Looking at these examples, what are some of the issues or questions that you might want to ask about in order to decide whether these formulae would give a reliable estimation of size.

(Ask the participants to call out questions and write them down on a flip chart. Some of the key issues are described in the next slide, prompt the participants as needed, to have them discuss as many of the key issues as possible.)
Summary of key issues/questions

- **What about geographic variability?** - Applying one percentage to a whole population may not be appropriate.
- **How local are the data?** - The % used for the extrapolation may not come from local data and may not be applicable.
- **Do the group definitions match?** Does the % apply to the specific group(s) of interest – e.g.
  - Does sex work in urban areas refer to all types of sex work? Brothel based, street based and entertainment establishment based? Or just a subset of these groups?
  - What type of MSM are included in the 2.5%? Higher risk MSM? Venue based MSM? Any man who has sex with a man?

Facilitator notes:

(Review some of the key issues/questions, summarize the examples or details given by participants based on what was written on a flip chart in the previous slide).

- The key message is that the extrapolation method should be as specific to the local data, context, and purpose for size estimation, as possible.
- Some formulae may be appropriate and give useful results in one context, but the same formulae may not be helpful in another area.
- (Extra discussion point: Ask the group, what could be done to improve the extrapolation formulae?)
Facilitator notes:

Another general point related to deciding what approach to extrapolation should be used in a particular country or for a particular size estimation exercise is that the approach can be very simple or very complex.

- A simplistic approach is to apply a single formula or generic percentage to the whole geographic area. This approach may be chosen if there is no other source of information available, or the level of precision or accuracy required for the size estimate is broad, i.e. only a very approximate figure is needed.
- A more complicated approach would try to use the more refined relationship between the source of proxy data and the size estimate of a particular HRG locally. To do this, some of the factors that influence the relationship may be identified, and areas where extrapolation is needed would be characterized by these factors. Then decisions would be made about how to adjust the formulae to account for these different factors in different areas.
- In practice, most extrapolations fall in the middle of the spectrum and would depend on the type of information that is readily available and the degree to which greater precision or accuracy is important for the exercise.
Example of complex extrapolation – Indonesia Female Sex Workers (FSW)

- Main data source: survey of village leaders (PODES) “are there sex work spots in your village?”
- Calculate the % of villages in each district whose leaders said “Yes”
- Rank districts by quintiles according to the % of villages with sex work spots
- Group districts with size data by quintiles of % of villages with sex work hotspots.
- In each quintile group, use size data to calculate the average % of adult female population that are FSW.
- Apply the average % to districts without data in the matching quintile group.

Facilitator notes: To get a better idea of what might be involved in an extrapolation exercise that is on the relatively complex side, let’s look at what was done to obtain national and provincial level size estimates for FSW in Indonesia (done in 2006).

- In this eg, the main geographic unit is district. In Indonesia, many districts have a direct size estimate obtained from some type of mapping. However, there are 440 districts in Indonesia, so it is not feasible to collect data on size in all the districts.
- One important asset that Indonesia has is a survey of village leaders in every village, and this survey, called PODES, is done every 3 years. In the last two rounds a few questions to help with size estimation of HRG were included. The relevant question for FSW was, “are there known sex spots where sex work activity takes place in your village.” Village leaders responded yes or no, and the percentage of villages whose leaders said yes, could be calculated for all 440 districts in the country.
- To extrapolate the data on size estimates in the districts where a type of mapping data was available to districts that only had PODES data, Indonesia wanted to apply the percentage of the adult female population that were FSW, to areas that had similar PODES “risk scores.”
- The match was done by ranking districts according to quintiles for the “% of villages in the district with sex work activity.” (Walk through the diagram in the slide to illustrate how this was done).
- Table is not clear. 1) maybe you should give example of what the quintiles were based on and 2) label clearly that the percentages represent proportion of women who are sex workers.
Facilitator notes:

We can take a closer look at how this calculation was done. The following table is a mock-up of the spreadsheet used.

- The first column is the district, the second column gives the adult female population in the district, and the third column shows the quintile of the district when ranked according to the % of villages in the district where sex work activity was reported.
- Using the data in this table and the information on the previous slide that shows the % of the adult female population who are FSW by quintile in the districts which had direct size data, try to calculate the extrapolated size of the FSW group in each district.

(Give the participants a few minutes to work on the calculations. When a majority is finished, go to the next slide for the answers.)
Facilitator notes:

- Here are the numbers that should have been filled in. The 2nd row of the table shows the formula that is used. WHERE? NOT CLEAR.
- Given this example of how extrapolation was done in Indonesia for FSW, what would you say are some of the strengths and limitations of what was done?
- (Get the participants to brainstorm their ideas. Write their responses on a flip chart. Prompt general ideas, as needed, then move to the next slide to summarize the discussion.)
Facilitator notes:

The key strengths of the Indonesia example were having a proxy data source that was available in all the areas and trying to refine the extrapolation into 5 groupings of districts.

The details were not presented here, but some key questions to ask about how the specific formulae were developed relate to the reliability and frequency and clear description of how the formulae were developed and the key decisions that were made, are an important part of presenting and distribution of the proxy data as well as the direct size estimates.

For example, (Read bullets).

Providing a succinct disseminating extrapolated data.
Where on the spectrum of complexity should we be?

- The level of complexity to choose depends on:
  - How you are going to use the size estimate? What level of precision do you need?
  - What is the quality of the proxy data you are using? What is the limit of the level of precision you have?

Facilitator notes:

How does a country make a decision about how complex the extrapolation process should be? There are two main considerations.

- The first is knowing how the size estimate is going to be used. This will dictate in part how accurate and precise the estimate of size needs to be. For example if the national estimate can be presented with a range of +/-10%, and the refinements for extrapolation wouldn’t make a difference within that range, it may not be worthwhile trying to develop more complex formulae.

- Another consideration is the quality of the proxy data being used. If the proxy data are not very reliable and have a great deal of error or imprecision, it doesn’t make sense to over-analyze the data you have and try to develop complex and nuanced formulae for size estimation.

For example, if the PODES data were not available in all villages, and the only measure of sex work activity were a subjective scoring of 1 (not a problem) to 5 (Vibrant sex work industry) rated by the district health officer, these data may not be reliable enough to develop a separate formula for each risk category of district.

Understanding the strengths and limitations of your data will help you to make the right decisions.
Facilitator notes:

- Let’s move to a concrete application of extrapolation for size estimation and walk through the steps for doing this exercise in your country. You may have already had to do this for doing estimates and projections using some of the tools put out by UNAIDS and WHO. When using EPP or the workbook method one of the inputs you may use is the size of different HRGs.
- Another way you may be familiar with using national level size estimates is for overall planning or budgeting for high risk group interventions. This number can be useful to estimating an overall budget, or setting targets monitoring coverage against a meaningful denominator.
- In almost all case some form of extrapolation will be necessary, and the formula used will have to be updated when the environment changed or the direct estimates of size become more stable and reliable.
One approach for planning a national size estimation with extrapolation

1. Clarify the use of the size estimate
2. Define the geographic scope and population definition
3. Collate the available data that may be used: both direct size estimates and proxy data
4. Examine the direct size estimates
5. Examine the proxy data
6. Take a step back: do you need to collect more data?
7. Develop an algorithm and formulae for extrapolation
8. Document the process

Facilitator notes:

Here the key steps for one approach in planning and executing national size estimation are listed here. (Read bulleted List)

We will go through each of these steps in a little more detail using some concrete examples. As we go through each step, think about how this might apply in your country. Or whether you have done something similar.
1. Clarify the application for the national level size estimate

- How precise an estimate will you need?
  - Ukraine IDU: 284-371 K (+/-13% of 327.5K)
  - Bangladesh FSW: 54,600-90,000 (+/-24% of 72.3K)
- What resources are available to collect/collate data that can be used for creating the estimate?
  - Is there a budget and timeline for the ability to collect additional data?
  - Or is this an effort that must be based on available data, with planning for future data collection and refinement of the estimates?

Facilitator notes:

- The first step is to clarify how you will be using the national level size estimate of a HRG.
- This gives you an initial sense of how much precision you need. For example, in some large countries the ranges given for their size estimates of FSW were:
- How you will use the size estimates also gives the context for the exercise and whether there are resources available for the extrapolation process. The level of resources will determine how extensive the data collation and collection process can be. For example, (read the bullets)
2. Define the groups and geographic scope for the size estimates

- Tie the definition to the purpose of the size estimates.
  - If this is for estimating the burden of infections, what (sub-) group is really at risk for HIV
  - If this is for planning/budgeting interventions, what (sub-) group is the most critical to cover with interventions?
    - Only groups that are realistically accessible to interventions?
    - Only groups that are in urban areas? Or large cities?
- The groups and geographic areas covered by the extrapolation formulae should be consistent with these definitions
  - Present the definition with the final size estimate results to be clear and avoid confusion.

Facilitator notes:

The next step follows from the first, because the definition of the group must be consistent with this way the size estimate will be used.

- If the purpose of the estimate is for inputs in calculating the burden of infections, the group definition should reflect the group who is most at risk for HIV. For example, when estimating MSM, it is important to define the sub-group of MSM who have high numbers of partners and frequent sex. This may translate to defining the group as venue-based MSM, those who cruise for partners in public venues.
- The specific definition used should be based on country-specific data on the context of risk, and how these HRG are best characterized.
- If the size estimate is being used for planning or budgeting the national programme, the definition of the group should reflect the sub-group that will be targeted for interventions.
- This should be partly based on the group who is most at risk for HIV, but it may also reflect practical considerations for establishing project sites. Sometimes this translates into limiting the definition to groups in urban areas or excluding home based sex workers which may not be easily accessed and who may be lower risk.
- Whatever the final definition for the group turns out to be, this should be reflected in the extrapolation formulae and presented with the final result to be clear to all audiences.
Facilitator notes:

Step 3. is to collate all available data to determine where and what type of data you have to work with.

- Start with the available data collected for size estimation. This will require setting up a spreadsheet or database that shows for every geographic unit (e.g. district) what size data is available. In some cases you may have more than one source of size data in the same location. Catalogue everything and be sure to make notes about the data source. If more than one group is being estimated, catalogue what is available on separate sheets or in separate columns to make things easier to organize.
- Make a note if there is a pattern to the types of areas where direct size data available. For example, is data only available in the largest cities, where there is an intervention, etc.
- Then turn your attention to potential sources of proxy data. For this effort you may have to be a bit creative and explore data sources which may be collected by groups in other sectors or groups. And this may require making contact with other groups involved in surveys or data bases. The kind of data available in each country will vary considerably, but the following list may help to give ideas of what to look for. (Read bullets)
- Organize and notate the proxy data in a similar fashion to what you did for the size data.
Brainstorming sources of proxy data

- Characteristics of useful proxy data
  - Data which characterize geographic areas by factors related to presence of HRG
  - Data related to the presence of HRG
  - Local data on the socio-demographic profile of HRG

- List some examples of what you’ve used for
  - FSW
  - MSM
  - IDU
  - Migrants (as a sub-group of clients of SW)

Facilitator notes:

Identifying good sources of proxy data are difficult. Let’s take a few minutes to brainstorm sources that you have used in the past or know about. Remember characteristics of useful proxy data are

- Those that characterize geographic areas related to the presence of HRG (such as the PODES data in Indonesia)
- Or data that is closely related to the presence of HRG and is available widely for different geographic areas (e.g. drug arrest data)
- Or data which characterizes the socio-demographic profile of HRG that can be used with general population size.

(Have the participant think for a few minutes about the data sources they know. Then go HRG by HRG and ask participants to call out ideas. Write these down on a flip chart. If it is not clear how the data source can be used as a proxy for size, ask the participant to elaborate until it is clear.)
4. Examine the direct size data

- Review the size data and talk to the persons responsible for collection/analysis of the proxy data
  - Understand how the data were collected
  - How reliable are these data?
  - How widely available are the data? Are there certain kinds of areas where direct size data are not available, e.g. rural areas, non intervention areas, etc.
  - Is there variability in the way data are collected in different areas?
- Where there are multiple sources of direct size data, reconcile the estimates for each areas where data are available
  - Is one source of size data clearly more reliable or complete than another?
  - If there is no clear hierarchy, can a composite value or range of values be used?
- Analyze the direct size data
  - What does the frequency distribution of the direct size data look like?

Facilitator notes:

Step 4. of this process is where the available size data are analyzed.

- First, make sure the nuances of how the data were collected and what its strengths and weaknesses are clear. This will require looking for inconsistencies in the data and talking to people involved with the data collection and analysis process.
- Some people may become sensitive when you try to understand the method used to collect the data, as they may take it as a form of criticism. Clarify that your intent is to be able to use the data most effectively in terms of what groups were included or excluded and to understand whether the data may be on the high side or low side of the true estimate.
- If there are multiple sources of data available in the same areas, decide which data source may be more reliable in each area or how to triangulate the results and use one other source of data to calibrate the other or as an upper or lower bound for a size estimate range.
- In all areas where data are available, finalize the size estimate for that area. Review the frequency distribution and understand the variance or whether there is clumping of areas that are similar in size. Look for outliers or areas where sizes may be much higher or lower than expected given what is known anecdotally about an area, or based on other characteristics such as population size, etc.
Facilitator notes:

The next step is to analyze the proxy data in a similar way.

- Again, start by understanding the strengths and limitations of the data by discussing the process of data collection and analysis with the groups or individuals involved.
- Begin to analyze the proxy data by looking at the frequency distribution. Is there enough variability in the proxy data by geographic areas, is the pattern of variability what you might expect given what you know about the areas.
- Compare the proxy data to other data sources including size data in the areas where both are available. Is there a correlation between the two data sources, is it consistent across different types of areas. If there is some difference in the correlation can it be explained by other characteristics or factors about the geographic areas?
Facilitator notes:

- After looking at the different data sources that are available for the extrapolation and understanding their strengths, weaknesses and potential relationship to each other, it is worth considering whether the extrapolation exercise will give you a satisfactory result or whether some additional data collation or data collection will greatly improve the process.
- If there is a piece or type of data identified, is it feasible to pursue this data. Consideration should be given to the financial cost, the time it will require, and if appropriate staff are available to participate.
- Another path may be to proceed with the extrapolation exercise, but to move forward with the data collation/collection in parallel and when the data become available to update the analysis.
- The right choice will depend on the country context, resources, and external timelines for the use of the national level size estimate.
7. Develop the algorithm/formulae to use

- Develop a general approach for converting proxy data into a size estimate
  - Which data sources to use
- Decide on which groupings of geographic units require different extrapolation values
  - Decision should be based on evidence that distinct characteristics of geographic units are related to size of the HRG population
- Develop a data-driven method for obtaining different extrapolation values for different groups
  - Based on local data as much as possible
- Test the algorithm/formulae
  - If the approach is used in areas where there is direct size data, is the result of the extrapolation similar?

Facilitator notes:

- The next to final step is to formalize the algorithm or formulae used in the extrapolation. This process is done by synthesizing what has been learned or decided in the previous steps.
- Depending on how complex the formula is, start with the general approach for converting the proxy data into a size estimate. Start by identifying the data sources that will be used and their use in a numerical formulae.
- If different extrapolation values or formulae are used for different types of geographic areas, decide on the groupings that will be used and make sure these groupings are based on evidence that the groupings are associated with a different proportions to HRG size.
- Using the same data, decide on the different extrapolation values/formulae to apply to each group.
- Then test the algorithm/formulae in the areas where size data is available. Does the result give something comparable to the size measured through mapping- or survey-based methods? Adjust the formulae or the groupings as needed.
Facilitator notes:

- After all these steps, the actual calculation of size is trivial. However, the last step is one of the most critical aspects of the extrapolation exercise: Document the process.
- The purpose of the documentation is to provide a clear and transparent record that can be understood by persons who were not involved directly in the extrapolation process. This will ensure that the exercise can be repeated and consistent calculations are possible. A clear summary of the approach, formulae, and extrapolation values used will also be critical for explaining the results to others who may question the strengths and limitations of the estimate.
- Key tips for adequate documentation include: (Read bullets)
What organizational structure is needed to support size estimation?

- How to develop an overall design?
- Who will plan and implement specific size estimation data collection/analysis?
- How to coordinate other groups with potential/opportunity to collect size related data?
  - Linking As, Bs, and Cs
  - Linking national to local
  - Linking health sector to other sectors
- How to collate and triangulate data as it becomes available?
- What resources are needed to sustain these structures?
Country Exercise for National Estimates
Country Exercise for National Estimates

Using your data for national level size estimates of HRG

This tool will walk through the issues related to designing a national level size estimation of a selected HRG for the purpose of estimates and projections.

If you already divide the country into regions when using epidemic modeling tools, continue to use those divisions for the purpose of this exercise, assuming that the process developed will be for estimating the size for each region.

(Follow this process for each HRG e.g. FSW, MSM, IDU, migrants, separately)

For the purposes of this workshop, prepare a summary of your group work to share with the other countries on the morning of Day 4.

- If you have the data available to develop an algorithm, present the algorithm you plan to use, the data sources that are available, and your rationale for using this approach. (see Step 7 for guidance on tips for documenting and presenting your approach)
- If you know what data sources are available but do not have all the data you need with you, describe your general approach, what data are currently available, and your plan for follow up.
- If data are very limited in your country, describe the general approach you think you will take, where you will obtain direct size estimate data, where you will rely on proxies and extrapolation, etc.

1. Define the population

For the purpose of this exercise, Choose to start with the HRG that is of greatest interest for your country.

Selected HRG: __________

a. Describe this population with respect to the group that is relevant for the size estimation. For each characteristic included in the definition, provide a short reason for including this in the definition, and cite any data source for why this aspect of the definition is important or relevant. Here are a few issues to consider in defining the group:

   i. describing the groups included in the definition (by type or

   ii. defining the group by a minimum level of risk intensity (e.g. by how frequently they inject or engage in risky sex acts, by how recent they last injected or engaged in risky sex acts, etc.)

   iii. are there geographic boundaries that are relevant (e.g. HRG in urban areas, or in districts greater than a certain population size, etc.)
Fill in each Table:

a. Definition of the population:

<table>
<thead>
<tr>
<th>Criteria of definition</th>
<th>Reason for criteria (data source)</th>
</tr>
</thead>
</table>

b. What are the important sub-groups for which separate size estimates may be useful. These sub-groups may be different enough that the process for estimating their size would come from different data sources or use different algorithms. Or these sub-groups may be important to distinguish in the process of estimates and projections because they have very distinct levels of risk, that should be modeled separately. When describing the sub-groups be sure to describe why this sub-group is important and any evidence that suggests that risk behaviour varies in different sub-groups or if the data sources used to describe them are different (e.g. surveys that cover only street based IDU).

i. for FSW these sub-groups may be defined by type of solicitation points (e.g. brothels-based, street-based, bar-based, massage parlour-based, etc.)

ii. for MSM these sub-groups may be defined by gender or sexual identity or sexual practice or whether they sell sex.

iii. for IDU these sub-groups may be by types of drug used, whether IDU live on the streets or in homes with their families, or places where drug users inject (e.g. shooting galleries or from professional injectors).

b. Important sub-groups

<table>
<thead>
<tr>
<th>Sub-group</th>
<th>Detailed definition</th>
<th>Reason to estimate separately/Comment</th>
</tr>
</thead>
</table>

2. Collate the available data

a. List out all the data sources (including date) that are available that have some relation to size estimates. Organize the type of data in the following categories. Mapping-based data, Survey based data, and other proxy data (including general population distributions, data sets on other health sectors that can be used for extrapolation). List data sources separately by date if they are different rounds of a survey that was conducted using the same protocol (e.g. BSS 2002, BSS 2005, BSS 2008, etc.) or if different protocols were used in different areas even for the same round (e.g. State AIDS Control Society Mapping, 2007; “XYZ” NGO Programme Mapping, 2007)

a. Data Sources

<table>
<thead>
<tr>
<th>Title of Data Source</th>
<th>Year</th>
<th>Implementer</th>
<th>Comment</th>
</tr>
</thead>
</table>
b. Create a table that lists all the geographic units in the country, starting with larger units (e.g. state or province) down to the smallest unit for which data are collected (e.g. cities or districts), in the first column. Head the subsequent columns with the major sub-types defined in #1.

In each cell of the table, list which data sources are available in which years for each sub-group and by geographic unit. If there is a data source that relates to an aggregated area and cannot be separated into smaller geographic units, make a note of this in the row for the corresponding higher geographic unit (e.g. state or province)

(Once this is done for all HRG, these matrices can be consolidated into a single spreadsheet.)

b. Catalogue of Available data by group and geographic area

<table>
<thead>
<tr>
<th>Geographic Area</th>
<th>Sub group 1</th>
<th>Sub group 2</th>
<th>Sub group 3</th>
<th>Overall Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District A.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District A.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District A. 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>…</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District B.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District B.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District B.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District B. 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>…</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State C…</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>…</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From looking at the completed matrix it should be clear where you have size estimation data and where you have useful proxy data for the different groups. From this initial assessment it can become obvious what kinds of methods for extrapolation may be possible to use, or where additional data may be most needed.

3. **Examine the direct size estimation data**

(If it is not possible to make a full assessment of each data source, note the areas which may require follow up when you return, and with whom or how you would do the follow up.)

For each data source review and describe the following issues

a. Do the population definitions match? Given how you defined the population in Step 1, well do the definitions match in terms of geographic scope, sub-groups, etc. Do you have enough information to determine whether there might be some difference between the intended definition of the population and selection bias which may have been introduced due to field conditions or the way the data were collected? Is there variation in how well the populations match in different areas.

b. What is the quality of the size data? Is there evidence that the data source may be an underestimate or overestimate of actual size? Or is there inconsistency in the way the data were collected in different areas, making the data quality better in some areas than others. Describe where the data may be less reliable.

**Example Format:**

Size Data Source 1: _______________________ (use the same titles as in earlier table of data sources)

<table>
<thead>
<tr>
<th>a. Population definition according to protocol/methodology:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b. Possible differences from desired population definition due to field conditions or implementation:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c. Geographic variation in the population definition of the size data:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
d. Data quality concerns:

e. Geographic variation in the quality of the data:

f. If not enough information to assess, what and with whom follow up will be done:

Data source 2: …Etc

4. Examining the proxy data

(if it is not possible to make a full assessment of each data source, note the areas which may require follow up when you return, and with whom or how you would do the follow up)

For each data source review and describe the following issues

a. How widespread are the data? Is it available in a large number of geographic areas? i.e. is it possible to use a uniform approach to extrapolation for most of the areas? Do you have proxy data in enough areas where direct size estimates data are available in order to develop an empirical relationship between the proxy data and actual size data. If multiple types of proxy data are needed for the algorithm, are these data equally available for specific geographic areas?

b. What is the quality of the proxy data? Does the method of data collection follow a rigorous protocol? Is it possible to assess bias or error? Is there geographic variation in the way the data are collected in different areas?

Example Format:

Proxy Data Source 1: ____________________________ (use the same titles as referred to in the table of data sources)
## TRAINING MANUAL ON METHODS FOR SIZE ESTIMATION OF KEY AT-RISK POPULATIONS IN THE ASIA-PACIFIC REGION

### a. Availability of proxy data

Proxy data are available in ____ out of ____ total geographic units  
(Describe generally the areas where the data are available. Describe the areas where there are major gaps in proxy data.)

### b. Availability of both proxy data and direct size data to develop the algorithm

Direct size data are also available in ____ out of the ____ number of geographic units where proxy data are available. (Describe the areas that have both direct size data and proxy data by key characteristics, e.g. at least one large metropolitan areas, or at least one area in each region, or data in the coastal area and highland area, but not in the lowland areas...)

### c. The additional proxy data needed for the size estimate extrapolation algorithm

(e.g. size of the adult (15-49 year old) female population for each geographic unit; the SES index, the number of AIDS cases reported in 2008, etc.)

Source 1:  
Source 2:  
Source 3:  

### d. Data quality concerns:

### e. Geographic variation in the quality of the data:

### f. If not enough information to assess, what and with whom follow up will be done:
5. Take stock of where you are

a. Are there any data that would be very critical to collect but that would improve the reliability of the correction factors or the use of the proxy data? – rank the data that is of highest priority (because of the epidemic potential, or because of the wide applicability of the data), e.g.
   - key geographic areas with large epidemic potential but no size data which would benefit from geographic mapping
   - key areas which are likely to have large HRG populations but weak or outdated data that can be quickly updated
   - data which would help to categorize geographic areas by risk level
   - data which would help to determine the appropriate adjustment factors

b. Rate each type of data by how easy or difficult and/or costly it would be to obtain. Specify whether field work is needed to collect the additional data, or whether it just needs to be collated from different sources/areas.

c. Identify whether there are resources available to collect/collate these data? From what source? And whether funds are available immediately, or only potentially.

Organizing priority data needs

<table>
<thead>
<tr>
<th>Type of data needed and for what sites (list in order of priority)</th>
<th>What is required to obtain data, i.e. requires field work and/or collation from different sources</th>
<th>Level of difficulty/resource intensity to collect 1(low) -5 (high)</th>
<th>Resources available to collect data (possible/already available)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

d. Develop a timeline and decision tree for either proceeding and updating estimates at a later time when data become available or waiting until specific data are collected/collated
6. Develop the algorithm/formulae for extrapolating the data.

Depending on what type of size data you are working with, different extrapolation techniques may be needed. Decide what type of situation you are working in for different geographic units and then decide what type of algorithm is needed.

1. Do you have size estimates for a part of the population (i.e. sub-group or limited geographic areas) that need to be inflated to reflect the whole geographic area?

For example:

- only male IDU that must be adjusted to include female IDU;
- only urban MSM that must be adjusted to include peri-urban MSM;
- only street based FSW that must be adjusted to include other part-time sex workers and home-based FSW, etc.

a. Draw a diagram to show how you think the relationship between the data you have relates to other sub-groups or parts of the population that may be out there and who are relevant (i.e. have a moderately high level of risk)

Example: Venn diagram of IDU populations, How to transform size estimation data available for urban male IDU into size estimation data for all IDU in the area.
i. Think about geographic areas that are not included (rural vs. urban areas, towns or villages not included, proportion of the population not included)
ii. For IDU, does the size estimate include both male and female IDU; does the estimate include IDUs who inject drugs (IDU)?

b. Are there any sources of data locally that give an idea of how the sub-group you have data for, may be related to the larger group of interest? If local data are not available, are there similar data in the region? In the literature?
   i. e.g. mapping data of IDU in one district that covered urban and rural areas and found that the % of IDU that lived in urban areas was 85% of the total IDU mapped. And the same mapping data showed that 6% of the IDU were female.

c. If there is more than one source of local data or more than one area where the data are available, look at the variation in the %. If it is wide, are there characteristics of the areas that might explain the large difference?
   i. e.g. districts near borders where lots of drug trafficking takes place had higher proportions of IDU in rural areas; or areas where sex workers were in large numbers had higher % of female IDU of the total IDU population. Consider dividing the geographic areas into multiple categories based on these characteristics and applying different extrapolation values (See 4.)

d. If there is more than one source of data and the variance is not too big, use an average value or a conservative value, or choose the value of the most well designed/most extensive data source. Document the reason for the decision made.

e. If data are not available locally and regional or data from the literature do not look realistic, can a local expert group be convened to discuss their experience to develop a consensus adjustment that can be used.

2. Do you have size estimates for a population that needs to be deflated to remove sub-groups that are not part of the defined population of interest?

For example
- estimates for drug users that need to be deflate to include only injectors;
- all MSM that need to be deflated to include only venue based MSM, or only MSW, etc.
Follow the same steps (above in 1.) to determine how the correction factor should be determined. Articulate how the group you are interested in estimating size for is a subset of the population for whom size data are readily available.

3. Do you have several sources of size data for the same geographic area?
   a. Does the quality of data for different methods vary across sites?
      i. Decide on criteria to use for excluding data sources that do not meet a minimum standard of quality.
      ii. Update the inventory of available data using these minimum quality standards. Document the decisions made to exclude size data where appropriate.
   b. Are there general issues around the protocol used to collect the size data that allow you to rank the source in terms of relative reliability compared to the other sources? (reasons may include the data are more up to date, more comprehensive in covering the group of interest, more rigorously implemented, etc.)
      i. Is there evidence that can be documented to justify this ranking.
      ii. Apply the ranking to choose which source of size data is used.
      iii. If these data need to be adjusted, go to 1.
   c. If it is not obvious that one source of data is more reliable than other, it may be necessary to conduct a detailed assessment of the potential bias for size data in each area.
   d. The same issues may apply to multiple sources of proxy data.

4. Do you have proxy data that needs to be transformed into a size estimate for a geographic area?

   For example
   o # of entertainment establishments that can be transformed into an estimate of # of entertainment based FSW
   o # of trucks passing through a transport hub in a week that can be transformed into an estimate of individual drivers and assistants
   o # of drug arrests that can be transformed into the number of IDU
   o % of household absentees that can be transformed into an estimate of single male out migrant
   o # of gurus in a city that can be transformed into an estimate of transgender sex workers

   a. If the proxy data is quantitative and has a wide spectrum of values look at the frequency distribution of the proxy data by geographic areas.
   b. Are there quantitative data that suggests how the proxy data are related to size in some areas?, e.g.
      i. Programme records for a city that can calculate the average # of FSW working in a bar, or a massage parlour, etc.
      ii. Survey data about how frequently truckers pass through specific transshipment points in a week or a month.
      iii. Administrative data on drug arrests that calculate the % of drug arrests who are current injectors
iv. Key informant interviews that estimate the average # of chela under each guru.

c. Does the same relationship between proxy data and size data hold in different areas where both pieces of information are available?
   i. If yes, is it possible to develop an average value across sites to use for all areas where the proxy data are available.
   ii. If no, go to #5.

5. Do you need to categorize geographic areas by risk to apply different extrapolation values to different types of areas?
   For example,
   o Urban areas are found to have a higher proportion of entertainment establishments FSW, while smaller towns have predominantly street based FSW. So different algorithms are used for districts with population size >400,000 and <400,000.
   a. Towns with large populations of FSW may have a higher % of MSM who are MSW.
   b. Test the formula in areas where there are size estimate data, does the formula give a reasonable result.

7. Documenting the algorithm:

   1. Describe the purpose of the size estimation and the desired population definition (as in #1).
   2. Give a brief overview of how the main source(s) of size data being used.
      i. how up to date is the data source?
      ii. how widely available it is?
      iii. Describe whether each source uses the same or different population definition as what is desired. Be clear if one source of data is used for estimating one sub-type and another data source is used to estimate another sub-type.
      iv. if there are more than one sources of size data, describe any hierarchy in choosing one estimate as the primary size estimate, or if the data are combined as some type of average, etc.

   3. Give the generic formula used to transform the main data source into a size estimate
      i. If a range of values is being presented, describe the generic formula for the minimum and maximum value separately.
      ii. If different formulae are used based on the availability of different data sources, describe the generic formulae used when a particular source of data is available and describe the number of geographic units where each formula is used.
4. If the geographic areas are divided into groups, in order to apply different adjustments factors, for each adjustment factor
   i. Describe the groupings and the basis for categorizing each area into one of the groupings (describe the source of data and how it is used; provide a rationale for the criteria used, if the groupings are based on a quantitative criteria, give the rationale for the cut-offs chosen, e.g. describe the frequency distribution and natural breaks points in the data.)
   ii. Describe how the different adjustment factors are determined for each grouping (give a rationale for the method chosen)
   iii. Summarize the process, creating a table that shows the grouping categories, the criteria used for the grouping, and the different adjustment factors assigned to each grouping.

5. Give a few concrete examples for applying the formulae in different geographic units, to illustrate each formulae or approach used.

6. Make notes on adjustments or changes if these were made and are exceptions to the decision tree and formulae generally used. E.g. when incorrect or unreliable data points are changed to more likely/realistic values based on information from key informants or efforts to investigate data quality.

7. Create spreadsheets for doing the calculations systematically
   i. use formula functions when possible so it will be clear which numbers are used as inputs and which numbers are calculated using the formula.
   ii. color code and label the spreadsheet cells, to make it clear to new users which cells should be population by data and which should be calculated from formulae.