Introduction to HIV, AIDS and sexually transmitted infection surveillance

MODULE 2

Surveillance of HIV risk behaviours

World Health Organization
Regional Office for the Eastern Mediterranean

UNAIDS
JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS

UNHCR
UNICEF
WFP
UNDP
UNFPA

UNODC
ILO
UNESCO
WHO
WORLD BANK

Introduction to HIV, AIDS and sexually transmitted infection surveillance

MODULE 2

Surveillance of HIV risk behaviours

World Health Organization

Regional Office for the Eastern Mediterranean

UNAIDS

JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS

UNHCR
UNICEF
WFP
UNDP
UNFPA

UNODC
ILO
UNESCO
WHO
WORLD BANK

Contents

Introduction

How to use this module

What you should know before the course

THE HUMAN IMMUNODEFICIENCY virus (HIV)/acquired immunodeficiency syndrome (AIDS) epidemic continues to grow worldwide and to devastate individuals, communities and entire countries and regions. Behavioural surveillance measures trends in the behaviours that can lead to HIV infection. It has been shown to make an important and useful contribution to national responses to HIV. Conducting behavioural surveillance requires many skills, including coordination among various partners. Although there are useful reference materials available for behavioural surveillance, there has not yet been a comprehensive effort to train surveillance teams. This training course aims to help address this.

This course is meant primarily for people involved in planning and using behavioural surveillance. You should already have a basic understanding of the epidemiology of HIV/AIDS and public health surveillance.

This module is part of a set of four modules that have been designed with a focus on the World Health Organization's (WHO) Eastern Mediterranean Region. The modules were designed for use in training workshops. The other modules are:

- Module 1: Overview of the HIV/AIDS epidemic with an introduction to public health surveillance
- Module 3: Surveillance of most-at-risk and vulnerable populations
- Module 4: Introduction to respondent-driven sampling.

Similar training modules have been developed for WHO's African, Americas, European and South-East Asia regions. Although the overall framework of the modules is the same, each region has different patterns of HIV epidemics and distinct social and cultural contexts. Also, different countries may have different HIV surveillance capacities and different needs. Thus, these modules were developed taking into account the specific context of the HIV epidemic in the countries of the Eastern Mediterranean Region.

The modules are also intended for use in the countries of the Joint United Nations Programme on HIV/AIDS (UNAIDS) Middle East and North Africa Region. For the purpose of this training course, all countries in the WHO Eastern Mediterranean Region plus Algeria are therefore the intended audience. We refer to these collectively as Eastern Mediterranean Region/Middle East and North Africa (EMR/MENA) countries.

Module structure

This module is divided into units. The units are convenient blocks of material for a single study session. This module can also be used for self-study.

Because you already know quite a bit about HIV/AIDS, we begin each unit with some warm-up questions. Some of the answers you may know. For other questions, your answer may be just a guess. Answer the questions as best you can. You will keep your answers to the warm-up questions in this module. No one will see your answers but you. We will study and discuss the unit, and then you will have time to go back and change your warm-up answers. At the end of the unit, the class will discuss the warm-up questions and you can check your work.

As you study this module, you may come across terms and acronyms that are unfamiliar. In Annex 1, you will find a glossary that defines many of these.

Annexes

More information is provided in the following annexes:

Annex 1: Glossary

Annex 2: Useful links

Annex 3: Answers to warm-up questions and case studies

Annex 4: Action plan for implementing HIV behavioural surveillance of most-at-risk populations in EMR/MENA countries

Additions, corrections, suggestions

We welcome feedback on this training module. Please send your suggestions for any changes or additional information that might be included to the following address for possible inclusion in future editions.

Mailing address:

World Health Organization
Regional Office for the Eastern Mediterranean
Division of Communicable Diseases
AIDS and Sexually Transmitted Diseases (ASD) Unit
Abdul Razzak Al-Sanhouri Street
P.O. Box 7608 Nasr City
Cairo 11371, Egypt

Email: asd@emro.who.int

Acknowledgements

This manual was prepared by the University of California San Francisco, San Francisco, USA, in collaboration with: WHO Regional Office for the Eastern Mediterranean, Cairo, Egypt; WHO headquarters, Geneva; Joint United Nations Programme on HIV/AIDS (UNAIDS), Regional Support Team, Cairo, Egypt; World Bank; and Centers for Disease Control and Prevention (CDC), Atlanta, USA.

Special thanks are due to Family Health International (FHI), whose behavioural surveillance materials provided a basis for this training module. The FHI materials were initially adapted by Tulane University, School of Public Health and Tropical Medicine, New Orleans, USA, through UTAP with CDC/GAP.

Unit 1

Behavioural surveillance

Overview

What this unit is about

IN THIS UNIT, you will learn basic information about behavioural surveillance, including its purpose and history and important things to consider.

Warm-up questions

1. Which of the following can be a use of behavioural surveillance?
   a. To explain changes in HIV prevalence over time
   b. To provide information for prevention programmes
   c. To raise the awareness of HIV among policy-makers
   d. To give us an early warning about which groups and areas infection is likely to spread in and between.
   e. All of the above

2. _____ surveillance involves regular, repeated, cross-sectional surveys to collect data on HIV risk behaviours and other relevant issues which can be compared over time.
   a. behavioural
   b. biological

3. True or false? Behavioural surveillance can be used to evaluate the effect of interventions or parts of programmes, if questions about them are included in behavioural surveys. Circle your answer below.

   True   False

4. True or false? In a generalized epidemic, everyone is at equal risk for infection.

   True   False

5. _____ sites are facilities such as sexually transmitted diseases (STI) clinics, antenatal care clinics, blood donation centres, drug treatment programmes, prisons and needle-exchange programmes.
   a. Sentinel
   b. Community

6. Which of the following is the definition of linking behavioural and biological data?
   a. Collecting HIV, STI and behavioural data from the same people at the same time
   b. Collecting HIV, STI and behavioural data from the same source population at different times
   c. Analysing HIV, STI and behavioural data from a similar source population, using whatever data are available
   d. Reporting behavioural and biological surveillance together
   e. All of the above

7. Collecting data at the _____ level provides more detailed information but requires larger sample sizes and more time and money.
   a. national
   b. subnational

8. True or false? Regardless of the state of the HIV epidemic (low-level, concentrated, or generalized), behavioural surveillance must be conducted among the general population every year.

   True   False

Introduction

What you will learn

By the end of this unit, you should be able to:

- define surveillance
- summarize the uses of behavioural surveillance
- summarize issues to consider when designing a surveillance system
- summarize what steps to take for a sustainable surveillance system.

Definitions

Surveillance is the systematic, regular and ongoing collection and use of data to take action for public health. Although they are often the first step in starting a surveillance system, one-time cross-sectional surveys should not be considered surveillance. HIV/AIDS surveillance can be divided into biological and behavioural surveillance.

Behavioural surveillance is the systematic and ongoing collection of data about risk behaviours related to diseases or health conditions. It is done to compare trends in behaviours with changes in the disease over time. Signs of HIV infection appear several years after a person is exposed to the virus. Behavioural data let us observe changes in the behavioural risk factors that drive HIV epidemics without having to wait for changes in the disease.

In biological surveillance, biological samples are collected and tested for HIV and other related illnesses, such as sexually transmitted infections (STIs) and tuberculosis (TB).

Cross-sectional surveys allow us to collect information from a sample of people from a target population at one point in time or over a short period of time. In behavioural surveillance, the same survey is repeated with the same target population, but with a different sample of people, at regular intervals. Doing this helps us to explore behavioural changes over time.

Very often, biological and behavioural HIV surveillance is done at the same time for a more complete view of HIV epidemics, but they can also be used separately. Linking biological and behavioural HIV data will be discussed in the next few pages.

Populations of interest in HIV surveillance

For the purpose of HIV surveillance and strategic HIV programme planning the concepts of risk and vulnerability of populations are important:

Risk is defined as the probability that a person may become infected with HIV. Unprotected sex, multiple sexual partnerships involving unprotected sex and injecting drug use with contaminated injecting equipment are the most important risk behaviours. Risk increases with the prevalence of HIV in the sexual or drug-using network.

Vulnerability to HIV is defined as the extent to which individuals are able to control their risk. Factors that make people vulnerable include: a lack of knowledge about HIV or lack of skills to avoid risk behaviours; inability to access condoms, clean needles, or other

protection; gender and income inequality that prevents sexual negotiation or contributes to forced sex; and discrimination and stigma, which deter people from changing risk behaviours. These factors, alone or in combination, when shared throughout a community, create collective or community vulnerability. Vulnerability is independent of whether HIV is highly present. Where HIV vulnerability is high it is likely that an individual or community will be less resilient to HIV risk if or when HIV increases within the network.

Populations of interest in HIV surveillance include the **general population**, **most-at-risk populations**, **bridging populations** and **vulnerable populations**. The definitions of these populations used in this module are the following:

- **General population:** The general population encompasses all people living in a defined community, e.g. a country, a province or a city. If surveillance data on the general population exclude certain subpopulations, e.g. specific age-groups of the population, foreign workers, refugees or others, this should be stated clearly. The general population is at relatively low risk of HIV exposure in low-level and concentrated HIV epidemics.
- **Most-at-risk population (MARP):** These are populations that experience the highest probability of being exposed to HIV and include injecting drug users, men who have sex with men (including male sex workers) and female sex workers.
- **Bridging populations:** Bridging populations are populations at intermediate risk of exposure to HIV and provide links between most-at-risk populations and the low-risk general population. Bridging populations may include the sexual partners of injecting drug users, sex workers and men who have sex with men. For example, a client of an HIV-infected sex worker may become infected with HIV. He may then have unprotected sex with his wife, transmitting the virus to her. In this scenario, he has acted as a bridge by which HIV infection has passed from the sex worker to his wife.
- **Vulnerable populations:** Vulnerable populations are found in all settings. Their risk of acquiring HIV depends on the prevalence of HIV in their networks and their ability to adopt safer behaviours and access prevention and treatment commodities and services. For that reason, populations such as people affected by humanitarian emergencies and migration, young people, women, prisoners and people living with disabilities may be at high risk in some places but not in others. Once these vulnerable populations adopt high-risk practices they become part of most-at-risk or bridging populations.
- **Key populations:** Key populations include most-at-risk, bridging and vulnerable populations, who, while being important to the dynamics of HIV transmission in a setting, are essential partners for an effective response to the epidemic, i.e. they are key to the epidemic and key to the response. Key populations will vary depending on the epidemic and community/country context.

Uses of behavioural surveillance

Uses of behavioural surveillance include the following:

- To give us an early warning about which groups and areas infection is likely to spread in and between.
- To explain changes in HIV prevalence over time. Without behavioural data, biological surveillance data are difficult to interpret. For example:

○ Does HIV prevalence that is stable or decreasing show that there are fewer new infections? Have more deaths removed more people from the population being tested? Has the population changed over time?

○ Does HIV prevalence that is increasing mean that prevention programmes are not good enough? Is life expectancy rising because of the success of treatment programmes? Is the increase because the epidemic has not reached the point where people are dying? Has the population being tested changed over time?

- To provide information about how to develop prevention programmes. Behavioural data let us identify the populations and behaviours that should be targeted.
- To monitor and evaluate the effects of prevention programmes. The national monitoring and evaluation strategies for HIV/AIDS should use indicators from behavioural surveillance data. However:

  ○ Without adapting it, behavioural surveillance gives evidence only for the effects of entire HIV programmes, not for specific parts of a programme.

  ○ Behavioural surveillance can be adapted to evaluate specific interventions by adding questions about exposure to the interventions. This must be done with care so that the focus of surveillance still is on finding and measuring risk behaviours.

  ○ Behavioural surveillance does not give conclusive evidence that the programme caused any changes in behaviour. Any change that is seen may have happened because of some other factor.

- To reinforce the findings of biological surveillance.
- To raise awareness and knowledge about HIV among policy-makers and donors.

Designing a behavioural surveillance system

What to consider

When you design a behavioural surveillance system, you should consider:

- whom to include in surveillance
- where to contact the surveillance populations
- how to link data from biological and behavioural surveillance
- how to make sure that the surveillance is right for the context.

Whom to include

The current guidelines about whom to include in surveillance differ according to the type of epidemic in a country. Epidemics can be classified as low-level, concentrated and generalized epidemics.

- Low-level: prevalence of HIV is below 5% in any most-at-risk populations (MARPs) and below 1% in the general population. HIV has not spread to significant levels in any subpopulation including most-at-risk populations.
- Concentrated: prevalence of HIV is above 5% in one or more most-at-risk populations, but remains below 1% in the general population. HIV transmission has taken root in one or more most-at-risk populations, however HIV is rarely transmitted to those outside the most-at-risk populations and their regular sex partners. Protecting most-at-risk populations would prevent a wider epidemic.
- Generalized: prevalence of HIV is above 1% in the general population. HIV has been

established in the general population due to high proportions of both males and females engaging in behaviours that facilitate HIV transmission. The epidemic would persist despite effective control of transmission in most-at-risk populations.

The guidelines for whom to include in surveillance for each type of epidemic are shown in Table 1.1.

Table 1.1 Whom to include in surveillance, by epidemic type

State of the epidemic	Biological surveillance (annually if possible)	Behavioural surveillance
Low-level	Most-at-risk populations	Most-at-risk populations every 2–3 years, general population every 3–5 years optional
Concentrated	Most-at-risk populations, general population in areas where HIV prevalence among most-at-risk populations is high and most-at-risk populations size is large.	Most-at-risk populations every 2 years for high priority sites, every 3-5 years for lower priority sites, general population every 3–5 years optional
Generalized	Most-at-risk populations, general population	Most-at-risk populations every 2 years for high priority sites, every 3-5 years for lower priority sites, general population every 3-5 years

Discussing the table

Looking at Table 1.1, answer the following questions:

1. What types of HIV epidemic are in your country?
2. Are different parts of your country experiencing different types of epidemic?

General population behavioural surveillance measures HIV risk behaviours in a sample that represents the people of a region or nation. It can be restricted to certain ages or social groups.

Behavioural surveillance of most-at-risk, bridging and vulnerable populations measures HIV risk in groups whose behaviours, occupations, circumstances or lifestyles could expose them to a higher risk of HIV infection and transmission than the rest of the population. In particular, most-at-risk populations are important in starting, hastening or maintaining the HIV epidemic. Bridging populations are populations that interact with both most-at-risk populations and persons of lower risk in the general population. Bridging groups can shift the epidemic from concentrated to generalized. Vulnerable populations may be included in behavioural surveillance in order to monitor trends in exposure to risk behaviour and practices.

Groups that are often included in behavioural surveillance are shown in Table 1.2.

Table 1.2 Populations often included in behavioural surveillance

Groups often included in behavioural surveillance
• injecting drug users
• sex workers and their clients
• men who have sex with men
• uniformed people, such as police, border personnel and military persons
• transport workers
• young people
• prisoners

Discussing the table

Looking at Table 1.2, discuss the following questions:

1. Which of the groups listed above are most at risk or vulnerable in your country?
2. Are there groups missing from the table?

Selection of groups for surveys

Introduction

Here are two frequently asked questions about how to choose the type of population for behavioural surveillance in different epidemics:

- If our country has concentrated epidemics and the general population is at low risk, should we do behavioural surveillance of the general population?
- If our country has a generalized epidemic, should we do surveillance of most-at-risk populations?

Surveillance of the general population in a concentrated epidemic:

In a concentrated epidemic, behavioural surveillance of the general population helps you to understand how HIV can spread beyond the groups in which it is concentrated. Therefore, surveillance is an early warning system that lets you explore the following:

- the size of the subpopulations in which HIV is concentrated
- links (bridges) among the subpopulations, their sexual partners and the general population
- the amount of risk behaviours in the general population.

Surveillance of most-at-risk populations in a generalized epidemic:

In a generalized epidemic, surveillance of most-at-risk populations is important because not everyone is at equal risk for HIV infection or has an equal role in spreading or continuing the epidemic. The behaviours of most-at-risk populations and their sexual partners may continue to drive an epidemic even after it has become generalized.

Surveillance of the general population in low-level epidemics:

In a low-level epidemic, if countries have the opportunity to include HIV behavioural indicators in other surveys of the general population, they may opt to do so.

Where to contact groups

The populations included in surveillance can be contacted either in the community or at sentinel sites.

Community sites are locations in the community where people gather, such as households or schools.

Sentinel sites are facilities where groups at low and/or high risk for HIV infection may be found. These may include STI clinics, antenatal care clinics, blood donation centres, drug treatment facilities, prisons and needle-exchange sites.

Linking biological and behavioural surveillance data

We have already discussed the importance of using behavioural data to interpret biological data over time. It is also important to plan behavioural and biological surveillance together so that we can be sure the data match and are useful.

When you are planning surveillance, decide the best way for behavioural and biological data to be linked. This decision can be difficult. Linking means different things to different people. It may mean:

- collecting HIV, STI and behavioural data from the same individuals at the same time
- collecting HIV, STI and behavioural data from the same source population at different times
- analysing HIV, STI and behavioural data from similar source populations, using whatever data are available
- reporting behavioural and biological surveillance together.

All of these methods are useful, but each has advantages and disadvantages. The decision about the best way to link data should be specific to the country. At the very least, behavioural and biological surveillance should be presented in a single report. Table 1.3 outlines the advantages and disadvantages of each type of linking.

Whatever type of linking a country selects, it should be done for trends over time rather than for one point in time. Behavioural and biological data linked at one point in time are impossible to interpret because of the relation between HIV and risk behaviours over time. For example, recent risk behaviours may not be related to a person's HIV status because he or she may have become infected a long time ago.

Table 1.3 Advantages and disadvantages of different types of data linking

Type of linking	Advantages	Disadvantages
Collect data from same individuals	● biological data are viewed over time, and behavioural and biological samples are alike ● behavioural and biological data can each be used to reinforce the findings of the other ● can be less expensive if data collection activities are combined ● can be more convenient, especially where access to groups is an issue ● only disrupts the community once	● time between risk behaviour and infection makes analysis less accurate ● blood draw and longer interviews may increase refusal rates ● required frequency, sample size and sampling strategies may be too different ● may be difficult to find interviewers with the skills to collect both types of data
Collect data from same source population	● helps us understand biological results over time as samples likely to be similar ● behavioural and biological data can each be used to reinforce the findings of the other ● no need to compromise on interviewers' skills, frequency of surveillance, sample size or sampling strategies ● refusal rate for behavioural surveillance not increased by blood being taken for biological surveillance	● time between behaviour and infection makes analysis less accurate ● no guarantee that behavioural and biological surveillance samples are alike ● no savings in cost or convenience ● community disrupted more than once
Analyse data from similar populations	● helps us understand biological results over time ● behavioural and biological data can each be used to reinforce the findings of the other ● no need to compromise on interviewers' skills, frequency of surveillance, sample size or sampling strategies ● refusal rate for behavioural surveillance not increased by blood being drawn for biological surveillance	● analysis at the level of the individual is not possible ● using different source populations means that behavioural and biological surveillance samples may not be alike ● no savings in cost or convenience
Produce a joint report	● provides data users with all the information they need in one document— no need to compromise on interviewer skills, frequency of surveillance, sample size or sampling strategies ● refusal rate for behavioural surveillance not increased by blood being drawn for biological surveillance	● unlikely that the populations are similar, so no integrated analysis ● no savings in cost or convenience ● behavioural data add little to the findings of biological surveillance

Discussing the table

Looking at Table 1.3, discuss the following questions:

    a.   What are some of the advantages of collecting data from the same source population?

    b.   What are your ideas for linking in your country? Please explain.

How to know if the type of surveillance is correct

There is no way to design a surveillance system that applies to all situations. The design should be made to fit the features of the HIV epidemic and the health care system in your country. In the same way, the things that we discuss in this training are not "rules". What you learn must be adapted for your own needs.

Steps for conducting behavioural surveillance

The nine steps

In this section, you will learn how the different parts of the survey interlink, and you will obtain an overview of the surveillance process. The steps for conducting behavioural surveillance are listed below:

1. Create a committee to organize the surveillance.
2. Agree on the purpose of the surveillance.
3. Identify methods for selecting populations and geographical areas.
4. Use formative assessment to help you decide about populations and geographical areas and to guide you in starting the survey.
5. Finish selecting subpopulations and geographical locations.
6. Develop a set of rules for the survey. Define the sampling frame and choose a sampling method.
7. Begin collecting data.
8. Analyse and use data.
9. Plan for the next round of surveillance.

Step 1: Create a committee to organize surveillance

The purpose of the committee is to guide and serve as a decision-making group for the surveillance. Its responsibilities include:

- defining the purpose of the surveillance
- making sure that the surveillance will meet the country's need for data
- identifying sources of funding
- convincing policy-makers and stakeholders about the importance of surveillance
- bringing surveillance partners together
- making final decisions about which populations and geographical areas to include in the surveillance
- monitoring the progress of each step of the surveillance
- helping interpret the data and form conclusions
- helping stakeholders agree about what the data mean
- making sure data are distributed and used.

Surveillance committees already may exist in some countries but may not include all the right people or may not be appropriate for each country's needs and context. Ideally, the Ministry of Health or National AIDS Control Programme should create these committees.

The committees should include members of the various national and international groups that have an interest in the surveillance system, such as representatives from:

- government ministries
- local and/or international nongovernmental organizations
- members of the target population and the nongovernmental organizations that represent them
- agencies selected to do the surveillance.

Step 2: Agree on the purpose of the surveillance

The first task of the surveillance committee is to agree on the purpose of surveillance and to find out what data the country needs. This important task can be difficult because committee members may not agree.

Agreement involves regular meetings and discussions among many groups, which can be difficult to organize.

Step 3: Identify methods for selecting populations and geographical areas

Before each round of surveillance, you should review the populations and geographical areas you will survey and make changes as needed. You may need to include new risk groups or remove groups if their data are no longer useful. All of this should be done using set methods.

Selecting populations: You should use the best evidence for the type of epidemic in your country and know how the epidemic works. The populations you select should be useful for monitoring and tracking the future course of the epidemic.

Selecting geographical areas: Choose the area based on what level of data will be most useful—the national or subnational level. It is important to remember that you can combine data from provinces to get a national figure, but you cannot take national data apart to get provincial figures. The main problem of collecting data at the subnational level is that it requires larger sample sizes and thus more time and money.

Step 4: Use formative assessment

Select your population by reviewing previous research to find what is currently known about the national epidemic. Make a list of potential subpopulations and geographical sites. Do qualitative assessments of the subpopulations you chose to determine:

- whether members engage in risky behaviours
- whether the subpopulation is large enough to be included in the surveillance
- how and where members can be found and sampled
- what defines the subpopulation and what are the eligibility criteria
- whether the members are willing to participate in the surveys.

Step 5:   Finish selection

All the information collected in the formative assessment should be combined so the surveillance committee can select the final population and geographical area.

Step 6:   Develop rules for the survey

You must create protocols or rules for the surveillance. Surveillance protocols make certain that consistent methods are used. This results in trend data that can be compared across studies, regardless of changes in staff between rounds of data collection. Protocols will also help you clarify the aims of the surveillance system and help you make sure that you have planned for all the elements of surveillance.

The elements that should be included in the surveillance are shown in Table 1.4.

Several terms and concepts used in Table 1.4 are defined below:

- Operational definitions of the target population are definitions that are useful specifically in sampling and fieldwork for surveillance. For example, the definition of a sex worker should clearly define what a sex worker is in length of time selling sex, form of payment, type of venue and so forth.
- An indicator is a behaviour that is measured by the surveillance system and that is essential to the spread of HIV. Indicators provide a way to track changes in behaviours over time and compare risk behaviours among different populations. Core indicators were developed by the United Nations General Assembly Special Session on HIV/AIDS (UNGASS) to help monitor global progress against HIV/AIDS. The exact wording of the indicators and further information are available online at http://data.unaids.org/pub/manual/2009/jc1676_core_indicators_2009_en.pdf.
- The appropriate sampling method. A good method must be chosen because usually we cannot interview all members of our chosen populations. Instead, we sample individuals and interview them in several ways. We will discuss sampling in detail in Unit 5.
- Define the sampling frame. Some sampling approaches require a list of units, or a sampling frame, from which the sample is selected. If a sampling frame is needed, you may need to map the population. This mapping can also be done during the formative pre-survey assessment, which is described in more detail in Unit 3.

Step 7:   Begin collecting data

The next step is to carry out the surveillance by starting to collect data. Some of the protocols may change during surveillance. It is important to make notes of any change so that you can collect data in the same way during the next round.

As the protocols are used, discuss with interviewers and supervisors the problems they had while collecting data and the solutions to the problems. This will improve data collection for the next round.

If you are doing behavioural surveillance for the first time, it is always a good idea to pilot test your protocol and survey. In this way, you can find possible problems that could not be found otherwise. Also, you can estimate the workload for your interviewers, test that the locations and times of interviews are appropriate, and improve how you collect data.

Table 1.4 The necessary elements of behavioural surveillance

Element	Description
1	A brief description of HIV/AIDS in the country and what surveillance system exists
2	The goals of the surveillance system, including how surveillance is made part of the national strategic plan for monitoring and evaluating HIV programmes
3	A reason for the selection of subpopulations and geographical locations. How often the surveillance will take place
4	A description of the methods to be used:
	• operational definitions of the target population (see next page)
	• definitions of important indicators (see next page)
	• the appropriate sampling method (see Unit 5)
	• definitions of the sampling frame (see Unit 5 and next page)
	• calculations of the sample size for each target population (see Unit 5 and next page)
	• development and validation of tools and instruments
	• selection and training of staff
	• data collection procedures
	• laboratory test kits and performance of tests
	• quality control
5	A plan for managing and analysing data
6	A plan for writing reports and distributing data. Also, a description of how the data should be used
7	A discussion of the ethics involved, including consent forms (see Unit 7)
8	The responsibilities of surveillance partners
9	The timeline
10	The budget and available resources

Step 8: Analyse and use data

You should have a plan to analyse data before you collect it. The plan should:

- identify personnel with appropriate skills in data management, data cleaning, and data analysis
- identify what computer software you will use
- define indicators
- include examples of important tables and graphs.

Using an analysis plan helps you make sure that all necessary data are collected and that unnecessary data are not collected. It also helps you make sure that data are analysed properly and are ready for timely distribution.

Collecting the data is not the end of surveillance. There is little point in collecting useful data that are not analysed and used. Data must also be easily understood and interpreted, and distributed quickly at national, regional and local levels. For this to happen, you must plan the analysis, distribution and use of the data at the beginning of surveillance.

Step 9: Plan for the next round

Although the surveillance steps have been presented as if they happen in a specific order, the process should actually be viewed as a circle (Figure 1.1). The purpose of surveillance is to collect trend data, which requires that similar methods be used for each round. However, surveillance systems must also be modified for new knowledge and other changes that will come with doing the surveillance. You must balance these two goals.
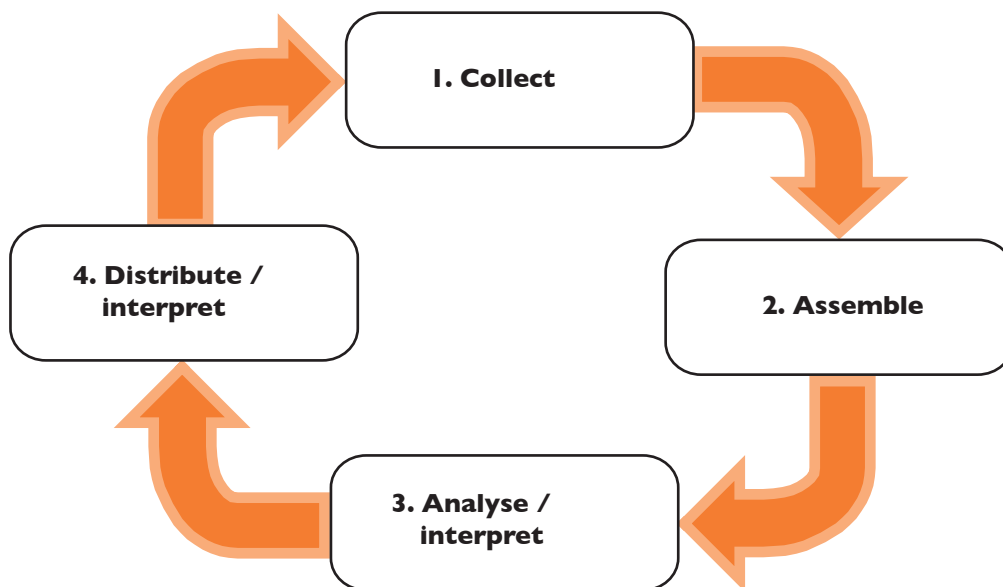
Figure 1.1 The flow of surveillance data

Summary

Surveillance is the systematic, regular and ongoing collection and use of data leading to interventions for public health. Behavioural surveillance involves repeated cross-sectional surveys and collecting data on HIV risk behaviours and other relevant issues that can be compared over time. When you design a behavioural surveillance system, there are several things you should consider: whom to include in surveillance, where to contact the populations, how to link data from biological and behavioural surveillance and how to be sure that the surveillance system is appropriate for the setting. You also should follow the steps required to build a surveillance system that can operate in a continuous cycle of collecting, assembling, analysing and interpreting data.

Unit 1 exercises

Warm-up review

Take a few minutes now to look back at your answers for the warm-up questions at the beginning of the unit. Make any changes you want to make.

Small group discussion

Get into small groups by country, region or district to discuss these questions.

1. How is surveillance organized in your setting? What works and what does not work?

2. When discussing guidelines for whom to include in surveillance for each level of epidemic, what do we mean by "the general population" and by "the most-at-risk population"?

Apply what you have learned/case study

Try this case study individually. We will discuss the answers in class.

1. Name an HIV most-at-risk population in your country. Describe the group and explain why you think it is at high risk for HIV infection.

2. Look at the diagram in Figure 1.2 which shows how the HIV epidemic works in a hypothetical community. The diagram helps us to understand the behaviours that drive the HIV epidemic, and the most-at-risk populations and their links to the general population. The ovals show the different population groups, and overlapped ovals mean some may belong to more than one group. The arrows show the links between the populations, the thickness of arrows shows the strength of the links, and dashed arrows represent potential links.

   a. Discuss the diagram and identify the most-at-risk populations, bridging populations and low-risk populations.

   b. Try to use data from the most-at-risk populations in your country and make a similar diagram. Is there any information lacking in your country that could help you to describe the situation?
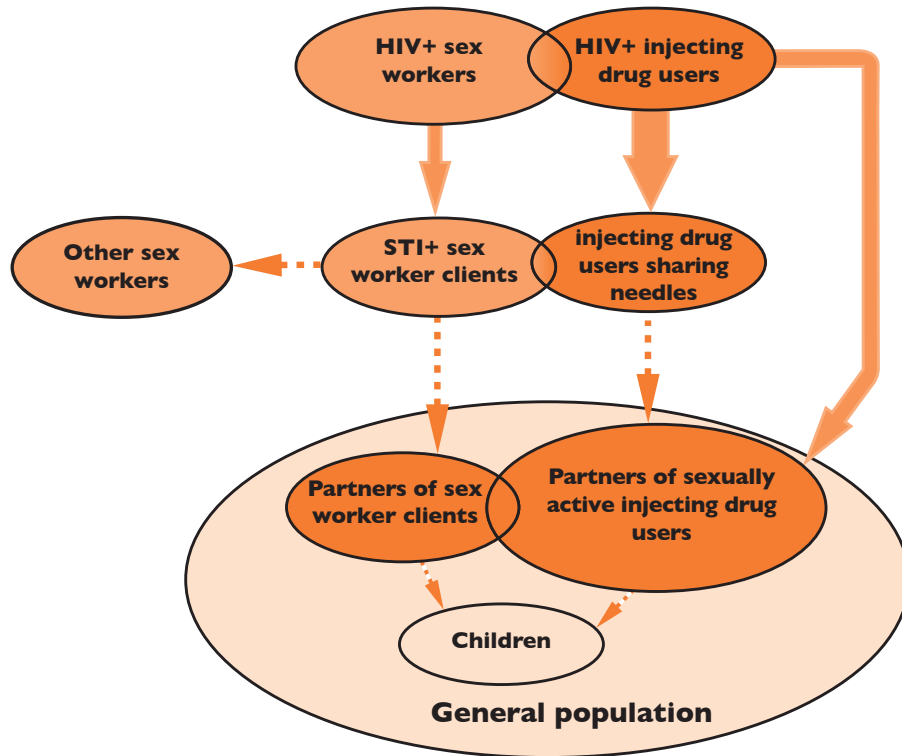
Figure 1.2 How the HIV epidemic works in a hypothetical community

Unit 2

Measures and
indicators for
behavioural
surveillance

Overview

What this unit is about

IN THIS UNIT, you will learn about selecting appropriate measures and indicators for behavioural surveillance.

Warm-up questions

1. What are two characteristics of a good indicator?

2. True or false? Behavioural surveillance indicators should measure aspects of behaviour that are necessary to spread HIV. Circle your answer below.

    True   False

3. When should indicators be selected during behavioural surveillance?
    a. during planning
    b. during analysis.

4. True or false? The two most difficult issues in defining indicators for behavioural surveillance are defining the behaviours and defining the time period to which the indicator should refer.

    True   False

5. True or false? Behavioural surveillance indicators do not need to be consistent over time.

    True   False

6. Fill in the blank spaces using one of the answers given below.

An indicator's reference period should be determined based on _____ of a risk behaviour that we want to measure. For instance, to accurately measure a very _____ risk behaviour, a _____ reference period should be used.

a. frequency, frequent, shorter

b. frequency, frequent, wider

c. importance, important, shorter

d. importance, important, wider

Introduction

What you will learn

By the end of this unit, you should be able to:

- understand the characteristics of a good indicator
- select indicators for behavioural surveillance
- understand the main problems with indicators
- discuss surveillance comparing standardized indicators and locally-adapted indicators.

Good indicators

An indicator is simply a measure of something. A good indicator:

1. measures something related to the surveillance
2. measures something accurately
3. is easy to interpret and is defined clearly. For example, if we were interested in measuring unsafe sex, we would need to:
   - define unsafe sex, such as "sex without using a condom with a sexual partner who is a sex worker"
   - define the population to which the measure can be generalized, such as "men aged 15–49 years"
   - define the time period to which the measure refers, such as "last sex".
4. can be compared across different population groups and across time, so definitions and field methods must be consistent over time
5. is reasonable to collect in terms of effort and cost.

Selecting indicators for behavioural surveillance

Indicators for behavioural surveillance

Indicators for behavioural surveillance should measure behaviours that are needed to spread HIV, including:

- behaviours that determine the likelihood that an uninfected person will come into contact with an infected person. Examples of these behaviours are number and type of sexual partners and patterns of needle exchange
- behaviours that determine the likelihood that transmission of HIV will occur if contact with an HIV-infected person occurs. Examples of these behaviours are level of condom use and practices for sharing injection equipment. The likelihood of transmission is also determined by other factors, such as the presence of other STIs.

Indicators should be selected during the planning of behavioural surveillance to focus on the country's data needs. It should also be ensured that all the data needed to construct the selected indicators are collected.

Assessing HIV risk

Select indicators based on the data that the country needs. Examples of essential indicators for the general population, injecting drug users, sex workers and men who have sex with men are shown in Table 2.1, below. You can learn about other common indicators online at http://data.unaids.org/pub/manual/2009/jc1676_core_indicators_2009_en.pdf.

Collect data on factors like alcohol or drugs that affect high-risk behaviours. Also, collect data on background information, so that indicators can be compared across different demographic profiles.

Table 2.1 Essential indicators for risk factors

General population	Injecting drug users	Sex workers	Men who have sex with men
Percentage of women and men aged 15–49 years who have had sexual intercourse with more than one sexual partner in the last 12 months	Percentage of injecting drug users who report the use of sterile injecting equipment the last time they injected	Percentage of sex workers who report the use of a condom with their most recent client	Percentage of men who report the use of a condom the last time they had anal sex with a male partner

Two issues in defining indicators

The two most difficult issues in defining indicators for behavioural surveillance are defining the behaviours and defining the reference period (time period) to which the indicators refer.

Defining behaviours

A good indicator is easy to interpret and is defined in clear terms. What may appear to be a small difference in a definition can translate into large differences in what an indicator represents, how it can be used and the data needed to create it.

For example, an indicator that measures the frequency of visits to sex workers by married men who have sexual intercourse with sex workers could be defined as the proportion of:

- married men who visit sex workers regularly
- men who visit sex workers regularly who are married.

The first indicator gives us an idea about the size of the link between sex workers and married men. If a large proportion of married men visit sex workers, then commercial sex could be important in driving the epidemic. However, this indicator lacks a clear definition for frequency of visit. "Regularly" is not clear enough and could be interpreted in different ways.

The second indicator does not tell us anything about the role of sex workers in driving the HIV epidemic. We do not know the size of the link between sex workers and married men, only that among men who visit sex workers, a certain proportion are married.

You should also be consistent with definitions across time so that you can measure trends. For example, defining consistent condom use among sex workers as "condom use during every sex act in the last week" in the first round of surveillance and as "condom use with every client in the last week" in the second round would result in indicators that cannot be compared.

Defining reference (time) periods

An indicator's reference (time) period can affect how accurately the indicator is measured. The reference period should be determined by the frequency of a behaviour. But it should be kept short to increase the likelihood that respondents will remember their behaviours. If a behaviour is very common, respondents may have difficulty remembering how they behaved. A shorter reference period may be better. If the behaviour is very rare, measurement over a short reference period may not capture the experience very well.

For example, it would be difficult for most current injecting drug users to remember how many times they injected in the last year. If we used one year as a reference period, we would get inaccurate information.

If we used a reference period that was too short (for example, in the last hour) most injecting drug users would remember accurately how many times they had injected. But most people would respond "not at all" and we would not get a good idea of the risk behaviours.

Standardized versus locally-adapted indicators

We have already discussed the importance of using standardized indicators over time so countries can measure trends. It can also be important to use a basic set of standardized indicators across regions or countries to make comparisons.

The behavioural indicators developed by UNGASS are recommended for use in EMR/MENA. Generally, the needs of each country should come before international needs. Indicators should be appropriate locally in terms of reference periods and definitions or they will be of little use. Ideally, countries will collect both locally-adapted indicators and the basic set of standard UNGASS indicators.

Summary

Behavioural surveillance indicators should measure behaviours that are needed to spread HIV. Indicators should be selected when you plan behavioural surveillance to focus on the country's data needs. Selecting indicators during planning also makes certain that all the data needed to construct the indicators are collected. In addition, it can be important to use standardized indicators across regions or countries to allow comparisons.

Unit 2 exercises

Warm-up review

Take a few minutes now to look back at your answers for the warm-up questions at the beginning of the unit. Make any changes you want to make.

Small group discussion

Get into small groups by country, region or district to discuss these questions.

1. What would be a good behavioural indicator to measure the likelihood of HIV transmission through sexual intercourse in the general population of your country?

2. Discuss whether these definitions are meaningful for your country. Explain which specific populations they refer to.
    a. steady sexual partners
    b. commercial sex

3. Can an unmarried person have a steady sexual partner? What if they do not live together?

Apply what you have learned/case study

Try this case study individually. We will discuss the answers in class.

1. Compare the indicators listed below and discuss the advantages and disadvantages of each one. In general, which one is more useful?
    a. Proportion of men who have sex with men who had anal sex without condoms with other men in their last sexual contact
    b. Proportion of men who have reported unprotected sex with other men in the last month

2. Discuss the differences between substance users and illicit drug users.

Unit 3

Formative assessment

Overview

What this unit is about

IN THIS UNIT, you will learn about the theoretical and practical knowledge needed to plan behavioural surveillance activities. You will learn what information should be collected during the pre-surveillance activities and what methods you can use. This is called formative assessment.

Warm-up questions

1. Which of the following is an example of a question you should answer to select geographical areas during formative assessment?
   a. What are the regional differences in HIV transmission and risk behaviour?
   b. Where are the interventions located?
   c. How much money and staff are available for surveillance?
   d. All of the above.

2. True or false? For some methods of formative assessment, you need to review existing surveillance data and published and unpublished literature. Also, you should talk to people who are knowledgeable about the epidemic. Circle your answer below.

   True   False

3. Methods used in _____ research include large sample size, random samples and shorter interviews.
   Methods used in _____ research include unstructured questionnaires, lengthy interviews and fewer well-trained interviewers.
   a. qualitative
   b. quantitative

4. Suppose you want to determine areas of a city where street-based sex work occurs. Which of the following is not an appropriate method to use during a formative assessment?

    a. literature/data review

    b. in-depth qualitative interviews

    c. general population survey

    d. all are appropriate

5. List three advantages to building alliances.

    a.

    b.

    c.

Introduction

What you will learn

By the end of this unit, you should be able to:

- identify and understand the purpose of formative assessment
- describe and select the methods used in formative assessment.

What is formative assessment?

Formative assessment is made up of many activities in which we use multiple qualitative research methods to collect the information needed to begin an effective HIV surveillance system. Doing a formative assessment, which is sometimes called pre-surveillance assessment, allows us to:

- select the populations to include in surveillance activities
- determine how practical it is to do surveillance in certain populations and geographical areas
- choose the geographical area for surveillance activities
- choose a sampling method to study the population
- define what makes members of the population eligible for inclusion in the survey.

To meet these goals of formative assessment, think of a list of questions that must be answered by the surveillance. This list will not be complete, and it is likely that you will add and remove questions as you go.

Select the populations

To select populations to be surveyed, you must understand how the HIV epidemic works in a country. Answering the following questions will help:

- Should surveillance concentrate on subpopulations, the general population or both?
- Who are the most-at-risk populations?
- For each group, how do the behaviours affect the epidemic? This can be answered by finding out the size of the group, how they are linked to other populations, what types of HIV risk behaviours members have, with whom they do them and how often?
- How will surveys of these populations add to our understanding of the epidemic?
- What interventions have been used or are planned among the populations?
- Is the population easy to contact for surveillance?

Determine if surveillance is practical

A successful surveillance system needs a practical plan to collect data. A good plan will help you get better involvement from populations and will help you in your fieldwork. To create a plan, several questions must be answered using formative assessment:

- Whose permission is needed to do surveillance?
- Who can assist us with the survey and make sure it runs smoothly?
- In what language should the interviews be done?

- What skills should interviewers have to make sure we get the most valid results?
- What are the locations where we could collect surveillance data privately?
- What are the best times of day to interview participants so that they will have time to complete the interview?
- What practical problems may interviewers have? For example, how will they assure participants of privacy and safety?
- How long will the surveillance activity take?
- Are skilled staff members available to do surveys?
- Can we get enough money and resources to complete surveillance? What would our budget be?

Choose geographical areas

You should select the area based on whether data are most meaningful if they are collected at a national or subnational level. Important questions must be answered before you select geographical areas:

- What is already known about the epidemic in different regions?
- What are the differences in terms of HIV transmission and risk behaviour in these regions?
- Where are the most-at-risk populations concentrated?
- What is the estimated size of the most-at-risk populations in each geographical area?
- Are there cross-border or internal transportation routes where transmission may occur?
- Are the data needs at the national or subnational level? For instance, do the data satisfy national or UNGASS reporting requirements?
- To what level do we want to generalize the data?
- Can some areas be proxies for others?
- Where are the intervention programmes located?
- Which locations currently have biological surveillance?
- How much money and staff are available for surveillance activities?

Choose a sampling method

When you are doing behavioural surveillance among most-at-risk populations that are hard to reach, such as injecting drug users, men who have sex with men and sex workers, probability-based sampling techniques such as simple random sampling or cluster random sampling are usually not practical.

Answers to these important questions will help you choose the best non-probability sampling methods for your survey:

- Who are the gatekeepers of the population?
- Does a large proportion of the group gather at certain locations? Are group members easily reached there?
- What are the best times of the day to find group members at the locations?
- Do group members know each other? Are they part of a network?
- Can the same people be found at more than one location?
- Is it possible to recognize people at the locations as members of the group?

Define groups and set eligibility criteria

One of the biggest mistakes in surveillance is not tracking populations consistently over time. This sometimes happens because a population is not defined clearly.

For example, there are many ways to define sex workers so it is important to have one clear definition of a sex worker in your survey. It is also necessary to have specific eligibility criteria. You must answer important questions to define populations and set eligibility criteria:

- Are the risk behaviours different within a population? For example, how, when and where do sex workers work? From apartments during the day or on the street at night?
- Can the group be divided by differences in their behaviours?
- What characteristics can help you identify group members?

Formative assessment methods

Key steps

The following steps should be conducted as part of formative assessment:

1. Assess what is currently known about the HIV epidemic or subepidemic in your country. This step involves reviewing existing surveillance data and published and unpublished literature. It also means talking to people who are knowledgeable about the epidemic. They can tell you about potential hotspots and at-risk or vulnerable populations. Consider gathering information on risk behaviours and HIV in the general population using information from existing surveys.
2. Do the formative assessment using qualitative assessment methods and mapping. This step involves fieldwork to find hotspots and to gather information. You can use the information to define populations that will be included in surveillance and to guide surveillance activities.

Review existing data

The purpose of reviewing available research is to find existing data, which you can use to understand HIV in the target population and geographical area. It can also show whether your target population is different from the population described in other sources. You may review:

- peer-reviewed scientific literature
- abstracts from regional and international AIDS conferences
- grey literature (printed but not necessarily widely published reports, such as programme evaluations, government surveillance reports, programme planning documents)
- basic surveillance data.

Peer-reviewed literature can be located using internet-based search engines, such as Entrez Pub Med from the U.S. National Library of Medicine (NLM) (available at www.pubmed. gov). Also available is NLM Gateway, which includes abstracts from the International AIDS Conferences and AIDSLINE (gateway.nlm.nih.gov), and the U.S. Bureau of the Census

HIV and AIDS surveillance site for prevalence studies and reports (www.census.gov/ipc/www/hivaidsn.html).

Grey literature can be found using online resources, including UNAIDS' epidemiological fact sheets about countries with HIV/AIDS prevention programmes (www.unaids.org). Others sites are the WHO's Communicable Disease Global Atlas (globalatlas.who.int) and Family Health International's guidelines for use of behavioural surveillance surveys (www.fhi.org/en/topics/bss).

Governments, donors or nongovernmental organizations produce monitoring and evaluation reports which sometimes can be accessed through the internet. Often, you will need to contact governmental officials, representatives of the donors, or nongovernmental organizations to get copies. You can also seek subnational data from local organizations.

Qualitative research

Qualitative research focuses on the characteristics or quality of things, and how and why things exist. Quantitative research focuses on the number, size or frequency of things.

The reasons why certain behaviours are present in a group or a certain area of a country cannot be obtained through statistics and numbers. Even countable things often fall under qualitative categories or definitions.

For questions we want to answer as part of our formative assessment, you do not need precise and generalizable quantitative measures of how many people have certain knowledge or perform certain behaviours. The surveillance itself can do this. Rather, you need in-depth qualitative knowledge of such things as:

- the performance of a behaviour (for example, sex work or injecting drugs)
- how that behaviour is organized in a group
- what people with similar behaviours or in similar groups do
- how and why people do certain things.

This information can give you the following:

- background information associated with the behaviour
- social and structural components of the behaviour in a group
- what the participants and larger society think about the behaviour
- characteristics of the people who perform the behaviour.

Equally, qualitative research can help explain how people think about the risk associated with the behaviour. It can explain how both the individual and the group deny or change their behaviour to accommodate the risk.

Researchers often discuss the need to present an insider's view (an "emic" or culturally-specific view) to understand what people mean when they reply to questions in an interview. Behaviours may make sense to an insider, but may not appear to make any sense from the outside. Therefore, it is important to understand the culture and point of view of the people in the group to correctly interpret qualitative data.

Difference between qualitative and quantitative methods

The differences between qualitative and quantitative research lead to differences in methods and approaches, as shown in Table 3.1. Some of these differences are not absolute, especially when both methods are used together.

Table 3.1 Differences between qualitative and quantitative methods

Quantitative	Qualitative
Large sample size	Small sample size
Random sample	Purposeful sample
Calculated sample size	Theoretical sample size
Structured questionnaire	Unstructured questionnaire
Questions asked exactly as written on questionnaire	Interviewers are encouraged to explore questions with participants
Large number of closed-ended questions	Smaller number of open-ended questions
More interviewers who need fewer special skills	Fewer interviewers, but they need to be specially trained

Below, we elaborate on three aspects of qualitative research methods:

- purposeful sampling
- theoretical sampling
- unstructured and semi-structured questionnaires.

Why purposeful sampling?

We purposefully select a sample of people in qualitative research because it does not help us to select them at random from the population. We would be unable to get a large amount of specific information, such as definitions of the different types of sex work so that we can divide sex workers into subpopulations. Instead we use key informants, such as sex workers or their clients.

Once you have decided on the key informants, you should select them for their abilities to be interviewed, including having knowledge about the topic, being easy to talk to, understanding the information you need and being willing to give it to you.

Theoretical (saturation) sampling

Unlike in quantitative surveys, the sample size in qualitative research is not calculated before data collection. Instead, it is determined by theoretical or saturation sampling. This involves continuing to interview informants until you cannot get more new information from them.

For example, we could ask informants to list all the different types of places they know where sex workers sell sex. When informants are not providing new information, an adequate sample size has been reached.

Open-ended semi-structured interviews

Once the different types of sex work have been identified, we need to determine if they should be considered separate surveillance groups. For this, we need to understand how the different types of sex workers or different places where they operate affect HIV risk and transmission. This is best explored using open-ended semi-structured interviews.

Open-ended questions have no set answers for participants to select. Instead, they must answer in their own words. This means topics can be discussed in more depth.

Probing is often necessary to avoid incomplete or unclear answers. Open-ended interviews tend to be much longer and require more skill than interviews for surveys with very large sample sizes.

A semi-structured interview is one that uses a written list of open-ended questions and topics that need to be covered, although not necessarily in a particular order. Interviewers are encouraged to depart from the guide to follow new leads. These methods are in contrast to methods in quantitative research, which tries to make sure that each informant is asked exactly the same questions in exactly the same way.

Qualitative data collection methods

Introduction

Qualitative data can be collected through a variety of methods for pre-surveillance activities such as:

- in-depth interviews with key informants
- expert interviews
- focus groups
- direct observation
- rapid assessment.

In-depth interviews with key informants

In-depth interviews are unstructured or semi-structured interviews. They use open-ended questions, which we have discussed above. The looser structure lets respondents talk about themselves and a wide range of issues related to the topic. In-depth interviews can last over several sessions and the key informants often become informal assistants to the researcher. These informants are members of the most-at-risk populations or are other knowledgeable insiders rather than external experts, such as government officials.

Expert interviews

Semi-structured open-ended questions are also used for interviews with experts. For example, you may need to interview colleagues involved in behavioural and biological surveillance from local offices of the Ministry of Health or nongovernmental organizations, clinics and donor agencies.

Focus groups

Like in-depth interviews, focus group interviews use open-ended questions, but the interviews are more structured. Also, focus groups are group interviews that stimulate discussion so the researchers can listen to how ideas about the topic are expressed. The primary goals are to:

- define the views, motivations and needs of the focus group
- generate ideas for interventions
- test survey materials
- determine social rules about things focus group members can talk about in front of each other.

Because the interviews are public, it may be difficult for a respondent to express private thoughts or talk about intimate behaviours. Similarly, the group interview can be manipulated by influential members and responses can be skewed.

To encourage openness in focus groups, researchers gather people from similar backgrounds or experiences to discuss a specific topic. A group usually consists of 8–12 people, is guided by a facilitator and has a record taker who takes notes. Because of the relationships of power among people in a group, facilitators need to have special training and be quite skilled.

Table 3.2 shows questions you may have about formative assessment and the methods you may choose to find the answers. To be accurate, it is best to use information from more than one source for each question.

Table 3.2 Sample questions about formative assessment

Question	How to find the answers
How can we find the venues or locate surveillance sites?	• Interview staff of government offices and nongovernmental organizations involved in HIV prevention. • Conduct focus group interviews with target population. • Observe and map areas where the target population spends time.
What kind of incentives should we offer?	• Use focus group interviews and individual interviews with members of the target population to determine type (money, gift, etc.) and size (equivalent to a standard amount of money). • Use expert interviews to learn about the proper amount of incentive for the specific target population. • Interview institutional review board (IRB) experts to make sure the amount you want to offer is not too large and therefore coercive, nor too small and considered by participants to be not worth their effort.

Discussing the table

Looking at Table 3.2, answer the following questions:

a. Have you participated in focus group interviews, conducted one or heard about them where you work? What are useful ways to use each method? What may be some problems associated with each method?

b. How might an incentive that is too high or too low affect the project?

c. Can you think of some other formative assessment questions that you should answer before doing surveillance?

Direct observation

In HIV behavioural surveillance, direct observation means watching members of the target population to find out where and when they gather. It is also used to understand their behaviours in certain environments, and to see how they interact with members of their own population and other populations. For instance, you may observe street-based sex workers to see whether they are organized in any way, such as in small groups that work together or groups that are controlled by pimps. You may also want to find out in what areas of the city they work and the times they begin and end work. Observers should not be involved in the setting that they are observing because it may cause significant change in the way group members behave.

In contrast, in another method called participant observation, the researcher participates in the setting being observed. This method usually requires a good deal of work over a long period of time to give investigators a chance to be accepted as a natural part of the group. For instance, investigators may live with homeless injecting drug users in shelters over several months, interact with them very closely and take notes on their observations and discussions.

Table 3.3 Eight aspects of direct observation

Observations	Questions to ask	Example answers: observing sex workers
Settings	Where does the observation take place? When? What is the physical area like? What objects are present?	A street corner, night
People	Who is present? What type of people are they? How old are they?	Women aged 19 to 35 years
Activities	What are the people doing? Why are they there?	A couple of women standing together, others off on their own, some talking on cell phones
Signs	Are there any clues that provide evidence about meanings and behaviours?	Men walk up to the women and they leave together
Events	Is this a regular occurrence, or is it a special event such as a meeting or a disagreement?	Regular occurrence
Time	In what order do things happen? Is there a reason for this?	Every once in awhile a man walks by, some drive by, they talk and leave together
Goals	What are the people trying to accomplish?	Buy and sell sex
Networks	How do the people present know one another? Is it a social relationship or a meeting set up by an organization/group? Does the relationship change over time?	Some of the women seem friendly with each other while others keep to themselves. They must all know each other or have seen each other before

Table 3.3. is a summary of eight things to consider when using the direct observation method. It also shows answers to questions a researcher might ask while observing sex workers.

Discussing the table

Looking at Table 3.3, answer the following questions:

a. What is the value of observing signs?
b. What aspect of observation addresses relationships?

Purpose of mapping

Mapping identifies locations frequented by the most-at-risk population. During key informant interviews and focus group discussions it is a good idea to draw a map (see Figure 3.1, below). You can also use a large map of the city and have people identify places where the most-at-risk population gathers; these could be night clubs, parks, truck stops, street corners and so on. Investigators have also gone to sites and asked people to identify additional sites their groups frequent. This is repeated until no more sites frequented by the most-at-risk population can be identified.

Reproduced from the PLACE manual, MEASURE Evaluation.

Figure 3.1 Example of a hand-drawn mapping exercise [1]

Mapping can also be useful to:

- identify local HIV/AIDS prevention programmes, STI clinics, methadone clinics and HIV/AIDS treatment and care sites
- track places where people meet new sexual partners or where injecting drug users socialize

- identify the places for prevention services that are most accessible to the population.

You may also want to make a spreadsheet of the venues to go along with your map. This is especially important when preparing for time-location sampling method (TLS). You can describe the people who attend the venue using the following as an example [1]:

- Male-to-female ratio
- Regular patrons
- Where patrons reside
- Do patrons include sex workers, men who have sex with men, military people, mobile individuals, youth, locals, unemployed persons, injecting drug users?
- Do people meet new sex partners at the venue?
- Do people have sex at the site?
- Do people inject drugs at the site?

You can describe the venue itself using the following characteristics as an example:

- type of venue
- busy times
- maximum occupancy
- safety
- number of staff
- sufficient light and privacy (to conduct an interview for TLS).

By interviewing someone knowledgeable about the venue you may be able to gauge the coverage and prevention potential [1], such as condom availability or evidence of HIV prevention activities.

Mapping for respondent-driven sampling

Mapping for respondent-driven sampling (RDS) is useful for planning where the surveillance site will be. It must be in a safe area and accessible to the population to be surveyed. You can find a full description of this method in *Introduction to HIV, AIDS and sexually transmitted infection surveillance. Module 4: Respondent-driven sampling* [2] and at: http://www.respondentdrivensampling.org

Mapping for time-location sampling

Mapping for TLS is the centre of the sampling method. You will spend much time before and during the process to identify venues that a most-at-risk population uses and to enumerate the people attending the venues. Enumeration is the process of counting people at each venue who appear to fit the eligibility criteria of the target population. During formative assessment, you can use enumeration to decide whether venues and time periods contain enough people to be included within sampling frames (generally at least eight people).

When selecting venues, you also should look for a good location to do the interviews, such as private areas at or near the venue.

Rapid assessment tools

Introduction

Several rapid assessment tools have been developed on HIV or drug abuse topics and can be used as a guide for formative assessment. These are not formative assessment tools, but you can take some ideas from them for formative assessment purposes. Three are discussed here.

Rapid assessment guide

The *HIV/AIDS rapid assessment guide* is a tool that provides five sources of data to give a spatial, quantitative and qualitative overview of a project area, including:

- mapping—for a geographic overview of a project area
- site inventory—a detailed enumeration of population, employment, infrastructure, transport routes and health and social services
- ethnographic guide—for a rapid definition of the sexual and social culture of a project area
- focus group guide—for further in-depth qualitative insights
- behavioural surveys—designed to provide essential data on sexual behaviour, condom use and STIs.

The tool is available at:
http://www.fhi.org/en/hivaids/pub/guide/rapidassessmentguide/index.htm

Rapid assessment and response

Rapid assessment and response (RAR) is a method used to assess the nature of a public health problem such as HIV and to bring in the resources needed to solve it. Using RAR, we form new ideas about prevention and think of new interventions to use in the field. In HIV surveillance among most-at-risk populations, RAR has been used widely to assess the spread of HIV among injecting drug users. It has characteristics that make it useful as a formative assessment tool:

- Speed—rapid assessments are completed in far less time than conventional research projects.
- Use of existing data—rapid assessments make use of existing data, and new data are gathered only when existing sources are inadequate.
- Use of multiple indicators and data sources—RAR makes use of multiple data sources and combines them to form conclusions.
- Investigation on many levels of society—problems are examined in social, cultural, religious, political and historical contexts.

When we want to use RAR for formative assessment, the part that is most useful is the rapid assessment component.

PLACE

Priorities for local AIDS control efforts (PLACE) is a method that can be used to identify areas that have high levels of HIV transmission and to focus prevention resources where

new sexual partners are being found within those areas. Although designed as an HIV prevention tool, PLACE also can assess where high-risk sex occurs and what most-at-risk groups may be involved.

The PLACE method collects four types of information useful in formative assessment:

- key-informant interviews
- interviews with knowledgeable people at the sites
- interviews with people socializing at the sites
- maps of site locations.

These data are collected in five discrete steps (see Table 3.4). Following the steps in the PLACE method may be very useful when you plan for the TLS method.

Table 3.4 Steps in the PLACE method

Step	Objective
Prepare	To adapt protocols, if necessary, and obtain community support and ethical approval
1	To identify and describe areas in a city or district likely to have high HIV incidence and select location(s) for implementing full PLACE assessment
2	To identify sites and events within selected areas where people meet new sexual and needle-sharing partners
3	To conduct site visits to verify key informant reports and obtain characteristics about the site and patrons that will help you develop prevention programmes
4	To describe patrons of sites and estimate the rate of new sexual partnerships among people socializing at sites
5	To summarize findings, estimate monitoring indicators, and prepare a map useful for the intervention

Source: [1].

For further information, please refer to:
http://www.cpc.unc.edu/measure/leadership/place.html

Discussing the table

Looking at the table, answer the following questions:

a. What are the preparatory steps for conducting a PLACE assessment?
b. How would you use PLACE? Where would you go? What would you ask?
c. What are four sources of information used to determine where high-risk sex is occurring?

Building alliances

Forming alliances

Alliance building is bringing people together for a purpose that will benefit all the members of the alliance. Alliances emerge when individuals, organizations, businesses or

other groups want to do more than they can do alone. Alliances for surveillance should consist of organizations and people that most-at-risk populations trust.

Surveillance means contacting most-at-risk populations. This can be difficult due to issues such as stigma, marginalization or illicit behaviour. Individuals, groups and organizations often need to work together to conduct the surveillance to:

- more fully understand the culture and diversity of the target population
- identify subgroups
- gain access to the target population
- help begin surveillance activities.

Steps to forming alliances

Building an alliance happens over time. In surveillance activities, you must understand the power structures that influence most-at-risk populations. People in your alliance will give you this information. They can also help you prepare for surveillance and to distribute its results. There are five main steps to building an alliance:

1) Establish a central working group:
   - Build a team of four to 10 members who represent the target population and organizations that have the most to gain from surveillance. These groups should be willing to give you the most resources toward the goal.
   - Duties of team members may be: making policies, goals and strategies for the alliance; recruiting other members; organizing and conducting meetings; evaluating progress; and dissolving the alliance once its mission is done.
2) Develop a general goal on which everyone can agree:
   - It is important to be brief and clear when you state the position of the alliance and describe the desired outcome.
3) Bring in effective members:
   - All participants should be informed of the surveillance and should have an interest in a successful outcome.
   - Members should be respected people who can motivate others. They also should be able to compromise, work well with people and represent their organizations.
   - Invite prospective members, including new and old allies and rivals.
4) Find out the benefits of surveillance for each member of the alliance.
5) Develop a strategy based on the goal and give responsibilities to each member.

Advantages and disadvantages of alliances

Although alliances are necessary to do surveillance, keep in mind some of the advantages and disadvantages (Table 3.5):

Table 3.5 Advantages and disadvantages of alliances

Advantages	Disadvantages
● increases resources	● requires time and skill
● results in more creative solutions	● results are not always obvious
● spreads the workload	● may require extra surveillance training for members
● creates a safer environment	● may be hard to keep people motivated if surveillance planning takes a long time
● raises community awareness	

Identifying stakeholders

You should know who the stakeholders are in your surveillance. Excluding an important stakeholder can undermine the process. It is a good idea to include people from groups that usually are under-represented in this type of planning. Stakeholders and alliance members usually include the following:

- governmental organizations and nongovernmental organizations who run HIV prevention and care programmes
- influential current and former members of the target population
- law enforcement, police and criminal justice staff
- national and international advocacy groups
- current public health prevention programmes
- health authorities
- gatekeepers
- local community leaders
- social welfare and service organizations
- staff at STI clinics.

The list of stakeholders will probably be specific to the project. However, it can overlap from one project to the next as the community becomes more able to participate.

Specific ethical considerations for conducting formative assessment

Possible institutional review board requirements

An institutional review board (IRB) is the ethics committee designated to approve, monitor and review biomedical and behavioural research involving humans with the aim of protecting the rights and welfare of research participants. Principal investigators should follow state and local procedures to decide whether doing formative assessment falls under local regulations for human subjects. If local IRB review is needed, you should get approvals before beginning formative assessment. Any person recruited for a focus group or key informant/expert interview has the right to refuse to participate and may refuse to answer any question. All behavioural surveillance staff must follow ethical principles by protecting the privacy, confidentiality and autonomy of participants.

Potential risks and benefits

Being interviewed for formative assessment has a low risk to participants. Privacy is protected, but some people in a focus group may recognize others. Loss of privacy, however, can also occur outside of surveillance; for instance, a person may be recognized when he or she joins in HIV prevention activities or sees health-care providers in public or clinical settings.

The benefit to participants can come from better knowing their own risks for HIV/STIs, learning more about local HIV prevention efforts, and getting prevention materials and referrals for health care. Participation will also benefit the wider community by helping in the planning of local HIV services and prevention efforts.

Summary

Before a surveillance system can begin, formative assessment must be done. In formative assessment, researchers or public health practitioners define and reach the target population. This is when they begin to understand the characteristics of the population related to specific public health problems.

Qualitative research methods let us understand how behaviours in different groups affect the HIV epidemic. It also helps us plan for surveillance activities. To collect information during formative assessment, there are several methods you can use, such as reviewing available research, holding focus groups and in-depth interviews, and observing the target population and people who work with them to learn more about them before starting the surveillance.

Also, formative assessment can help you decide what sampling methods are right for the population you plan to study. It can help you identify biological and behavioural indicators and decide what ethical issues are important.

Unit 3 exercises

Warm-up review

Take a few minutes now to look back at your answers for the warm-up questions at the beginning of the unit. Make any changes you want to make.

Small group discussion

Get into small groups by country, region or district to discuss these questions.

1. What groups in your country, region or district would be considered most at risk for HIV infection? Why? What links do they have to other populations?

2. For each most-at-risk group you have identified above, indicate the following, based on your country, region or district:

    a. Is the population easy to identify and contact?

    b. Where would you find these people? List two places.

    c. Do group members know each other or are they part of a network?

    d. Who are the gatekeepers for these most-at-risk populations?

    e. Are there any risks in including this group in surveillance? If so, how can the risks be minimized?

3. Outline how you would use qualitative research methods to do formative assessment in one of these populations in your country, region or district.

Apply what you have learned/case study

Try this case study individually. We will discuss the answers in class.

You are the HIV surveillance officer in Menaland, a country in EMR/MENA that has a concentrated epidemic based on a consistent HIV prevalence of 5% in STI-clinic patients. You decide that you would like to start surveillance among sex workers in the capital city, but first you must do a formative assessment to define and understand the characteristics of sex workers and to find ways of accessing the population.

To date, the National AIDS Programme has not done any surveillance or prevention activities in this population, so very little data exist. To get information, you decide to form an alliance with community stakeholders who work closely with the population.

a. Who would these stakeholders be?

b. How would you identify and contact them?

c. How would you establish alliances?

d. Once you have established alliances, what are the steps that you would take to do formative assessment before starting surveillance activities?

e. Describe the benefits of using PLACE for formative assessment among sex workers in Menaland.

Unit 4

Survey methods

Overview

What this unit is about

IN THIS UNIT, you will learn about selecting and adapting instruments and methods for behavioural surveillance.

Warm-up questions

1. What is measurement error?

2. In HIV behavioural surveys, which of the following is a source of measurement error?
   a. interviewer
   b. participant
   c. questionnaire
   d. all of the above

3. True or false? You can always control for bias during data analysis. Circle your answer below.

   True   False

4. True or false? If we use face-to-face interviews to survey most-at-risk populations, we can collect more accurate data compared with other survey methods, such as self-administered paper-based or computer-assisted interviews.

   True   False

5. Which of the following is an advantage of using the self-administered paper-based questionnaire?
   a. inexpensive to administer
   b. no literacy requirement
   c. data-entry step eliminated

6. Which of the following are advantages of adapting survey questions from surveys that have already been used successfully?

   a. builds on current best practice of how questions can be expressed

   b. saves time and money

   c. makes certain your data are consistent with other available data sources

   d. all of the above

7. What is a solution to the problem of interviewer safety when working with most-at-risk populations?

   a. selecting interviewers who know the area

   b. selecting interviewers who know the population

   c. having interviewers work in pairs

   d. all of the above

8. True or false? When you are having difficulty getting members from most-at-risk populations to come to interviews, one solution is to use incentives.

   True   False

Introduction

What you will learn

By the end of this unit, you should be able to:

- select appropriate data collection methods
- design and adapt survey instruments
- understand measurement error and how to reduce it
- think about potential difficulties in your fieldwork and find solutions
- select appropriate interviewers and supervisors.

Data collection methods

The type of method you use for data collection can influence how participants answer questions. The answers influence how accurately the data measure the behaviour you are studying. This is particularly true of data on topics that are difficult to talk about, such as sexual and drug-use practices. It is important to make sure that people know that their answers will remain private. Data collection methods you can use in behavioural surveillance include the following:

- Face-to-face interviews: An interviewer reads the questions to the participant. The participant gives a spoken response. The interviewer records the response on the questionnaire.
- Self-administered paper-based questionnaire: The questionnaire is given or mailed to the respondent. The respondent reads and completes the questionnaire.
- Computer-assisted questionnaire: The respondent uses headphones to hear the questions. The respondent answers the questionnaire on a computer.
- Audio-assisted questionnaire: The respondent uses headphones to hear the questions. The respondent writes answers on paper.

The best data collection method will depend on the context. Each method has advantages and disadvantages, as shown in Table 4.1.

Survey instruments

You have already learned that a characteristic of a good indicator is that it measures behaviour accurately. You can improve accuracy by phrasing the question properly. Luckily, you do not have to start over every time you develop a questionnaire. You can adapt your questions from surveys that have already been successful. Adapting questions does the following:

- builds on current best practice: the best ways to ask the question
- saves time and money
- makes sure your data are consistent with other available data sources.

Whether you adapt a questionnaire or write your own questions, it is essential to pretest and pilot-test the questions.

Table 4.1 Advantages and disadvantages of various data collection methods

Method	Advantage	Disadvantage
Face-to-face	- no literacy needed - interviewer-respondent rapport can increase truthfulness - high rate of response - you can have different lengths or styles of survey	- labour intensive - low degree of privacy and anonymity - interviewers may distort answers without meaning to
Self-administered (paper-based)	- higher degree of privacy and anonymity - inexpensive to administer	- respondents must be literate - higher rate of non-response - questionnaire must be short
Computer-assisted	- higher degree of privacy and anonymity - data-entry step eliminated	- respondents must be literate and at ease with computers - requires expensive equipment - less useful in venue-based surveys (when you are on site collecting information, rather than in a clinic or office)
Audio-assisted	- higher degree of privacy and anonymity - minimal interviewer training - minimal literacy needed	- requires expensive equipment - less useful in venue-based surveys

Table 4.2 Content of surveys useful for behavioural surveillance

Survey	Content
Demographic and Health Survey (DHS), AIDS Indicator Survey (AIS), ICF MACRO http://www.measuredhs.com/	- knowledge and source of knowledge of HIV/AIDS and other STIs - knowledge of how to avoid HIV/AIDS - condom use at last and next-to-last sex - relationship to last and next-to-last sexual partner - length of time knowing last and next-to-last sexual partner - age at first sex
Multiple Indicator Cluster Survey (MICS), UNICEF http://www.childinfo.org/index2.htm	- knowledge and source of knowledge of HIV/AIDS - knowledge of how to avoid HIV/AIDS - knowledge of testing sites and if ever tested
Behavioural Surveillance Survey (BSS), Family Health International http://www.fhi.org/en/topics/bss.htm	surveys for sex workers, men who have sex with men, injecting drug users, youth and adults on HIV-related risk behaviours

There are several sources for general population survey instruments that are widely used, that use good sampling procedures and thorough statistical analysis, and that are considered a reliable source of behavioural data. See Table 4.2 for some of these sources.

Measurement error

Definition

Measurement error occurs when the data collected do not accurately measure the behaviour. This error affects the validity of your data.

Measurement error terms

Measurement error is best shown with an example. Imagine the surveillance system wants to measure how often people have unsafe sex. To do this, the participant is asked if he or she has had unsafe sex in the last month.

Measurement error could occur in one of two ways:

- The person had unsafe sex but is too embarrassed to say it, so answers 'no'. In this case, we do not measure behaviour accurately because the person does not say the truth.
- The participant thinks of unsafe sex as sex with the risk of pregnancy instead of sex without use of a condom. In this case, we do not measure behaviour accurately because the respondent does not interpret the question as we intended.

Some measurement errors cause more problems than others. Suppose, for example, we are interested in measuring the average number of sexual partners in the last month for men and women, but in our imaginary population, men are more likely to overreport the number of partners, and women to underreport the number of partners. Because of this over- and underreporting, the difference between men's and women's behaviour appears larger than it is. This is systematic error, which can result in biased findings. Bias refers to systematic error in the collection or interpretation of data. Bias cannot be controlled for during analysis or by increasing the sample size.

Measurement errors also occur because of sampling. Sampling refers to choosing a number of people who then represent a larger population. If the answers these people give are different from the answers that the entire population would have given, random error, or non-systematic error, will result. Random error makes the data vary widely. Unlike bias, you can control these variations statistically or by using different study designs.

Systematic and non-systematic error: example

Table 4.3 shows three situations of measurement error for a sample of six people. The samples are described below the table.

Table 4.3 Examples of measurement error

	Condition 1: True answers in entire study population	Condition 2: Non-systematic error (random error)	Condition 3: Systematic error
Person 1	Yes	Yes	No
Person 2	Yes	Yes	No
Person 3	Yes	Yes	No
Person 4	No	No	No
Person 5	Yes	Yes	Yes
Person 6	Yes	Yes	No
Person 7	No		
Person 8	No		
Person 9	Yes		
Person 10	No	Not included in sample	Not included in sample
Person 11	No		
Person 12	No		
% unsafe sex	50%	5/6 (83%)	1/6 (17%)

Condition 1: The entire population was included in the survey, and all 12 persons in the population understood and answered the question correctly. The true prevalence of unsafe sex in the last month in the population would be measured as 50%.

Condition 2: The six people picked by chance to represent the total population of 12 people have a measured prevalence of 83%. This is an example of random error.

Condition 3: Persons 1, 2, 3 and 6 misunderstand the question and report that they did not have unsafe sex when, in fact, they did. The error in reporting is systematic, meaning that all persons who misunderstood the question underreport unsafe sex. The measure of unsafe sex in our sample is only 17%, so the measure is biased. This is an example of systematic error.

Factors in measurement error

In surveys, there are several sources of measurement error:

- questionnaires
- interviewers
- respondents.

Questionnaire errors

How questions are phrased and in what order they are asked can affect how accurately the data reflect the behaviour. Measurement errors may occur if the questionnaire contains the following:

- Culturally inappropriate questions: Some words and phrases may not be understood or may have different meanings in different cultures. Some questions are understood, but are considered too personal or negative in some cultures. To avoid these possibilities:
  - Fit the wording of the question to the type of participant.
  - Explore local terms with formative research and pretest questionnaires.
  - Soften the effect of questions that may seem negative or too personal. Place personal questions near the end of the questionnaire.
  - If translation is required, check accuracy by back translation to the original language.

- Unclear wording: Unclear questions lead to unclear answers.
  - Keep words simple and avoid jargon.
  - Present only one concept at a time.
  - Be specific when asking questions. Do not ask, "Have you had unsafe sex recently?" Instead, be specific and ask, "Have you had vaginal intercourse without a condom in the past 30 days?"

- Leading wording: Some questions are phrased so that the respondent is more likely to choose a certain answer. The question may mention only some of the choices, such as: "When should a condom be used: for oral sex, vaginal sex or what?" They may be leading in other ways, such as indicating a socially-desirable response, threatening self-esteem or being emotionally charged.
  - Either mention all possible answers to the question or mention none.
  - Do not use emotionally charged words.
  - Avoid threats to the respondent's self-esteem.

- Too many or too difficult questions: Asking too many questions or questions that are too difficult can make people impatient or bored with the interview.
  - Make questionnaires shorter and questions easier to understand.
  - Exclude questions you will not use.
  - Use screening questions with skip patterns.
  - Allow "not applicable" responses.

- Poor ordering of questions: Putting questions that people may be sensitive about at the start of a questionnaire can make them uncomfortable and decrease any rapport with the interviewer. Suddenly changing topics also can be confusing and cause respondents to lose concentration.
  - Place questions about the same topic together.
  - Move from general to specific questions.
  - Place sensitive questions towards the end of the questionnaire.

Interviewer errors

Even with a well-designed questionnaire, the behaviour of the interviewer can affect the accuracy of the data. Interviewer errors include:

- Not following the fieldwork protocols: In a survey, interviewers should ask the questions exactly as they are written on the questionnaire, and not interpret questions in their own way. If interviewers do not follow protocol, participants may not be asked the same questions and measurement error may occur.

- Produce an interviewers' manual.
- Select and train interviewers well.
- Provide good supervision of interviewers.
- Check quality of data and procedures.

• Being judgmental: If respondents think that the interviewer will disapprove, they may be less willing to report behaviours that could seem socially undesirable.
- Do not show disapproval to respondents.

• Misinterpreting responses: An interviewer can misinterpret the respondent's answer. For example, the respondent may use slang that the interviewer does not understand.
- Use interviewers who understand the language of the survey population.
- Train interviewers to take notes whenever something is unclear.

• Making mistakes in recording answers: Sometimes interviewers may makes mistakes such as recording answers incorrectly or skipping questions.
- Train interviewers well.
- Produce an interviewer's manual.

Respondent error

The actions, behaviours and perceptions of the respondent can affect how accurately the data collected reflect behaviour. A respondent may:

• Misunderstand the question: The questions may be asked using language that is too sophisticated or that does not consider the educational level or cultural norms of participants.
- Fit the wording of the question to the participant.

• Show faulty recall: A respondent's memory of past behaviours can be influenced by behaviours or events that happened more recently. This is often referred to as recall bias.
- Make sure the time reference periods are right to help participants remember.

• Misreport answers: A respondent may hesitate to give accurate answers if there seems to be a threat to privacy. For instance, because of its private nature, sexual behaviour is difficult to study. Also, a person may overreport socially desirable behaviours and underreport socially undesirable behaviours. Mismatching the personality, sex, ethnic identity or age of the interviewer to the participant also can influence the answer.
- Use the right data collection method.
- Place private questions towards the end of the questionnaire.
- Use members of the target community as interviewers.
- Use interviewers of the right age, sex and ethnicity.

Some of these strategies can be applied before the beginning the survey by pretesting the questionnaire and pilot testing the survey procedure.

Pretesting

Often, HIV behavioural surveys are developed by adapting existing questionnaires to match the sociocultural context of the population you will study. Pretesting the individual questions and the questionnaire as a whole are important steps in developing good survey instruments. Pretesting usually means collecting comments on the questionnaire from key informants (members of the target population). Use these comments to determine:

- the best order for sections and questions
- whether the wording of questions is clear and understandable to people being interviewed.

Revise unclear wording before the questionnaire is completed.

Pilot testing

After pretesting, pilot testing improves questionnaires even more. Collect data from small samples of respondents to check how well the fieldwork procedures will work in practice. Piloting will help you to:

- understand how long it takes to administer the questionnaires
- improve the training of interviewers by finding sections or questions that cause problems, ideas and issues that are not clear, and skills that need to be strengthened
- plan fieldwork by finding out how well interviewers cope with conditions, if there are any problems with sites or equipment, and if there may be problems recording or checking data.

Fieldwork practicalities

Most-at-risk populations included in surveillance may be hard to reach because:

- They engage in illegal and secret behaviours.
- They often do not want to be identified because of high levels of stigmatization and discrimination.
- Their existence is denied by the general population and government.
- There are restrictions on who may approach the group and how the group may be approached. Gatekeepers, such as pimps or bar owners, may not want sex workers interviewed, the government may not want non-military personnel interviewing military personnel, and so forth.
- Groups do not want to be found for surveillance because they fear authorities will persecute them.
- Groups do not want outsiders entering the group.
- Group members are difficult to tell apart from non-group members. For example, on entering a bar known to be a place where sex is sold, how do you differentiate a sex worker from a customer of the bar?

Examples of practical difficulties and potential solutions for working with most-at-risk populations are shown in Table 4.4.

Table 4.4 Practical problems and solutions for working with hard-to-reach groups

Problem	Solution
How to identify group members	• Collect information on how to identify group members and on the places where they may be identifiable and accessible. • Use members of the group, such as ex-injecting drug users, ex-sex workers, other men who have sex with men, as interviewers.
Fear of persecution decreases participation	• Do activities that let you enter the community. • Work with authorities to make sure that surveillance activities do not lead to police response. • Make sure there are some real benefits for group members that result from the surveillance activities. • Use group members as interviewers or to inform interviewers about issues specific to the group. • Work with nongovernmental organizations that are accepted by the target population. • Realize that it can take more than one surveillance round to gain a group's trust. Try again if there are problems in initial rounds.
Restriction of surveillance by gatekeepers	• Explain to gatekeepers the importance of surveillance and how the results will ultimately benefit the population. • Secure gatekeepers' support for the project. • Involve gatekeepers in the surveillance activities. • Work with nongovernmental organizations that are accepted by the target population.
Scheduling	• Work around schedules of groups and group members. • Make interviews as short as possible. • Use incentives for participation.
Consent and confidentiality	• Consider doing anonymous surveys, if this is appropriate (depending on whether you intend follow-up of treatment, for example). • Train interviewers on the importance of confidentiality and consent. • Make sure that the gatekeepers know that participation is voluntary. • Identify potential problems with privacy and confidentiality.
Interviewer safety	• Select interviewers who know the area. • Have interviewers work in pairs.

Selecting appropriate interviewers and supervisors

Interviewer and supervisor characteristics and skills

To fulfil the roles described below, supervisors and interviewers need to have specific characteristics and skills. Some of these things can be taught through training, such as how to ask questions and record information. Certain skills cannot be taught, such as fluency in different languages, previous experience with research and with the target populations, and education. Characteristics that cannot be changed, like gender, ethnicity and age, help in interactions with respondents and community members who have the same characteristics. Your interviewers and supervisors should have these qualities.

Interviewer's role

The interviewer is responsible for using questionnaires to collect and record information from respondents.

Supervisor's role

The supervisor's role includes:

- representing the project to the community
- negotiating with and contacting communities
- managing fieldwork and fieldwork teams
- organizing practical matters, such as travel, places to stay and payment of interviewers
- supervising interviewers and reinterviewing respondents to check the quality of interviewers' work
- checking questionnaires
- controlling recruitment of participants
- making certain that ethical guidelines are followed
- identifying problems and finding solutions.

Summary

Measurement error occurs when the data collected do not accurately measure the behaviours being studied, which affects the validity of the data. It is important to select the right data collection methods because the method can influence how accurately respondents answer questions. You should also design and adapt survey instruments that are widely used, are considered to use good sampling procedures and thorough statistical analysis, and are considered to be a reliable source of behavioural data.

Unit 4 exercises

Warm-up review

Take a few minutes now to look back at your answers for the warm-up questions at the beginning of the unit. Make any changes you want to make.

Small group discussion

Get into small groups by country, region or district to discuss these questions.

1. Based on your experiences, discuss what you would do to reduce measurement errors in HIV behavioural surveillance surveys.

2. What are some general population surveys that you have used that ask questions about HIV/AIDS?

Apply what you have learned/case study

Try this case study individually. We will discuss the answers in class.

1. In a study of 1417 participants of syringe exchange programmes, two methods of interview were used: face-to-face and audio-computer assisted self-interviews (ACASI). The survey contained questions about several "stigmatized behaviours". After completion of the survey, researchers compared the frequency of responses to questions related to the stigmatized behaviours based on the interview methods. They found that participants who were in the ACASI group had reported more frequent stigmatized behaviours than participants in the face-to-face group [3].

   a. Discuss potential reason(s) that may explain this difference.

   b. Imagine that you want to do a survey among injecting drug users who go to drug treatment centres in your country. What would be the advantages or disadvantages of using each of the methods?

2. What topics should be taught during field supervisor training? Outline what you would include in a field supervisor training course.

Unit 5

Sampling

Overview

What this unit is about

IN THIS UNIT, you will learn key terms and approaches to sampling. This will allow you to discuss and select sampling options for behavioural surveillance activities.

Warm-up questions

1. The _____ is the group that meets a survey's measurement objective (for example, all sex workers in a city).
   a. target population
   b. survey population

2. Is choosing names randomly for sampling an example of probability sampling or non-probability sampling?
   a. probability sampling
   b. non-probability sampling

3. The accuracy of a sample survey depends on which factor?
   a. the size of the area included in the survey
   b. the size of the sample
   c. the number of interviewers
   d. the number of months set aside for the survey

4. True or false? Non-probability sampling is prone to selection bias. Circle your answer below.

   True   False

5. You can increase the precision (that is, decrease standard error) of your estimate by increasing the _____.
   a. sample size
   b. quality of interview training

6. Point prevalence is the total number of times something occurs in the population at one point in time divided by the total size of the population.

   True   False

7. What is the estimate of precision that is used to construct a range of values in which the true measure of the population is likely to fall?
   a. standard error
   b. systematic sample

8. True or false? You should use non-probability sampling techniques whenever a sampling frame is not available or it is not possible to develop one.

   True   False

Introduction

What you will learn

By the end of this unit, you should be able to:

- explain how sampling a population is different from a complete count
- outline the purpose of sampling
- explain the most important concepts in sampling
- explain the need for sampling designs that are more complex
- describe sampling issues and options for behavioural surveillance
- describe how to choose a sampling approach.

Concepts, terms and definitions

Target population

Often, you want to survey a group of people to measure certain characteristics. The target population is the group that meets the measurement goals of a survey, such as all people in the entire country, the adult population in a country or sex workers in a capital city.

Complete count

There are several ways to gather information about the target population. One way is to do a complete count, in which you contact every person in the target population to ask your questions.

A national census is a good example of this type of information gathering. To do a national census, you need a very large team of interviewers, a great deal of funding and several months (or longer) to interview everyone.

Sometimes even interviewing a subpopulation in a large city is costly, time consuming and difficult. For example, if you want to measure condom use among sex workers in a capital city, it is not possible to interview every sex worker.

Sampling

When you want to measure characteristics of a target population, it is not necessary or practical to survey every person. Instead, you can select a number of people from the population. How you select these people is called sampling, and the people you select are your sample.

The goal of sampling is to estimate certain indicators from your sample that you can generalize to the target population. The ways of selecting these people are called sampling approaches.

How well the sample reflects the entire population is called representativeness. How well the results from a sample would reflect the true results from the entire population is called generalizability. The more representative the sample is of the entire population, the more generalizable the results will be.

For example, when you want to measure condom use among sex workers in the capital city, you can get an estimated proportion of sex workers who use condoms without knowing the exact number in your target population. Instead, you can survey a sample for condom use. If the survey is done properly, the estimated proportion is representative of the entire target population.

Survey population

To do HIV behavioural surveillance among most-at-risk populations, such as injecting drug users, sex workers, and men who have sex with men, you may focus only on subpopulations in a certain geographical area instead of focusing on the entire target population. The survey population is the target population modified for practical reasons. This means you can generalize your estimates only to the defined survey population. For example, if sex workers participating in a survey are 18 years of age and older and work only in bars, you cannot generalize estimates from this sample to sex workers who are younger than 18 years of age and work on the street. The survey population is sometimes called the accessible population.

Selecting a sample

Introduction

When we select a sample, it is because we want to obtain estimates of certain indicators from a target population by collecting data only from selected individuals. We also want to make sure that what the sample tells us is accurate. Accuracy refers to how well the estimates from the sample reflect the true estimates of the target population. To improve accuracy, you should make sure that your sample gives unbiased and precise estimates of the indicators for your target population.

Bias, also called systematic error, is a type of error that causes the estimates based on your study to depart systematically from the true value of your indicators. Because there are many possible samples that could be selected from the population, the people in a sample can be slightly different from the people in another sample just by chance. This difference produces a different estimate of your indicators. This possibility must be considered when you use a sample to draw conclusions about a larger population. Precision refers to how well the results can be reproduced if you repeat the same study in the same population using a different sample.

You can get a more precise estimate by repeating the same study many times and using the average as your estimate. This average will come close to the true estimate that you would get from the entire population and will become unbiased. However, we cannot repeat a study many times. A more practical method is to use a larger sample size, as we will discuss in the next few pages.

The issue relating to bias and precision is demonstrated in Figure 5.1.

(1)	(2)	(3)	(4)
Biased	Biased	Unbiased	Unbiased
Imprecise	Precise	Imprecise	Precise
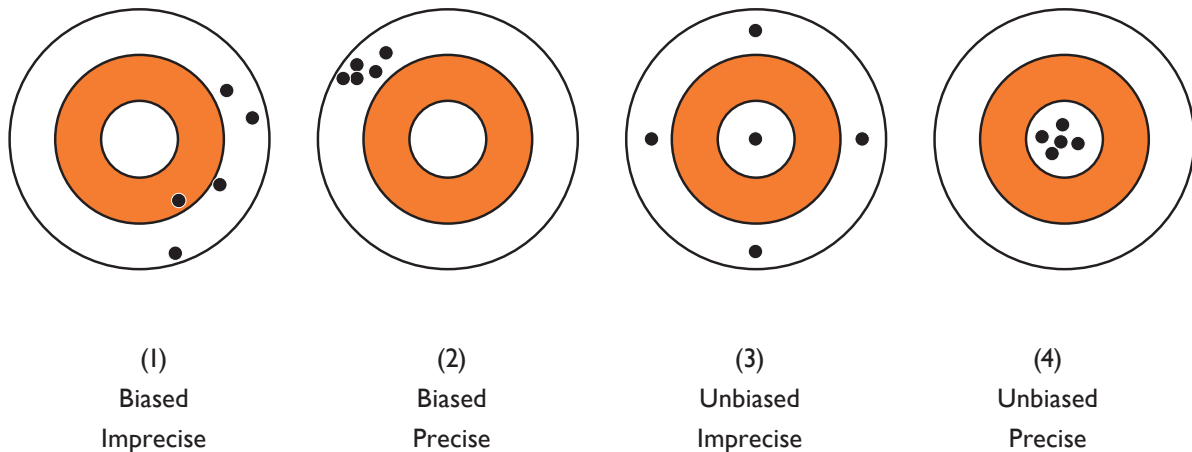
Adapted from [4].

Figure 5.1 Illustration of the relation of bias and precision in estimating health indicators of the target population

Group discussion

Looking at Figure 5.1, discuss which of the images represents the most accurate estimates.

Selection bias

Bias (systematic error) can occur during the design of the survey or during collection, analysis or interpretation of data. If you avoid selection bias (bias due to sampling) you know that your sample is representative of the whole population. To prevent selection bias from entering the results during data collection, the selection of people for the sample must be random. This means that every person in the target population should have the same opportunity to be selected.

For example, you are doing a household survey to study characteristics of marriage and sexual activity of women living in a city. While sampling, you may find that married women are more likely to be at home than unmarried women. There may be a difference that is not clear between women who are not at home during sampling and women who are at home. If you end sampling here, you may bias your estimates toward married women who are at home at the time of your survey. If you make repeated visits to the area to sample women who were not at home or not in the area, you will avoid selection bias.

Probability sampling

Avoiding selection bias requires a random or probability sample. In a probability sample, each person in the population has a known, non-zero probability of being selected.

As an example, imagine that prisoners in a jail are your target population. You want to know what percentage of the prisoners practise unsafe sex, so you conduct a probability sample by writing each prisoner's name on a piece of paper, putting all the papers into a box, and sampling names by taking papers randomly from the box, like a lottery. Thus, everyone has a chance of being chosen or a non-zero probability of selection.

To say it another way:

- If there are 1000 people in the prison, the probability of any one person being sampled would be 0.001 (1 out of 1000).
- If there are 200 people in the prison, the probability would be 0.005 (1 out of 200).

This is an example of simple random sampling (SRS). Everyone has a known and equal chance of being selected. SRS is the simplest sampling method and the most basic method of probability sampling.

Further sampling basics are discussed in the next few pages, but we will see later in this unit that we have only a few options when we sample most-at-risk populations for behavioural surveillance.

Sampling frame

Developing a sampling frame is a fundamental part of probability sampling. A sampling frame is a list of people from which a sample may be selected. Figure 5.2 shows the relation among the target population, survey population, sampling frame and sample.

Using the example above, the list of prisoners' names to draw from the box is the sampling frame. You could also give each prisoner a number, then select numbers from a random numbers table. Statistical software can also be used to generate random numbers. Go to Appendix A.5.3 to find out how a random number table works.
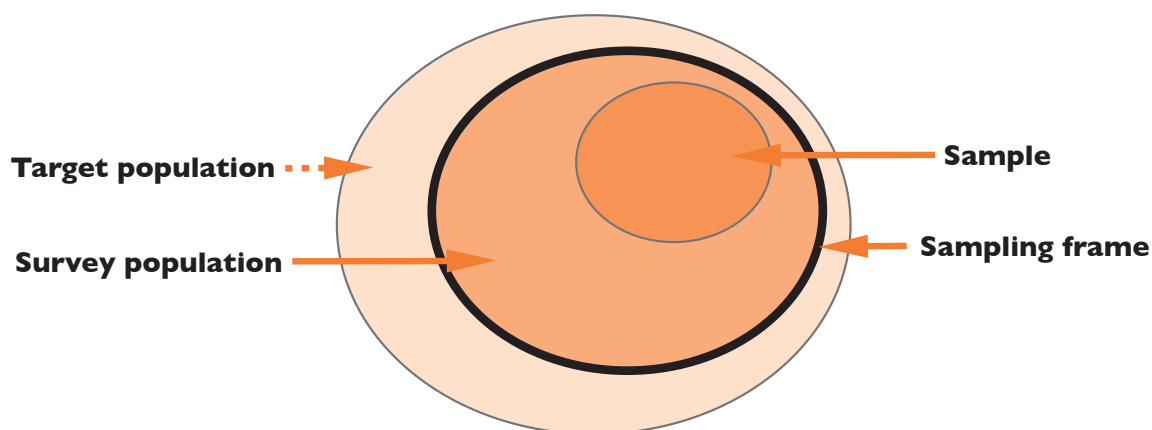
Figure 5.2 The relation among target population, survey population, sampling frame and sample

Calculating sample size

Effect of sample size on precision estimates

When you use probability sampling, you can figure out how large your sampling variation could be. That is, how much the estimate from your sample may vary from the true value in the population. This estimate of precision is called standard error.

You can use standard error to construct a range of values. You usually want to be 95% sure that the true population measure lies in this range. This range is called a margin of error or confidence interval.

When you increase the precision, you decrease the standard error. You can increase the precision of your estimate by increasing the sample size. Standard equations are available to calculate sample size. The example below shows the effect of sample size on an estimate.

Imagine you estimate that the true percentage of women who have heard of AIDS is 15% in a population of 100 000. You want to know whether this estimate is correct for the entire population:

- A sample size of 50 women would provide us with 95% confidence that the true percentage of the population is within ±10% of your estimate. This means that if you estimate 15% for the sample, the true percentage in the population could be as low as 5% (15%–10%) or as high as 25% (15% + 10%). Thus, 95% confidence interval (CI): 5%–25%. This is a wide confidence interval, which means a large margin of error and a less accurate estimate.
- A larger sample size, of 1250, would provide us with 95% confidence that the true percentage in the population is within ± 2% of your estimate. If you estimated from your sample that the percentage of women who have heard about AIDS is 15%, the true percentage of the population could be as low as 13% (15%–2%) or as high as 17% (15% + 2%). Thus, 95% CI: 13%–17%. This is a narrow confidence interval, which means a small margin of error and a more accurate estimate of the total population.

This example shows that the larger your sample size, the smaller your margin of error and the more accurate your estimates.

Factors in sample size

Calculating sample size is important for any behavioural survey. There are several things you should consider, including:

- what types of things (variables) you want to measure, such as age or whether married
- whether there are data on the variable you chose
- whether you want to measure your variable at one point in time or to measure the change in your variable over time.

Types of variables

There are two types of variables:

- Categorical variables are used to list characteristics. The variable may provide two options, such as urban or rural, or more options, such as married, divorced, or single.
- Continuous variables are used to register numerical values, such as age or CD4 cell count

This unit focuses on categorical variables, which are used most often in point prevalence estimates (described below). We do not discuss continuous variables because it would entail explaining more about the formulas used to calculate them than is possible in this module.

Using existing data

Point prevalence is the total instances of a condition in the population at one point in time divided by the total population size. For example, HIV point prevalence is defined as the proportion of the population infected with the virus.

Imagine that you are interested in estimating the point prevalence of condom use among sex workers. You first need to have a rough idea of this percentage:

- If your country has not yet started to use HIV behavioural surveillance, you would:
  - Conduct a pilot study or use existing data on related indicators, such as family planning, STIs, or condom use. You also would check for data in similar regions or countries to get an estimate of the percentage of condom use.
  - Use this estimate to calculate the size of the sample needed to get an accurate estimate of this percentage, given that you accept a certain margin of error, such as plus or minus 1%.

If HIV behavioural surveillance is used in your country, you can use available data to find the percentage of condom use among sex workers. Then you can calculate sample size for your new survey round or change in the percentage over time.

Sample size for point prevalence

The sample size needed to conduct behavioural surveys can be based on the number of participants needed in each round (or year) to detect a point prevalence of an indicator or a change in the proportion of an indicator from one round to the next.

Formula to detect a point prevalence

The general formula for the sample size (n) needed to detect a point prevalence of a certain indicator is [5]:

$$n \quad = \quad \frac{[Z_{1-\alpha}]^2 \times P \times (1-P)}{E^2}$$

Where:

n = Sample size required per survey round (year)

$Z_{1-\alpha}$ = The z score for the desired confidence level, usually 1.96 for 95%

P = The proportion of the estimated indicator

E = Margin of error

The $Z_{1-\alpha}$ score is a statistic that corresponds to the level of significance desired. Usually a significance level of 0.05 (or, equivalently, a 95% confidence level) is selected and it corresponds to a value of 1.96. The smaller the significance level (that is, the higher the confidence level), the larger the sample size you will need.

P is the estimated proportion of an indicator. For example, you wish to show that condom use at last paid sex for sex workers was 20% in 2007, so *P* is equal to 0.2. The closer *P* is to 50%, the larger the sample size you will need.

E is the margin of error, which is defined as the radius of a confidence interval for the estimated proportion. For example, if the margin of error is 5%, it means that you wish to show that the estimated proportion of condom use (in this case 20%) falls within ± 5% of 20% (15%–25%). The smaller the margin of error, the larger the sample size you will need.

Additional issues in calculating sample size

Design effect

The design effect (*D*) accounts for the similarities people have when they are sampled within the same cluster. For example, injecting drug users within a particular cluster may be similar in condom use because of the social norms, condom availability or intervention programmes of their specific setting. The design effect can be thought of as a correction factor for how much a cluster sample differs from a simple random sample. Effectively, the design effect multiplies the sample size by a factor of *D*. Choosing a design effect is difficult without prior survey data. Design effects from 1 (that is, none) to 2 (moderate) cover a typical range.

For respondent driven sampling (RDS) surveys, the minimum design effect of 2 is recommended. For cluster sampling and time-location sampling (TLS), a moderate design effect of 2 is recommended. These sampling techniques will be presented in the next few pages. The bigger the *D*, the larger the sample size needed [5].

$$n \;=\; D \times \frac{[Z_{1-\alpha}]^2 \times P \times (1-P)}{E^2}$$

The size of the target population has very little effect on the size of the sample needed.

Table 5.1 shows the minimum sample size needed for detecting different point prevalence estimates (ranged from 5% to 50%) with different margins of error (ranged from 1% to 3%) for hypothetical risk behaviours within a population.

Note that if all other factors remain the same, the size of the required sample increases as the estimated point prevalence increases up to 50%. Because of the characteristics of the formula for sample size, the required sample size reduces for prevalence estimates above 50%. For example, the required sample size to estimate a point prevalence of 60% is the same as for a point prevalence of 40%; 70% is the same as for 30%; 80% is the same as for 20%; and so on. Also, if all other factors remain constant, the size of the required sample decreases as the margin of error increases.

Table 5.1 Minimum sample size needed to detect different point prevalences and margins of error within a 95% confidence interval

Point prevalence of an indicator for HIV risk behaviour	Margin of error	Sample size
0.05	0.01	931
0.1	0.01	1764
0.15	0.01	2499
0.2	0.01	3136
0.25	0.01	3675
0.3	0.01	4116
0.35	0.01	4459
0.4	0.01	4704
0.45	0.01	4851
0.5	0.01	4900
0.05	0.02	233
0.1	0.02	441
0.15	0.02	625
0.2	0.02	784
0.25	0.02	919
0.3	0.02	1029
0.35	0.02	1115
0.4	0.02	1176
0.45	0.02	1213
0.5	0.02	1225
0.05	0.03	103
0.1	0.03	196
0.15	0.03	278
0.2	0.03	348
0.25	0.03	408
0.3	0.03	457
0.35	0.03	495
0.4	0.03	523
0.45	0.03	539
0.5	0.03	544

Calculating sample size for change over time

The sample size needed for behavioural surveys can be based on the number of participants needed in each round (or year) to detect a change in the proportion of an indicator. This is shown in Table 5.2. More detailed information about how to determine minimum sample size for this type of study is provided in Appendix A.5.2.

Table 5.2 Sample size needed to detect a change in the proportion of an indicator

Indicator level at time 1 (P_1)	Indicator level at time 2 (P_2)	Sample size needed each year
0.10	0.20	199
0.10	0.25	99
0.20	0.30	293
0.20	0.35	138
0.30	0.40	356
0.30	0.45	162
0.40	0.50	387
0.40	0.55	173
0.50	0.60	387
0.50	0.65	169
0.60	0.70	356
0.60	0.75	152
0.70	0.80	293
0.70	0.85	120
0.80	0.90	199
0.80	0.95	75

In Table 5.2, above, P_1 and P_2 are the measures for which you wish to see a change between survey rounds. For example, if you wish to show that sex worker condom use at last paid sex increased from 20% in 2006 (P_1) to 30% or greater in 2007 (P_2), you would need to survey at least 293 persons each year. Here are some additional points:

- P_1 is usually based on previous surveys in the same or similar populations or on an educated guess at what the level will be.
- P_2 is ideally set at the goal you would like to achieve, such as a 10% or greater increase in condom use.
- In practice, it is usually set at the smallest change you think is meaningful. For example, a 10% increase in condom use would be considered a meaningful improvement. A 1% increase would not be considered meaningful.

HIV behavioural surveillance often measures trends of indicators over time. It is essential that the different survey rounds define and sample a subpopulation consistently over time. If you are not consistent, you do not know if any of the changes are real or just due to differences in the methods you used.

Sometimes it is appropriate to change the sampling strategy because better techniques are developed. In this case, you can sample the population using both the old and new method to compare estimates.

Probability sampling techniques

Introduction

As we saw earlier in this unit, SRS is the most basic form of probability sampling. It uses a lottery procedure to sample and everyone in the population has a known and equal chance of being selected.

SRS has limitations, but there are other techniques discussed and defined below that can help, including:

- systematic sampling
- cluster sampling
- stratified sampling.

Systematic sampling

Scenario: You know the total number of people in the target population, but there is no pre-existing list of people. To do SRS, you must count a large number of people first.

Solution: A better method than counting everyone in the population is to sample from it systematically. In systematic sampling, you construct the sampling frame as you sample. You divide the total number of people in the target population by the necessary sample size to obtain blocks of individuals. For example, if you want to sample 100 from a total of 1000 people, then your block size is 10 (1000/100 = 10). You select every tenth person.

Implement: First, make sure that the list is not ordered in any way that would bias who is selected. Then do the following:

- Select a random starting point from this block (for example, from 1 to 10, you randomly select 3).
- Sample people at regular intervals (10) down the list, starting at 3 (3, 13, 23, 33, and so on).
- Continue this process to reach your sample size of 100.

Cluster sampling

Scenario: A sampling frame for the whole population does not exist. The target population is spread over a wide area. Travel and time costs to cover the whole area are high.

Solution: When it is difficult or impossible to make a list/sampling frame of each individual in the target population, you can develop a sampling frame of some larger unit. These are called clusters or primary sampling units.

Implement: You sample in stages, by first sampling clusters and then sampling people within the clusters. A cluster is any aggregate of the population, such as departments, villages or health facilities.

Cluster sampling is the most common method of sampling in surveys. It has the advantage that a sampling frame (a list of every person in the target population) is not required. Instead, a sampling frame of clusters is used.

Once the clusters are selected, you need only to list people in the selected clusters. All members of the target population still have a chance of being sampled (a non-zero probability) as long as all the clusters that make up the target population are included in the list of clusters.

When the population is spread out, you can use cluster sampling, as described above, to concentrate fieldwork in specific areas to reduce time and travel costs.

If clusters are of unequal size, you need to take this into account. If you do not, people in smaller clusters will have a higher probability of being selected than people in larger clusters and your sample will be biased. There are statistical methods that you can use so that the sample gives unbiased estimates, but they are beyond the scope of this module.

Cluster sampling results in a less precise estimate of your indicators than SRS, because people within clusters may be more similar to each other than to people in other clusters. You can correct this by increasing the sample size.

Stratified sampling

Scenario: The target population consists of distinct subgroups. You may need to make precise estimates of your indicators for each subgroup.

Solution: When the population consists of distinct subgroups (for example, age groups or regions), use stratified sampling.

Implement: First, calculate the sample size needed to measure your indicator. Then define the subgroup (stratum). Finally, randomly select the necessary sample size in each stratum.

Because you want to make precise estimates of your indicator for each stratum, your sample size will be much larger than if you wanted an estimate for the entire population. You can combine strata estimates to obtain a population estimate for your indicators. However, you must know the proportion of the population in each stratum.

Stratification is the classification of a population into subgroups or strata by certain characteristics. Stratified sampling is the selection of separate (that is, independent) samples from each stratum. Stratification can result in a more precise estimate and is sometimes used purely to improve precision.

Sampling most-at-risk populations

Most-at-risk populations in household surveys

Household surveys, also known as general population surveys, are rarely an appropriate method for locating members of most-at-risk populations:

- The group members may not be found in sufficient numbers through a household survey.
- Group members may have behaviours that are too sensitive to discuss in a household setting.

For example, to study HIV risk behaviour among men who have sex with men, you may have to visit many households to find the required number of men who have sex with

men. When you find them, they may be unwilling to discuss their sexual behaviours in the presence of other household members or at home.

Most-at-risk populations are hard to identify

Injecting drug users, sex workers and men who have sex with men, are usually hard to reach for surveillance. Members of the target population may be hidden due to their stigmatized or illegal behaviours and may be unwilling to be identified or to acknowledge their risk behaviour. In this case, constructing a sample frame of the target population can be difficult or impossible.

Most-at-risk populations are hard to reach

Sometimes you can identify the places that most-at-risk populations congregate, define these as clusters and sample them. Examples of possible clusters for most-at-risk populations are shown in Table 5.3 below. Opposition from gatekeepers, who control access to the location, may make including clusters in the sample difficult. Also, some most-at-risk populations may be concentrated in high crime, crowded or slum areas, making physical access difficult.

Table 5.3 Examples of possible clusters

MARP/vulnerable population	Possible cluster
Sex workers	Streets, bars, hotels, guesthouses, cafes
Men who have sex with men	Parks, bars, cafes
Injecting drug users	Shooting galleries, injecting sites, needle-exchange programmes, methadone clinics
Truckers	Loading, unloading and halting points
Migrants	Refugee camps, households, workplaces

Most-at-risk populations form unstable or no clusters

Some most-at-risk populations do not congregate, making cluster sampling impossible. For example, it would be difficult find a cluster for home-based sex workers unless they all live in the same area.

For other most-at-risk populations, only subgroups of the population congregate. You can use parks as a cluster for men who have sex with men, but not all men who have sex with men visit these areas, so an important section of the men who have sex with men population could be missed.

Finally, constructing sampling frames within the selected clusters may be difficult. People may not want to be identified as a member of the population.

Because developing a sampling frame is required for probability sampling methods, these methods are usually not appropriate for most-at-risk populations.

As a general rule, if you are able to develop a sampling frame for your target populations (e.g. prisoners), then choose probability sampling methods for your survey.

Sampling options in the absence of a sampling frame

Non-probability sampling

There has been much debate about the best way to sample when surveying most-at-risk populations. When making a sampling frame is not possible, rely on non-probability sampling methods. In non-probability sampling, the study subjects are selected by preference or convenience rather than by non-zero probability.

Usually, most-at-risk populations are recruited through institutions, such as hospitals, or through visible venues, such as bars, using convenience methods.

These methods have limitations that may cause biased estimates because their procedures are not random and because of selection bias.

We do not discuss all non-probability sampling techniques in this unit in detail. Many are not recommended for HIV behavioural surveys. To obtain more information about these sampling techniques, refer to "Review of sampling hard-to-reach and hidden populations for HIV surveillance" [6].

Convenience and snowball sampling

The two types of non-probability sampling that have been used in HIV surveillance are convenience sampling and snowball sampling.

- In convenience sampling, you use a group of high-risk persons that are easily approached, such as patients at an STI clinic. All members of the target population who present at the clinic are enrolled until the sample size is met.
- Snowball sampling, also called chain-referral sampling, uses persons who have been enrolled in the survey to refer more people in their group to the survey. These people then refer more people whom they know from their group, and so on.

Convenience sample example

An example of convenience sampling would be if you interviewed sex workers visiting an STI clinic during a condom promotion campaign over one month to measure condom use among sex workers in a city. In this example, you did not select sex workers from the entire city at random, but instead selected people who attended that clinic during a certain period of time.

Convenience sampling is easier than SRS. You do not need to develop a sampling frame, which in this case would be the list of all sex workers in the city. Instead, you conveniently sampled only from sex workers who visited the designated clinic. The problem is that these sex workers might not be the same as the rest of the sex workers in the city.

In convenience sampling, all sex workers do not have an equal probability of being selected for the sample. Sex workers who did not visit the clinic had a zero probability (no chance) of being sampled. If you try to generalize what you estimate about unsafe sex or access to condoms from these sex workers to the entire sex worker population, you could get a biased estimate.

See Table 5.4 for a comparison of probability and non-probability sampling.

Table 5.4 A comparison of probability and non-probability sampling

Issue	Probability sample (such as SRS)	Non-probability sample (such as convenience sampling)
Prone to selection bias	No	Yes
Can generalize results to survey population	Yes	No
Can estimate precision of survey estimates (use statistical techniques accurately)	Yes	No
Usually requires sampling frame	Yes	No

Table 5.4 discussion

If your sample is not a probability sample, or a probability sample was planned but not done correctly, you cannot generalize your findings beyond the sample. This is because you do not know whether the subjects in the sample are different from those in the rest of the target population. Statistical tests are not accurate when performed on non-probability samples.

New sampling techniques

Introduction

Recently, new sampling strategies have been developed to address the limitations of non-probability sampling. The two most widely used sampling techniques, time-location sampling (TLS) and respondent-driven sampling (RDS), provide more accurate data if they are done correctly.

Time-location Sampling (TLS)

TLS is like cluster sampling, but it avoids the problem of clusters that are not stable (that is, clusters where the number and type of people varies by, for example, time of day). TLS allows the same site to be included in the sample more than once—at different times of the day, for example, or on different days of the week.

TLS example

If brothels are open seven evenings per week, from 20:00 to 04:00 and if there are 20 brothels in a city, there would be 7 × 20, or 140 TLS units.

If the sex workers who worked from 22:00 to midnight are different from the sex workers who worked from midnight to 04:00, then you could also divide the timeframe into two halves, from 22:00 to midnight and from midnight to 04:00. You would then have 7 × 20 × 2, or 280 TLS units. These units are essentially clusters.

Once the time-location clusters are defined, a certain number are selected randomly. The number needed will depend on the number of persons in each TLS cluster and how large the sample size needs to be.

For instance, you could choose 30 time-location clusters, and randomly select the same number of respondents in each cluster. Then you can adjust the data during analysis for some TLS clusters having more people in them.

This sampling method avoids the fact that the risk behaviour in a cluster may vary, in this case by time of day. It also means that it is not necessary to count the total number of people in a cluster, only the number of people there during the sampling time. You also need to have ways to keep people from being sampled more than once.

Clusters should not consist solely of places that group members congregate for HIV prevention activities. These locations are likely to be associated with people already concerned about HIV/AIDS.

Respondent-driven sampling (RDS)

RDS is a chain-referral sampling technique, an adaptation of snowball sampling, and does not require a sampling frame.

Members of the target population refer each other for survey because they are assumed to be best at identifying other members and encouraging their participation.

Unlike snowball sampling, RDS restricts the number of people that one person can refer, and it requires statistical adjustment. You can find a full description of this method in Module 4 of this training course and at: http://www.respondentdrivensampling.org/

RDS example

Imagine you want to study injecting drug users in a city. You start your RDS sampling with 8–10 initial injecting drug users (seeds) from the population. These injecting drug user seeds are purposefully selected to be socially active members of their community, with a large network of peers. The process is as follows.

- Each of these seeds is interviewed and trained to become a recruiter.
- Seeds are each given a fixed number of coupons (often three) to use to recruit other eligible injecting drug users from among their network of peers.
- Each new recruited injecting drug user brings his or her coupon to the survey site and is interviewed.
- The recruit and recruiters are linked using unique codes on coupons.

- To encourage participation, the recruiters are given incentives for successful recruitment of other participants.
- Each recruit becomes a recruiter. In this way, the chain of referral progresses for several waves until the sample becomes stable.

Given sufficiently long referral chains (five to six waves), regardless of the type of injecting drug users with which you started, the final sample will reflect the population from which it is recruited.

For example, even if you selected five male and five female injecting drug users as your initial seeds (50% of each group), as the chain of referrals progresses, the proportion of male injecting drug users gradually increases. After the six wave, the proportion of male injecting drug users in your sample reaches 80% and does not change in the seventh, eighth and ninth waves. You then know that this percentage reflects the true population. This is called reaching equilibrium.

RDS should result in unbiased estimates if:

- links between recruiters and recruits are documented
- the sample has reached equilibrium
- you ask about the size of each person's network so you can statistically adjust for the fact that the recruiters are likely to recruit people like themselves and also so you can find differences in personal network size.

Computer packages can assist in analysis of RDS data. Respondent Driven Sampling Analysis Tool (RDSAT) is free software that can be accessed at: http://www.respondentdrivensampling.org

Table 5.5 contains details about TLS on the left and RDS on the right for comparison.

Determining sampling approach

Guidelines to selecting the sampling approach

It is always difficult to find the right sampling technique to recruit members of the target populations. Answering the following questions can help.

Please note that these are only suggested guidelines. Take into account the ethical, practical and sociocultural context in your country.

1. Is there any existing sampling frame for the target population? If yes, a conventional probability sampling technique is recommended. Example: household surveys or census.
2. If no existing sampling frame is available, do members of the target population gather in accessible locations in high proportions? If yes, developing a sampling frame and using a conventional probability method is recommended. Example: prisoners or youth in school.
3. If the target population is one of the most-at-risk populations that gather in accessible sites, but the population is floating, is it possible to construct a list of all sites associated with the target population? If yes, TLS is recommended when group members can

Table 5.5 The advantages and disadvantages of TLS and RDS

Time-location sampling (TLS)	Respondent-driven sampling (RDS)
Steps:	**Steps:**
• Write protocol and calculate the sample size needed.	• Write protocol and calculate sample size needed.
• Conduct formative assessment.	• Conduct formative assessment.
• Identify clusters through ethnographic mapping.	• Start with initial contacts or seeds who are surveyed and become recruiters.
• Construct a sample frame of clusters, defining clusters by both location and time if the population is floating.	• Recruiters each recruit up to three people they know in the high-risk group.
• Select clusters and individuals in clusters using equal probability sampling.	• The new recruits become the recruiters.
	• Several recruitment waves occur.
Requirements:	**Requirements:**
• Extensive mapping and ethnographic work is needed.	• The target population must have networks.
• A sample frame of location and time (clusters) and individuals in selected clusters is needed.	• It is necessary to keep track of links between recruiters and recruits.
• Coordination with gatekeepers (bar owners, pimps, etc.) is needed.	• The size of each person's network must be documented.
• The target population should be identifiable and approachable for interviewers.	• Sample must reach equilibrium.
• Estimating cluster size in advance is difficult, so adjusting the estimates during data analysis is important.	• Special statistical software must be used to adjust the final data.
Special considerations:	**Special considerations:**
• We can do a probability sample of populations that go to certain venues and sites.	• Can reach subgroups who do not visit certain venues or clinics.
• It is difficult to maintain randomness while selecting respondents within clusters.	• If conducted correctly, can provide unbiased estimates.
• Clusters/sites can close quickly.	• There is much less need for ethnographic mapping, and field operations can be done in office settings.
• Samples can be concentrated in geographical areas.	• As the target population recruits their peers, it still works when it is difficult to identify the target population.

be reached easily and identified by interviewers. Otherwise, RDS is recommended. Example: men who have sex with men who attend certain bars or clubs and can be recognized and approached by your field staff.

4. If there is no certain location where target populations usually congregate, and members of the target population are hard to reach in public venues, does the target population have a social network? If yes, RDS is recommended. Example: injecting drug users who are living in a city and are connected to each other due to their need for drugs and safe injecting venues.

Exercise

Think of a most-at-risk population in the context of your country. Try to decide on the best sampling approach using the information above.

Summary

You sample when you want to measure characteristics for a specific target population, but lack the time and resources to obtain information from every member of the target population. Concentrating survey time and resources on questioning a sample of people can result in better quality data than spreading resources over the entire population. It is also important to understand sampling issues and options for behavioural surveillance. Follow the criteria for choosing the most appropriate sampling approach.

Unit 5 exercises

Warm-up review

Take a few minutes now to look back at your answers for the warm-up questions at the beginning of the unit. Make any changes you want to make.

Small group discussion

Get into small groups by country, region or district to discuss this question.

    a. What sampling strategies have you used?

    b. What difficulties and successes did you have with these strategies?

    c. For each of the following groups, decide what is the best sampling strategy:

        Group 1: Youth

        Group 2: Injecting drug users

Apply what you have learned/case study

Try this case study individually. We will discuss the answers in class.

    1. Below, you will find three brief descriptions of surveys and sampling methods. Discuss each approach based on your experience and the information we have just covered on sampling:

- Identify a problem or two in each approach.

- Explain what you would do to adjust the sampling approach.

    a. You are participating in the planning for a population survey of all men and women in the country aged 16–55 years. Interview teams will visit homes in 50 clusters (areas chosen for the survey by the National Bureau of Statistics) from 10:00 to 14:00 daily, Monday–Friday. Men and women who consent to the survey will be interviewed on various health and economic topics.

b. A nongovernmental organization conducted a survey among men who have sex with men in a large city to determine the percentage of condom use. In March, 100 men were interviewed. Condom use was 20%. In September, interviewers returned to the same area to interview another 100 men. This time, condom use was 24%.

The nongovernmental organization's report said that condom use was increasing in all men who have sex with men due to its outreach efforts.

c. At a truck stop, a university group interviewed (and tested for HIV) a total of 150 truck drivers and their assistants. HIV prevalence was determined to be 8% (range of 5% to 11%). The university's final report stated that 8% (range of 5% to 11%) of truckers in the country are infected with HIV. The report included a plan to provide more care and treatment along all truck routes.

2. A survey of men who have sex with men using RDS in your city found that 80% could correctly name the ways that HIV is transmitted and ways that HIV can be prevented.

Health education campaigns are planned for the coming year using men who have sex with men peer educators. You wish to show that HIV/AIDS knowledge has improved by at least 10% by the following year using the same method.

How many men who have sex with men do you need to include in this year's survey and next year's survey? Use Table 5.2. Assume that knowledge may increase or decrease and that you wish to have 80% power. Try to check your answer using the sample size formula to detect the change in proportion (Appendix A.5.2).

3. You would like to study sex workers in the capital city of Menaland using RDS. Your main objective is to obtain a population-based estimate of health care utilization among sex workers in the city. You have defined the health care utilization indicator as the proportion of sex workers who have ever sought treatment or consultation for reproductive or sexual health from public clinics, private clinics or pharmacies in the past year.

To the best of your knowledge, there is no data on the rate of STIs or STI clinic attendance among sex workers in Menaland. Furthermore, there is little data on health care utilization in Menaland and other similar countries. Therefore, the sample size calculation has to be done based on your projection considering data on health care utilization from studies conducted elsewhere. You project that "ever" health care utilization in your proposed survey could be 20% among sex workers in Menaland.

a. Using the sample size formula for point prevalence, a two-sided alpha of 0.05 and a 4% margin of error, calculate how many subjects you must recruit to show that the population-based estimate of the prevalence of health care utilization is 20%.

b. Would you consider adjusting your estimated sample size because you used RDS? Demonstrate how.

c. Would it make any difference if you knew that the estimated size of the sex worker population in Menaland is about 2000?

Appendix A.5.1

Summary of probability sampling techniques

Sampling	Steps	Advantages	Disadvantages
Simple random	1. Construct sample frame for survey population 2. Select people randomly from the sample frame using a random numbers table or lottery draw	• Concept is easy to understand and analyse	• Requires sample frame of entire target population • Logistically difficult if sample is geographically dispersed • Using a random number table/lottery draw is time-consuming
Systematic	1. Create a list of the target population when you know its total number 2. Calculate sampling interval (SI) by dividing total target population by the sample size 3. Select random start between 1 and SI and select that person 4. Add SI to random start and select person, etc.	• Random numbers table or lottery draw not required • Easy to analyse	• Logistically difficult if sample is geographically dispersed
Stratified	1. Define the strata and construct a sample frame for each stratum 2. Take a simple/ systematic sample from each stratum 3. Calculate indicator estimates for each stratum and for the population	• Produces unbiased estimates of indicators for the strata • Can increase precision of indicator estimates	• Requires sample frame of entire survey population • Logistically difficult if sample is geographically dispersed • Requires a sample large enough to make precise estimates for each stratum • Population estimates require weighting

Sampling	Steps	Advantages	Disadvantages
Cluster Probability proportional to size (PPS) or equal probability sampling	1. Construct sample frame of clusters 2. Calculate SI, select random start between 1 and SI 3. Select cluster whose cumulative size contains the random start 4. Add SI to random start and select cluster 5. Sample equals numbers of people from selected clusters	● Only need sample frame of clusters and individuals in selected clusters ● Sample is concentrated in geographical areas	● Decreases precision of estimates and thus requires larger sample size ● Size of clusters required prior to sampling
Cluster Equal probability, fixed cluster size	1. Construct sample frame of clusters 2. Select clusters using simple/systematic sampling 3. Sample is equal numbers of people from selected clusters	● Only need sample frame of clusters and individuals in selected clusters ● Sample is concentrated in geographical areas ● Need cluster sizes only prior to sampling	● Decreases precision of estimates and thus requires larger sample size ● Weighted analysis required for unbiased estimates ● Size of clusters required for weighted analysis
Cluster Equal probability, proportional cluster size	1. Construct sample frame of clusters 2. Select cluster using simple/systematic sampling 3. Sample is equal proportions of people per cluster	● Only need sample frame of clusters and individuals in selected clusters ● Sample is concentrated in geographical areas	● Decreases precision of estimates and thus requires larger sample size ● Size of clusters required for proportional sampling ● Sample size and thus precision of estimates is unpredictable

Appendix A.5.2

Formula for sample size calculation to detect change in proportion [7]

The general formula for the needed sample size (n) to detect change in the proportion of an indicator from one round to the next is:

$$n = D \frac{[Z_{1-\alpha} \sqrt{2P(1-P)} + Z_{1-\beta} \sqrt{P_1(1-P_1) + P_2(1-P_1)}]^2}{(P_2 - P_1)^2}$$

Where:

n = Sample size required per survey round (year)

D = Design effect (see above)

$Z_{1-\alpha}$ = The z score for the desired confidence level, usually 1.96 for 95%

$Z_{1-\beta}$ = The z score for the desired power, usually 0.83 for 80%

P_1 = The proportion of the sample reporting indicator in year 1

P_2 = The proportion of the sample reporting indicator in year 2

P = $(P_1 + P_2)/2$

Example: you would like enough sex workers in your survey rounds to show that condom use at last paid sex increased from 20% in the year 2006 to 30% in 2007. Choosing the values of these numbers is based on the following considerations.

Suppose you are planning a survey of sex workers using a two-stage cluster design. You wish to show that condom use will increase from 20% in the baseline survey (this year) to 30% or greater in the survey wave next year. How many sex workers do you need to include each year?

Solution:

D = 2 (moderate)

$Z_{1-\alpha}$ = 1.96 (95% confidence level)

$Z_{1-\beta}$ = 0.83 (80% power)

P_1 = 20% condom use in year 1

P_2 = 30% condom use in year 2

P = (0.20 + 0.30)/2 = 0.25

N = $2\{1.96\ SQT[2\times0.25(1-0.25)]+0.83\ SQT[0.20(1-0.20)+0.30(1-0.3))]\}^2/\ (0.30-0.20)^2$

= 582 sex workers per survey wave

P_1 and P_2: P_1 and P_2 are the measures of interest for which you wish to see a change between survey rounds. For example, you wish to show that condom use at last paid sex for sex workers increased from 20% in 2006 (P_1) to 30% or greater in 2007 (P_2). P_1 is usually based on previous surveys in the same or similar population, or an educated guess at what the level will be. P_2 is ideally set at the goal you would like to achieve such as a 10% or greater increase in condom use. In practice, it is usually set at the smallest change you think is meaningful. For example, a 10% increase in condom use would be considered a meaningful improvement, whereas a 1% increase would not be considered meaningful. The smaller the change you wish to detect, the larger the sample size you will need. Also, the closer P_1 and P_2 are to 50%, the larger the sample size you will need.

$Z_{1-\alpha}$: The $Z_{1-\alpha}$ score is a statistic that corresponds to the level of significance desired. Usually, a significance level of 0.05 (or, equivalently, a 95% confidence level) is selected and corresponds to a value of 1.96. This value is used when the change in the indicator might be either up (increase) or down (decrease) from year to year (a "two-tailed" statistic). The smaller the significance level (that is, a higher confidence level), the larger the sample size you will need.

$Z_{1-\beta}$: The $Z_{1-\beta}$ score is a statistic that corresponds to the power desired. Usually, 80% power is selected and corresponds to a value of 0.83. This value is used when the change in the indicator might be either up (increase) or down (decrease) from year to year (a "two-tailed" statistic). The higher the power, the larger the sample size you will need.

Appendix A.5.3

Random numbers table and instructions [8]

NOTE: Several computer applications, such as Excel, Access, SAS, and CSPro, can generate random numbers. If you are familiar with these programs, you can generate numbers electronically rather than using the random numbers table (Table A.5.3.1).

Instructions for using the random numbers table

1. Determine how many digits you need your random number to be, based on the total number of households.

2. Choose a direction (right, left, up or down) in which you will read the numbers from the table. You will read the numbers in this direction for all random numbers selected for the sample.

3. With your eyes closed, use a pointed object, such as a pen or pencil, to touch the random numbers table. Your starting point is the digit closest to the point where you touched the table.

4. In the direction you chose, read the number of digits required. Numbers that are not within the range needed are discarded. Continue reading the numbers in the chosen direction until all random numbers have been selected.

Examples for random sampling

For each area to be sampled, you must determine how many households there are and how many need to be selected for your sample. You determine that you need to select 75 households from a total of 228 households in Zone 1. So you will need to select 75 numbers between the numbers of 1 and 228 following the steps below. In this example, you are choosing a number between 1 and 228, so you need 3-digit numbers.

Example 1

You have decided that you will move to the right of where your pencil lands and you need 3-digit numbers between 1 and 228. Your pencil lands on the last digit in the cell in Column C, Row 2.

Your pencil should be on the digit "0." Reading to the right to get a 3-digit number gives you the number "084." This means that you will ask household #84 on your list of numbered households to participate in the survey. Reading to the right to get the next 3-digit number gives you the number "443." Since this number is not between 1 and 228, you continue to the next number until it is in the correct range. The next number that fits the range is "015." Repeat this process until all 75 random numbers have been selected.

Example 2

You have decided that you will move down from where your pencil lands and you need 3-digit numbers between 1 and 228. Your pencil lands on the second digit in the cell in Column H, Row 16. Your pencil should be on the digit "1." Reading down to get a 3-digit number gives you the number "123." This means that you will ask household #123 on your list of numbered households to participate in the survey. You will repeat this process until all 75 random numbers have been selected.

Example 3

You have decided to move to the left of where your pencil lands and you need 3-digit numbers between 1 and 228. Your pencil lands on the first digit in Column D, Row 27. Your pencil should be on the digit "8." Reading left to get a 3-digit number gives you the number "879." Because 879 is not between 1 and 228, you must choose a new starting point. You try again and your pencil lands on the fourth digit in Column J, Row 11. Your pencil should be on the digit "1." Reading to the left to get a 3-digit number gives you the number "117." You will repeat this process until all 75 random numbers have been selected.

Examples for cluster sampling

You will need to use the random numbers table for two tasks in cluster sampling. The first task is to choose a starting point between 1 and your sampling interval to select the clusters.

Example 4

In our example, we need a number between 1 and 39 (our sampling interval). You have decided that you will move to the right of where your pencil lands and you need a 2-digit number between 1 and 39. Your pencil lands on the last digit in the cell in Column C, Row 45. Your pencil should be on the digit "2." Reading to the right to get a 2-digit number gives you the number "29." The number "29" falls between 28 and 60, so you would start your selection of clusters with cluster #2.

For the second task, you need to randomly select households from each selected cluster. For each selected cluster, you need to choose 25 households. Let's say that you have chosen cluster #15 which has 40 households. You will need to choose 25 numbers between 1 and 40. In this example, you are choosing a number between 1 and 40, so you need a 2-digit number. The numbers can range from 1–40.

Example 5

You have decided that you will move to the right of where your pencil lands and you need 2-digit numbers between 1 and 40. Your pencil lands on the last digit in the cell in Column E, Row 10. Your pencil should be on the digit "2." Reading to the right to get a 2-digit number gives you the number "22." This means that you will ask household #22 on your list of numbered households to participate in the survey. You will repeat this process until all 25 random numbers have been selected.

Example 6

You have decided that you will move up from where your pencil lands and you need 2-digit numbers between 1 and 40. Your pencil lands on

the second digit in the cell in Column F, Row 23. Your pencil should be on the digit "0." Reading up to get a 2-digit number gives you the number "03." This means that you will ask household #3 on your list of numbered households to participate in the survey. You will repeat this process until all 25 random numbers have been selected.

Example 7

You have decided to move to the left of where your pencil lands and you need 2-digit numbers between 1 and 40. Your pencil lands on the first digit in Column I, Row 26. Your pencil should be on the digit "5." Reading left to get a 2-digit number gives you the number "53." Because 53 is not between 1 and 40, you must choose a new starting point. You try again and your pencil lands on the fourth digit in Column D, Row 8. Your pencil should be on the digit "3." Reading to the left to get a 2-digit number gives you the number "39." You will repeat this process until all 25 random numbers have been selected.

Table A.5.3.1 Random numbers table

	A	B	C	D	E	F	G	H	I	J
1	8450	6992	6563	0340	2649	6933	9446	6182	2601	7800
2	5952	1443	7100	8444	3904	0159	1849	2601	9763	9058
3	5711	6779	9388	9668	4167	1423	2744	4622	2179	8503
4	2681	8047	0494	7853	8411	5406	8127	9577	8530	2350
5	0739	3114	3997	3482	3226	2216	6874	0620	8521	2938
6	8985	2463	5054	3448	6357	0187	6342	4740	4064	5068
7	7644	9339	8375	4583	7715	6355	6827	2055	9328	3287
8	6277	6631	8797	3693	6370	1436	1599	6267	2758	0323
9	6355	7590	7628	9054	0022	4241	7499	3430	3644	6576
10	7828	0589	3075	1954	5972	2266	0055	1097	9706	9009
11	6026	4546	4119	1554	4895	3123	9849	2094	5062	6711
12	8416	1972	9345	1593	2943	2379	5062	4829	5952	8292
13	1433	8823	7706	5273	6160	2161	5510	8617	7894	0175
14	0622	4884	8113	4447	5735	6347	7280	2301	2330	0693
15	4104	7164	1184	3964	2119	6968	0469	3827	0845	8400
16	4272	4979	1471	0942	9573	4283	1557	0161	3957	2516
17	1225	4171	3433	8700	0042	5884	2508	3250	1520	6366
18	7442	6575	1927	7267	7182	3960	4341	0350	1126	5945
19	4911	9007	3048	0319	0916	3002	1466	4421	7246	7662
20	3143	7402	4486	0909	1858	7961	1211	6296	5545	4588
21	8055	9294	2578	0426	4322	6925	2487	5677	9491	4301
22	9240	5260	7134	8001	0140	3394	8437	4066	2855	0933
23	7923	8630	3654	2638	2868	1059	0903	3114	6351	8261
24	0020	5104	4344	3324	9214	6615	5926	7012	9052	9205
25	3312	5923	5469	9171	4877	5392	3394	5077	3750	5637
26	3466	4193	5330	4680	0456	5891	3175	5733	5678	0956
27	1677	1694	1697	8921	2520	2811	3597	1355	9605	3637
28	3846	6283	0969	0051	5857	1043	1671	2013	8955	7706
29	8084	2327	0550	7231	1087	4830	9742	5654	5458	8290
30	2715	2247	4504	1374	9236	7340	1773	0693	2749	1335
31	6537	5815	9312	1460	6593	7678	4312	7537	9360	7195
32	4263	8931	1642	6694	1925	2661	1274	7346	234	3159
33	7468	4077	6691	3961	7640	2355	9938	8485	9398	8364
34	4884	3324	3690	7433	1245	0523	4483	5933	5634	0512
35	7222	7299	1346	8937	0933	1569	5562	3735	2982	5966
36	5040	0820	8606	4006	4743	6343	4873	1002	4757	1075
37	2980	4860	5694	1501	5791	9414	7246	1283	9766	7427
38	8660	5480	7436	9745	8869	3307	4916	6543	9830	6099
39	7627	4959	6417	3542	1877	0370	5464	9590	5184	7379
40	1890	7664	7144	3523	8465	0385	8174	4740	3654	5543
41	3175	2580	3919	7436	0796	1018	5565	1142	4577	0457
42	7616	9338	6304	0283	6502	9085	5443	1531	9724	4140
43	5223	4525	0895	9930	0050	2201	5270	6447	1850	2070
44	9384	9794	8418	0374	4119	2075	0067	4535	7769	4719
45	5862	9165	5302	9789	5771	9670	7523	9280	2604	0212
46	9450	9307	6597	7183	5243	8854	6735	2415	0364	3096

Unit 6

Data analysis and use

Overview

What this unit is about

IN THIS UNIT, you will learn the skills required to ensure behavioural surveillance data are analysed, disseminated and used appropriately.

Warm-up questions

1. True or false? It is better for staff members to share overall responsibility for data management. Circle your answer.

   True   False

2. Data management does not include which of the following?
   a. data coding
   b. data entry
   c. data cleaning and checking
   d. data framing

3. True or false? The data manager should not be involved during the questionnaire design.

   True   False

4. Which type of behavioural surveillance analysis is performed to determine whether one variable is related to the distribution of another (for example, an association between a respondent's age and use of condoms)?
   a. univariate
   b. bivariate
   c. multivariate

5. Most of the indicators defined for behavioural surveillance purposes are calculated using _____ analysis.

    a. univariate

    b. bivariate

    c. multivariate

6. True or false? The surveillance cycle ends when the official report is published.

    True   False

7. True or false? Data from other sources, including other countries, can be used to fill in any important gaps as long as the source is made very clear.

    True   False

Introduction

What you will learn

By the end of this unit, you should be able to:

- discuss data management issues
- describe the types of data analysis commonly used in behavioural surveillance
- understand the steps for appropriate data analysis and use
- list the different audiences for the data from behavioural surveillance
- interpret and package data appropriately for the different audiences.

Data management issues and activities

What is data management?

Collecting data is expensive and time consuming, so efficient data management is important. Data management includes:

- data coding
- data entry
- keeping the data safe
- cleaning and checking the data
- getting the data ready for analysis (merging, coding).

Errors in these tasks can affect the quality and usefulness of the data. You should plan data management before doing data collection. Establish simple guidelines for coding, storing questionnaires, data entry and database management. You should follow these guidelines from the start and document any change in them. Make sure that data management is discussed in surveillance meetings so that it is considered by the whole surveillance team.

Issues to consider in data management

Personnel: Data managers with experience and skills are needed to manage large and complex datasets. It is usually better for one person to have overall responsibility for the data than to share responsibility among different staff members. The data manager should be involved during the design of the questionnaire to review the questions and the way the questionnaire is organized. Make sure that data-entry clerks and interviewers get adequate training to minimize errors. You should also make sure that their work is monitored carefully.

- Computer capacity: Analyse what you may need for computer capacity. Databases often take up more space than we expect. It is important to have enough capacity for data storage.
- Back-up routine: Back up your data onto a separate hard drive often to prevent loss of some or all of your data files.
- Audit trail: Document everything you do and when you do it, such as data entry, verification and corrections. Include notes about any queries. Record changes to the

master dataset on paper to keep an audit trail. You can also have an electronic logging system that records all changes.

- Data checking: Data errors can be introduced at any stage of the surveillance process. Be cautious throughout the survey. Data checks should include:
  - consistency checks by supervisors after data collection
  - double data entry (it is unlikely that two people will make the same mistake)
  - data entry programmes to check that values stay within specified realistic ranges
  - programmes written specifically to check for inconsistencies in the data, so any errors can be checked back to the questionnaires and the interviewers.
- Database merging: With several data entry clerks working on the data, and with several copies of the database, you need a method for merging the data after data entry.
- Management of paper questionnaires: Make sure questionnaires are protected from damage and are kept confidential. Signing questionnaires at each stage of the data collection (data checking) is a useful way to keep track of which data are already entered.
- Storage and archiving: Develop a system for archiving raw data. The data management system should:
  - document data effectively
  - allow you to find the data you need quickly
  - keep data archives and back-ups safe, current and usable.

Data analysis

What is data analysis?

Data analysis is processing raw (recently collected) data to produce study results. Data analysis includes summarizing, presenting and interpreting data. Behavioural surveillance analysis can be:

- cross-sectional, meaning that the data are analysed for just one point in time—for instance, from one surveillance round
- trends, meaning that changes in indicators are analysed between two or more points in time—for instance, across several surveillance rounds.

If probability sampling was used, statistical tests can tell us how well the sample measure estimates the true population measure. That is, they can tell us how likely it is that our findings could have occurred by a chance sampling variation or that they represent true population values. You need special analytical techniques if the sample was a cluster sample or if the cluster sample did not use a self-weighted design.

Types of data analysis

There are three main types of analysis conducted in behavioural surveillance:

- Univariate analysis: the most basic type of behavioural surveillance analysis. It is often the most important, because it shows the distribution of each variable. Most of the indicators defined for behavioural surveillance are calculated through univariate analysis. These indicators include variables like the proportion of young men who

have had sex with more than one partner during a given time period. When trends are analysed, statistical techniques are used to calculate how likely it is that changes in the proportions could have occurred by chance or that they reflect real change.

- Bivariate analysis: analysis to find out whether one variable is associated with another. For example, there might be an association between a respondent's age (the explanatory or predictor variable) and their use of condoms (the outcome variable). Variables are associated if the value of one tells you something about the value of another. Statistical tests in bivariate analysis determine how likely it is that any observed difference may be due to chance.
- Multivariate analysis: analysis to look at the influence of at least two variables on another variable. Relations between variables are often complex and interwoven. You can use multivariate analysis to find the individual effects of several explanatory and possibly related variables on an outcome variable.

Steps in data analysis and use

Follow these four steps when performing data analysis:

1. Develop an analysis plan.
2. Explore the data.
3. Use appropriate statistical techniques.
4. Interpret the data.

Step 1: Develop an analysis plan

Before data collection begins, you should make an analysis plan. This helps you make sure that the correct data are collected and that enough time and resources are given to data analysis. The plan should:

- List the questions you want the surveillance to answer. This should be done in consultation with the final data users so that their data needs are met. Consider the social, cultural and political context in which the final result should be interpreted.
- Define the indicators needed to answer these questions.
- Make examples of important tables and graphs to be sure that you know how all the data will be used.
- Identify staff with the proper skills. It is important to involve statisticians from the start. If they understand the purpose of the research and are familiar with the questionnaires, they will be more helpful than if they are brought in only at the end to do multivariate analysis.
- Provide enough money and time for analysis. Cross-checking, recoding and creating indicators is often what takes the most time.

Step 2: Explore the data

Data analysis is usually done by a trained statistician. Once the data have been collected and cleaned, the statistician should explore the data first to understand the data coding and find any errors in the data. This should be followed by making frequency distributions of the variables and recoding the data to create the indicators. To make

bivariate analysis even more relevant, univariate frequencies should be closely examined. Once the descriptive summaries have been done, the statistician should look carefully at the results, such as checking for consistency in sample sizes and for missing values, to see if the data make sense.

Step 3: Use appropriate statistical techniques

The results of your survey will guide further analysis. Most behavioural surveillance analysis will be univariate and bivariate. Analysis should start by being simple but with a broad focus. Important findings can be missed if analysis immediately focuses on details.

Remember that complex multivariate analysis is not always appropriate, particularly if data quality is poor or sample sizes are too small.

Response rates always should be reported in your results. Data analysis must take into account possible biases for non-respondents. Recording the following data can be useful:

- number of target group members present at the cluster at the time of interview
- number interviewed
- number of people who refused to be interviewed, each person's reason for refusing and basic sociodemographic information about each person
- number of people who were not interviewed for other reasons
- number of interviews rejected as duplicates.

Step 4: Interpret the data

Surveillance data can often be interpreted in more than one way. How indicators are defined and presented can affect how the data are interpreted. Also, it is import to be careful of personal biases, which can cause some researchers to give greater weight to certain interpretations.

It is very important to examine the data on population size, HIV prevalence, prevalence of STIs and risk behaviour together.

The validity of self-reported data on sexual behaviour and drug use can often be questionable. Surveillance data are worthwhile, but we must be cautious about coming to quick conclusions about small rises or falls in infection rates or surprising findings in behavioural surveillance. Investigate further.

Using behavioural surveillance data

Introduction

In most surveillance systems, a gap remains between the collection and use of data. One reason data are not better used is that surveillance systems can be fragmented. When no single group or individual is responsible for compiling, analysing and presenting data in

an organized way, the different groups involved in surveillance often consider their job done when they finish their survey. Before data are used, form a surveillance committee to help resolve this issue.

There are two steps to follow in using behavioural surveillance data:

- Develop a plan for data use
- Present data.

Develop a plan for data use

Before data collection begins, you should use the data use plan that was developed as part of the surveillance protocol described in Unit 1, Table 1.4. This helps ensure that information gets to the right people in a timely manner and in a way they understand.

This plan should have been developed to identify the country's data needs and therefore the data's audience, meaning the people who will use the surveillance data.

At this stage, it can be useful to identify surveillance outputs and how they will be distributed. Outputs could include a national report on HIV/AIDS, policy briefs, press releases, and so forth.

Once reporting methods are identified, enough money and time can be set aside for these activities. If possible, begin to build networks for future advocacy before you begin surveillance activities. It is very important to establish good relationships with policy-makers and other data users, including the following:

- politicians and policy-makers
- donors
- AIDS programme managers in health and other sectors
- persons from international agencies
- nongovernmental organizations
- researchers
- populations included in the surveillance system
- surveillance system staff
- persons in the private sector
- the press
- persons in the legal profession.

Present data

Presenting data appropriately and accurately helps us make sure that it is used. When data are presented, consider the following issues:

- Data need to be packaged appropriately for the various audiences: different audiences will respond best to different ways of packaging the data. For example, face-to-face meetings, videos or briefing sessions are often more effective ways to interest decision-makers than lengthy documents that they do not have time to read. The process of choosing the right product for the right audience will be covered in detail later in this unit.

- Be clear about the limitations of the data (sample size, response rates, level of quality): surveillance data do not need to be perfect. They just need to be accurate enough to give a reliable idea of the major trends in HIV infection and related risk behaviour. Once numbers are presented as facts, the information that would help people figure out the quality of the data can be lost. Limitations in the data in terms of sample size and response rates need to be reported. The data should not be presented as more "scientific" than they are. For example, surveillance data are sometimes presented to two decimal points of accuracy, giving an air of statistical solidity. But presenting rounded numbers might be a more accurate reflection of your data quality.
- Use all data sources available: surveillance systems rarely produce all of the data that meets local advocacy needs. Data from other sources, including other countries, can be used to fill in any important gaps as long as the source is made very clear.
- Be careful how you present the data: most people who work in public health can look at a table or graph and understand what it means, but many people in other fields cannot. If needed, you can use headlines to tell people what the data mean, rather than simply describing the data. Generally, graphics should be kept as simple as possible and free of clutter.

Involve important stakeholders

To make sure that your data will eventually be used for programmes, try to involve stakeholders in interpretation of data, build consensus around the most important findings, and get support from partners on use of the data for future programme planning.

Choose the right approach for the audience

Public health officials often use one approach to data use, believing that surveillance work ends when the official surveillance report is published. The report may contain all the important information about the levels and trends of HIV and risk behaviours. However, the same data need to be presented differently for different audiences to emphasize the findings and get them to act on the data.

Successful advocacy follows a number of rules. *Guidelines for the effective use of data from HIV surveillance systems* (2004) [9] contains information about how to present data and good examples of how HIV data can be used. Table 6.1, outlines some of this information.

Table 6.1 Tips for successful advocacy

What you want to do	Things to think about
Define your goals	Surveillance data can meet several goals and data needs. Data for each goal may have to be presented to different audiences in different ways.
Define your audience	Who must the public health official communicate with to make sure that the goals are met? Who has the potential to push things forward? Who could potentially obstruct progress?
Find out what influences your audience's thinking and how to get their attention	Determine the best way to communicate with your audience. Find what messages are important for different audiences. The best communication occurs when people have something in common. Therefore, it is important to understand the concerns, motivations and goals of each audience from its own point of view. Once you understand your audiences, try to present the data in ways that will reach them.
Use the right language	After you identify the audiences, their goals and messages that will appeal to them and how to present the data, the language you use is important. For example, will confidence intervals mean anything to the group or will that term confuse people?
Get the length right	A report or presentation has no value if no one reads it or listens. This means fitting the information into the time that people are prepared to dedicate to it. If HIV/AIDS is the core interest of your audience, they may appreciate a full report they can read through and keep for their reference. If HIV/AIDS is not a core area of interest, the audience is more likely to digest the information if it is presented in a one-page fact sheet or short brochure.
Choose the best messenger	People listen to the people they trust. For example, the Minister of Health may be the best person to present surveillance data at the cabinet level, whereas young people may pay more attention to a pop star acting as an AIDS ambassador.
Get the timing right	HIV/AIDS is not the only issue on people's agendas. One way of increasing the attention the message gets is to time the release appropriately. For example, avoid clashing with important events and make use of events already scheduled, such as World AIDS Day.

Summary

Efficient data management is very important because error at any stage can affect the quality and usefulness of your data. Analysis of behavioural surveillance data can give information about the status of the HIV-related risk behaviours that drive the epidemic for one point of time (cross-sectional) or over a period of time (trend). It is important to use these steps to be sure that data are analysed correctly: develop an analysis plan, explore the data, use appropriate statistical methods and interpret the data.

You also should understand the different audiences to which behavioural surveillance data will be presented and understand how to present and package the data for their specific interests and needs.

Unit 6 exercises

Warm-up review

Take a few minutes now to look back at your answers for the warm-up questions at the beginning of the unit. Make any changes you want to make.

Small group discussion

Get into small groups by country, region or district to discuss these questions:

1. Discuss the strengths and weaknesses of the data management, data analysis and data use system used for behavioural surveillance in your setting.

2. Use data from your country's annual report. Write a paragraph describing the data for the general public.

3. Design a flyer for the general public that conveys the data in Question 2.

Apply what you have learned/case study

Try this case study individually.

What would be the best way to present your data to a senior politician?

Unit 7

Ethical considerations

Overview

What this unit is about

IN THIS UNIT, you will learn about ethical matters and requirements in behavioural surveillance.

Warm-up questions

1. Match each ethical principle with its definition.

   _____ Respect for persons

   a. Refers to maximizing benefit and minimizing risk to individuals—not only physical risk, but also risk of psychological harm and stigmatization.

   _____ Beneficence

   b. Requires investigators to see study subjects not as passive sources of data, but as persons whose decisions and personal choices must be respected.

   _____ Justice

   c. Risks and benefits from studies should be distributed fairly and evenly in populations.

2. Complete this sentence: It is important that everyone in contact with the surveillance system, from community leaders and local officials to the people surveyed, is treated with_____.

3. What is informed consent?

4. Name two pieces of information that should be given before a person can make an informed decision to take part in a survey?

   a.

   b.

5. True or false? Obtaining verbal consents to survey a most-at-risk population is sometimes better than written consents because we can protect the confidentiality of participants more rigorously. Circle your answer below.

    True   False

6. What are two ways to ensure a participant's confidentiality?
    a.
    b.

7. Name two strategies to protect survey participants from loss of earnings.
    a.
    b.

Introduction

What you will learn

By the end of this unit, you should be able to:

- explain the basic ethical principles of working with human subjects
- define informed consent and the procedures that are used
- understand the importance of confidentiality and how to ensure it
- discuss ethical considerations unique to behavioural surveillance.

Ethical principles of conducting human subjects research

The principles of medical ethics are most often applied to medical and public health research. Behavioural surveillance is modelled on public health research and uses informed consent. We describe below the principles of medical ethics as they apply to behavioural surveillance.

Three accepted ethical principles for conducting research involving human subjects are respect for persons, beneficence and justice:

- Respect for persons requires investigators to see study participants not as passive sources of data, but as persons whose decisions and personal choices whether to participate in the study must be respected.
- Beneficence refers to increasing benefits and reducing risk to individuals. The results of the survey must benefit the participants. Also, the risks of participation must be low in relation to the benefits. The investigators must be capable of doing the study scientifically. They must also protect participants against physical risk and risk of psychological harm and stigma.
- Justice means that risks and benefits from studies should be distributed fairly and evenly in populations.

For further information on ethical principles, please refer to the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects [10] which can be found in Appendix A.7.2. Also see the International Ethical Guidelines for Biomedical Research Involving Human Subjects [11] prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the WHO. Available online at: http://www.cioms.ch/frame_guidelines_nov_2002.htm

It is important that everyone in contact with the surveillance team is treated with respect, from community leaders and local officials to the people who are surveyed. Respect is important in every part of the study, including the following:

- engaging people in the study
- conducting the interviews
- analysing and presenting data
- discussing results with the participants and communities involved.

Informed consent

What is informed consent?

Before being interviewed, respondents need to decide if they want to participate in the survey. Informed consent involves telling people enough about the nature of the surveillance for them to make a proper and knowledgeable (informed) decision about whether to take part.

No project staff member should pressure or deceive respondents to get them to participate. Staff should make sure that respondents are not pressured to participate by other people, including family or community members. Being certain that the decision to participate is not influenced by staff members may involve discussions with gatekeepers about the voluntary nature of participation. The interviewers' training and procedure manual for the surveillance should include sections on informed consent.

Obtaining informed consent

A person can only make an informed decision to take part in a survey (give their consent) when he or she has information about the procedures and purpose of the study.

During the informed consent process, subjects are provided with:

- information about the survey
- an opportunity to ask questions about the survey
- an opportunity to decide whether to participate.

The details of informed consent are usually on an information sheet that the interviewer reads to the participant. Studies have shown that getting voluntary and informed consent is sometimes very difficult. Make sure that the information sheet uses appropriate language, is not so long that participants stop reading or listening, and contains all the necessary information.

You should provide the following kinds of information to potential participants:

- the nature of the survey, such as purpose of the survey, length of interview, types of questions
- who is conducting the survey, such as the Ministry of Health or a university
- the potential risks and benefits
- how the information will be used
- how the participant's privacy will be protected; for instance, names or addresses will not be taken
- that participation is voluntary and there are no negative consequences for refusing to participate; prisoners, for example, may believe that refusal could lead to loss of their normal rights
- that participants have the right to refuse to answer any questions or stop the interview at any time, especially if they find some of the questions uncomfortable
- that participants have the right to have deleted any recorded data they consider harmful to themselves.

Documenting informed consent

Written consent forms are generally needed to record that the process of informed consent has occurred. In surveillance, to be certain of total confidentiality, it is usually best to obtain verbal consent. The name of the respondent does not need to be recorded. Instead of requiring the participant's signature, interviewers can sign a statement verifying that the respondent has been given the information and has decided to participate.

Balance between consent and participation

Although consent must be voluntary, you will want to maximize participation and reduce refusal. Reducing refusal is important. People who refuse to be interviewed may be different from people who participate. If there are many refusals, our sample may not be like the population from which it was taken and thus may be biased.

To address non-participation bias, you need to aid participation without pressuring:

- Keep interviews as short as possible.
- Conduct fieldwork at times that are convenient to the participants.
- Stress truthfully the benefits of participating.

As we discussed in Unit 6, refusals should always be reported so that the analysis is complete.

Using incentives

Incentives can be cash payments for participation or small gifts, such as T-shirts, food stamps and hygienic goods. In general, incentives are considered appropriate for compensating or thanking study participants for time away from work and personal expenses, such as transportation. Higher payments may endanger the voluntary nature of informed consent. They can create a situation where a person's decision to participate is overly influenced by money or gifts.

Using incentives may result in a sample that is not like the population of interest because the sample is biased toward people who have a greater financial or other need. However, not using incentives may bias the sample toward people who are more cooperative.

In RDS, surveillance staff may provide additional incentives to participants to recruit other members of the population to the study. These incentives can be considered payment to the participants, who act as fieldworkers in their role as recruiters. This part of the methodology may be controversial in some settings and may require explanation to the IRB reviewing and overseeing the study.

Confidentiality

What is confidentiality in behavioural surveillance?

Confidentiality protects participants from negative outcomes that may happen if other people know their responses or know that they participated. For example, if information about a person's sexual practices is disclosed, he or she may suffer discrimination, stigma or

even prosecution. Potential threats to confidentiality, as well as measures taken to minimize them, should be discussed with the participants as part of the informed consent process.

Conducting surveys anonymously is the most secure way to protect confidentiality. Participants are not required to reveal their identities to be able to take part in the survey. However, to reduce the number of duplicated participants, you will need to record some sort of identification information.

Securing confidentiality

Confidentiality can be improved by holding interviews in a place that is private, such as your surveillance facility.

When surveys are done outside of your surveillance facility, such as in places where target populations congregate, the presence of other people breaches confidentiality. This may cause the participant embarrassment and might influence some of his or her answers. In such cases, the interviewer can explain that some questions are very personal and ask the participant to suggest a place where they are unlikely to be disturbed.

Sometimes interviewers hear stories during interviews that make them so sad that they must to talk with others about it. Interviewer training should stress that, although it is important to talk about sad issues, this should only be done with other team members and in a way that does not identify the respondent.

Measures that can be taken to protect confidentiality include:

- Make sure that names or other means of identification, such as addresses or telephone numbers, are not recorded on surveys.
- Store surveys and electronic data safely and appropriately.
- Train interviewers in the importance of confidentiality.
- Have clear disciplinary procedures for staff members who breach confidentiality.
- Identify problems and possible solutions related to confidentiality.

Ethical considerations unique to behavioural surveillance

Working with most-at-risk and vulnerable populations

Some members of most-at-risk and vulnerable populations, such as men who have sex with men, sex workers and injecting drug users, practise illegal or stigmatized behaviour. Surveillance activities may make stigmatization worse. A list of potential harms is listed in Table 7.1, below.

Table 7.1 Potential harms caused by behavioural surveillance

Potential harms	Example
Physical	Public attack, abuse, loss of health-care services
Legal	Arrest, prosecution
Social	Disclosure to family and/or workplace, discrimination, loss of employment, isolation

If members of a most-at-risk or vulnerable population fear that information about their behaviour may be used against them, they will probably refuse to participate in the surveillance activities. Worse, they may participate in these activities and the results of the surveillance may lead to greater stigmatization or repressive measures by law enforcement agencies.

Other potential ethical issues and potential solutions are identified in Tables 7.2 and 7.3.

Table 7.2 Potential ethical issues and solutions regarding inclusion in behavioural surveillance

Potential ethical issue	Potential solution
Increase in stigmatization and discrimination against the group	• Make certain of fully informed consent and confidentiality. • Keep data reporting neutral (without negative opinions or judgments). • Along with data reporting, do public health communications about the negative effects of stigma and discrimination on the epidemic.
Loss of earnings	• Keep interviews as short as possible to reduce missed work time. • Pay back lost earnings. • Conduct interviews outside of work times.
Gatekeeper gets angry at people who participate	• Involve and work with gatekeepers. • Emphasize the benefits of surveillance.
Gatekeeper forces participation	• Involve and work with the gatekeepers. • Emphasize the importance of voluntary participation.
Illegal activities are exposed, resulting in a police response	• Involve and work with law enforcement agencies. • Help them understand the purpose of surveillance and the damage that could result (scattering the high-risk groups, driving groups underground) from repressive measures.
Participants get no direct benefit from surveillance	• Report findings back to survey population. • Explain the indirect benefits during the informed consent procedure. • Do not foster false expectations.
The Ministry of Health does not have an established ethics committee	• Establish an ethics committee. • Use the ethics committee of a local university, hospital or other research institution.
Government authorities demand names of people living with HIV, men who have sex with men, sex workers, foreign workers with HIV infection, etc.	• Use informed consent. • Do not do the study. • Find legal ways to get around the law. • Talk to the authorities.
Being an immigrant is not a risk factor for HIV. Injecting drug use or sexual behaviour is the risk factor. Be careful of calling immigration a risk factor because it may cause stigma.	• To plan interventions, study the population and identify risk factors that contribute to HIV transmission. • Be careful not to stigmatize people.

Table 7.3 Potential ethical issues and solutions regarding inclusion in biobehavioural surveillance

Potential ethical issues	Potential solutions
Choosing between returning HIV test results to persons tested during surveillance activities and using the unlinked anonymous testing (UAT) approach.	• When treatment is available, efforts should be made to return HIV testing results to the people tested. • Even if there is voluntary and counselling (VCT), a referral to VCT does not necessarily mean people will go and you could lose the chance to refer a person to care. • Returning results means that the patient is linked to their test result. Return test results to all persons tested using unique codes.
If HIV testing is mandatory, informed consent may not be practised. Everyone should have a choice and be given options.	• Use informed consent. • Use UAT. • Use strictly confidential testing linking identity with code numbers and no identifiers.
RDS sites can become drug dealing venues, so risk behaviour may occur at or near the site. It is not a confidential setting if many people are standing around the site.	• Give participants appointment times. • Post rules inside the RDS site explaining the potential problems and dangers of it becoming a drug dealing venue. • Have a site supervisor to restrict access only to people participating in the survey.
Conducting RDS seed meetings before beginning the survey when populations may want to remain anonymous to each another. If seeds cannot have a meeting together, it may be a sign of a weak network.	• Conduct seed meetings separately. • Discuss the idea of social networking and surveillance as being beneficial to most-at-risk populations and vulnerable populations.

Working with adolescents

Different countries have different laws and standards about the age at which an adolescent can participate in research concerning sexual behaviour and when parental consent is required. Become familiar with these laws in your country as part of your initial planning.

Generally, surveillance tries to minimize the number of participants in the age range of 15–18 years and avoids including people under 15 years of age. If it is necessary to include children under the age of 15, special guidance on research with children should be sought.

Balance the inclusion of adolescent participants with ethical concerns and local demographic patterns in your country. Adolescents may represent a large segment of the population you wish to survey.

Benefits to participants

In some countries, surveillance is accompanied by an HIV prevention intervention. Some would argue that surveillance is unethical in the absence of an intervention. However, incentives are usually the only direct benefit of participation.

Do not foster false expectations of direct benefits to participants. Participants may ask interviewers for assistance, such as transporting a sick person to a health facility. There are limits to what sort of assistance is practical. Each country needs to make these decisions with reference to the local context.

Interviewers and fieldworkers should not make any promises they cannot keep. However, there are indirect benefits that should be stressed during the informed consent procedure and during entry into the community. These include the following:

- improving HIV prevention and care programmes in general
- raising public awareness of and sympathy for the burden of disease in the population
- reducing stigma and causing social change, especially around HIV infection
- providing information about results to the community so the situation is better understood.

Surveillance as research

In some settings, public health surveillance does not have the same requirements as formal medical or public health research for documenting and reviewing ethical procedures. IRBs consider surveillance to be non-research when it involves routine case reporting. However, some IRBs and national authorities believe that periodic cross-sectional surveys should go through formal review as research projects.

Although these ethical safeguards may not be required legally in all countries, they are essential for sustainable and ethical surveillance activities. To find out about specific ethical requirements involved in conducting surveillance activities, refer to your local ethics committees and IRB procedures. If there is any doubt, you should seek formal review and approval.

Summary

During behavioural surveillance, you should adhere to the accepted ethical principles of working with humans: respect for persons, beneficence and justice. A person can only make an informed decision to give his or her consent to participate when he or she has information about the procedures and purpose of the survey.

Confidentiality protects subjects from negative outcomes that may happen when other people know their responses or that they participated. It is important to understand the potential ethical issues that may arise as a result of being included in behavioural surveillance, as well as some potential solutions to these problems.

Unit 7 exercises

Warm-up review

Take a few minutes now to look back at your answers for the warm-up questions at the beginning of the unit. Make any changes you want to make. We will discuss the questions and answers in a few minutes.

Small group discussion

Get into small groups by country, region or district to discuss these questions.

1. What are some of the potential social harms caused by behavioural surveillance in most-at-risk populations in your setting?

2. What ethical issues/difficulties have you/could you experience when conducting surveillance in your country?

Apply what you have learned/case study

Try this case study individually.

Design a consent form to be used with sex workers.

Compare your consent form with the two sample consent forms provided in Appendix A.7.1 and discuss the following questions:

1. Which consent form is more similar to what you have?

2. Which sample is more appropriate for the context of your country/region/district?

3. Does Sample 2 capture all information needed for participants to make an informed decision? Why or why not?

Appendix A.7.1

Samples of verbal consent forms for participation in HIV biological and behavioural surveillance surveys

Sample 1

Introduction: "My name is _____. I am working for _____. We are interviewing people here in [*name of city, region or site*] in order to find out about [*describe purpose of study*]. Have you been interviewed in the past few weeks [*or other appropriate time period*] for this study?"

If the respondent has been interviewed previously during this round of the behavioural surveillance survey, do not interview the person again. Tell them you cannot interview them a second time, thank them and end the interview. If they have not been interviewed before, continue.

Confidentiality and consent: "I am going to ask you some very personal questions that some people find difficult to answer. Your answers are completely confidential. Your name will not be written on this form, and will never be used in connection with any of the information you tell me. You do not have to answer any questions that you do not want to answer, and you may end this interview at any time. However, your honest answers to these questions will help us better understand what people think, say and do about certain kinds of behaviours. We would greatly appreciate your help in responding to this survey. The survey will take about [*interview time*] minutes. Would you be willing to participate?"

(Signature of interviewer certifying that informed consent has been given verbally by respondent)

Sample 2

Dear volunteer,

Thank you for your time. The goal of this survey is to understand the risk of sexually transmitted diseases, such as HIV, in the city of _____. We will do this through an interview and a blood test of each woman who is willing to participate. This information will be used to plan better prevention and treatment programmes.

The investigators are physicians and researchers from _____ [*name of organization(s) implementing the survey*]. Contact information for investigators of this study is listed at the end of this information sheet.

We anticipate recruiting about _____ [*approximate sample size*] eligible individuals. After you give your consent, we will interview you. Your name will not be recorded at any time during this study.

Now I am going to explain the steps that we will take to conduct our survey:

- The total time expected for this study will be approximately one hour.
- I will ask you a series of questions about your health. This should take 45 minutes.
- A professional will collect a blood sample from you for HIV testing. This will take about 15 minutes.

- After that, you will receive an envelope containing reimbursement for your participation.

- Your test result will be available in two weeks. You may drop in to this site at this same time in two weeks to collect your result. Since we do not collect your name or other identifying information, the only way we can find your test result is through your study ID number. Please keep this number in a safe place and present it to our staff when you want to receive your test results.

- In the event that your test result for HIV infection is positive, you will be referred to our designated clinic to receive treatment and follow-up visits.

Your responses to the questionnaire and your sample will be saved for future analysis. Your questionnaire will not have any information that identifies you. Only the investigators of this study will have access to your response and specimen.

Some of the questions I will ask are personal and include information about sexual behaviour. If you are embarrassed or uncomfortable about answering any questions, you may choose to skip them or you may end the interview at any point. We will try to provide you maximum privacy during questions and sampling procedures. Please be advised that although a professional will collect your blood sample, it may cause you local discomfort.

To minimize risk and discomfort that taking part in this interview may have for you, all information you give us will be anonymous (will not identify you personally). This means that you do not have to say your name or give any identifying information at any point. Only an assigned identification number will be attached to your responses and your specimens. You will also use this number to get your test result.

Your answers will not be shared with anyone outside this study and will remain confidential. The information from the study will be presented with all participants' results together; no individual responses will be shared. Your survey will be kept in a locked filing cabinet at the study office, and data will be stored in a secure computer to which only investigators will have access.

The study is entirely voluntary. By participating, you will help create treatment and prevention programmes for female sex workers [*You can use a more locally appropriate phrase instead, if one exists.*] In addition to getting your test result and counselling, you will also be invited to participate in an educational workshop on how to prevent HIV and other sexually transmitted diseases. You will also receive a prevention kit and some refreshments.

Besides your time, the study will not cost you anything. You will be reimbursed for your travel expenses and time with _____ [*include amount and kind of currency*].

If for any reason you would like to have another interviewer, please let us know. We will make arrangements so that your privacy will be fully protected.

Participation is voluntary. Even if you decide not to participate in the study at this point, you are still eligible to take part in the workshop and receive the prevention kit.

Appendix A.7.2

Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects [10]

World Medical Association Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects

Adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964, and amended by the:
29th WMA General Assembly, Tokyo, Japan, October 1975
35th WMA General Assembly, Venice, Italy, October 1983
41st WMA General Assembly, Hong Kong, September 1989
48th WMA General Assembly, Somerset West, Republic of South Africa, October 1996
52nd WMA General Assembly, Edinburgh, Scotland, October 2000
53rd WMA General Assembly, Washington 2002 (Note of Clarification on paragraph 29 added)
55th WMA General Assembly, Tokyo 2004 (Note of Clarification on Paragraph 30 added)
59th WMA General Assembly, Seoul, October 2008

A. Introduction

1. The World Medical Association (WMA) has developed the Declaration of Helsinki as a statement of ethical principles for medical research involving human subjects, including research on identifiable human material and data. The Declaration is intended to be read as a whole and each of its constituent paragraphs should not be applied without consideration of all other relevant paragraphs.

2. Although the Declaration is addressed primarily to physicians, the WMA encourages other participants in medical research involving human subjects to adopt these principles.

3. It is the duty of the physician to promote and safeguard the health of patients, including those who are involved in medical research. The physician's knowledge and conscience are dedicated to the fulfilment of this duty.

4. The Declaration of Geneva of the WMA binds the physician with the words, "The health of my patient will be my first consideration," and the International Code of Medical Ethics declares that, "A physician shall act in the patient's best interest when providing medical care."

5. Medical progress is based on research that ultimately must include studies involving human subjects. Populations that are under-represented in medical research should be provided appropriate access to participation in research.

6. In medical research involving human subjects, the well-being of the individual research subject must take precedence over all other interests.

7. The primary purpose of medical research involving human subjects is to understand the causes, development and effects of diseases and improve preventive, diagnostic and therapeutic interventions (methods, procedures and treatments). Even the best

current interventions must be evaluated continually through research for their safety, effectiveness, efficiency, accessibility and quality.

8. In medical practice and in medical research, most interventions involve risks and burdens.

9. Medical research is subject to ethical standards that promote respect for all human subjects and protect their health and rights. Some research populations are particularly vulnerable and need special protection. These include those who cannot give or refuse consent for themselves and those who may be vulnerable to coercion or undue influence.

10. Physicians should consider the ethical, legal and regulatory norms and standards for research involving human subjects in their own countries as well as applicable international norms and standards. No national or international ethical, legal or regulatory requirement should reduce or eliminate any of the protections for research subjects set forth in this Declaration.

B.   Basic principles for all medical research

1. It is the duty of physicians who participate in medical research to protect the life, health, dignity, integrity, right to self-determination, privacy, and confidentiality of personal information of research subjects.

2. Medical research involving human subjects must conform to generally accepted scientific principles, be based on a thorough knowledge of the scientific literature, other relevant sources of information, and adequate laboratory and, as appropriate, animal experimentation. The welfare of animals used for research must be respected.

3. Appropriate caution must be exercised in the conduct of medical research that may harm the environment.

4. The design and performance of each research study involving human subjects must be clearly described in a research protocol. The protocol should contain a statement of the ethical considerations involved and should indicate how the principles in this Declaration have been addressed. The protocol should include information regarding funding, sponsors, institutional affiliations, other potential conflicts of interest, incentives for subjects and provisions for treating and/or compensating subjects who are harmed as a consequence of participation in the research study. The protocol should describe arrangements for post-study access by study subjects to interventions identified as beneficial in the study or access to other appropriate care or benefits.

5. The research protocol must be submitted for consideration, comment, guidance and approval to a research ethics committee before the study begins. This committee must be independent of the researcher, the sponsor and any other undue influence. It must take into consideration the laws and regulations of the country or countries in which the research is to be performed as well as applicable international norms and standards but these must not be allowed to reduce or eliminate any of the protections for research subjects set forth in this Declaration. The committee must have the right to monitor ongoing studies. The researcher must provide monitoring information to the committee, especially information about any serious adverse events. No change to the protocol may be made without consideration and approval by the committee.

6. Medical research involving human subjects must be conducted only by individuals with the appropriate scientific training and qualifications. Research on patients or

healthy volunteers requires the supervision of a competent and appropriately qualified physician or other health care professional. The responsibility for the protection of research subjects must always rest with the physician or other health care professional and never the research subjects, even though they have given consent.

7. Medical research involving a disadvantaged or vulnerable population or community is only justified if the research is responsive to the health needs and priorities of this population or community and if there is a reasonable likelihood that this population or community stands to benefit from the results of the research.

8. Every medical research study involving human subjects must be preceded by careful assessment of predictable risks and burdens to the individuals and communities involved in the research in comparison with foreseeable benefits to them and to other individuals or communities affected by the condition under investigation.

9. Every clinical trial must be registered in a publicly accessible database before recruitment of the first subject.

10. Physicians may not participate in a research study involving human subjects unless they are confident that the risks involved have been adequately assessed and can be satisfactorily managed. Physicians must immediately stop a study when the risks are found to outweigh the potential benefits or when there is conclusive proof of positive and beneficial results.

11. Medical research involving human subjects may only be conducted if the importance of the objective outweighs the inherent risks and burdens to the research subjects.

12. Participation by competent individuals as subjects in medical research must be voluntary. Although it may be appropriate to consult family members or community leaders, no competent individual may be enrolled in a research study unless he or she freely agrees.

13. Every precaution must be taken to protect the privacy of research subjects and the confidentiality of their personal information and to minimize the impact of the study on their physical, mental and social integrity.

14. In medical research involving competent human subjects, each potential subject must be adequately informed of the aims, methods, sources of funding, any possible conflicts of interest, institutional affiliations of the researcher, the anticipated benefits and potential risks of the study and the discomfort it may entail, and any other relevant aspects of the study. The potential subject must be informed of the right to refuse to participate in the study or to withdraw consent to participate at any time without reprisal. Special attention should be given to the specific information needs of individual potential subjects as well as to the methods used to deliver the information. After ensuring that the potential subject has understood the information, the physician or another appropriately qualified individual must then seek the potential subject's freely-given informed consent, preferably in writing. If the consent cannot be expressed in writing, the non-written consent must be formally documented and witnessed.

15. For medical research using identifiable human material or data, physicians must normally seek consent for the collection, analysis, storage and/or reuse. There may be situations where consent would be impossible or impractical to obtain for such research or would pose a threat to the validity of the research. In such situations the research may be done only after consideration and approval of a research ethics committee.

16. When seeking informed consent for participation in a research study the physician should be particularly cautious if the potential subject is in a dependent relationship with the physician or may consent under duress. In such situations the informed consent should be sought by an appropriately qualified individual who is completely independent of this relationship.

17. For a potential research subject who is incompetent, the physician must seek informed consent from the legally authorized representative. These individuals must not be included in a research study that has no likelihood of benefit for them unless it is intended to promote the health of the population represented by the potential subject, the research cannot instead be performed with competent persons, and the research entails only minimal risk and minimal burden.

18. When a potential research subject who is deemed incompetent is able to give assent to decisions about participation in research, the physician must seek that assent in addition to the consent of the legally authorized representative. The potential subject's dissent should be respected.

19. Research involving subjects who are physically or mentally incapable of giving consent, for example, unconscious patients, may be done only if the physical or mental condition that prevents giving informed consent is a necessary characteristic of the research population. In such circumstances the physician should seek informed consent from the legally authorized representative. If no such representative is available and if the research cannot be delayed, the study may proceed without informed consent provided that the specific reasons for involving subjects with a condition that renders them unable to give informed consent have been stated in the research protocol and the study has been approved by a research ethics committee. Consent to remain in the research should be obtained as soon as possible from the subject or a legally authorized representative.

20. Authors, editors and publishers all have ethical obligations with regard to the publication of the results of research. Authors have a duty to make publicly available the results of their research on human subjects and are accountable for the completeness and accuracy of their reports. They should adhere to accepted guidelines for ethical reporting. Negative and inconclusive as well as positive results should be published or otherwise made publicly available. Sources of funding, institutional affiliations and conflicts of interest should be declared in the publication. Reports of research not in accordance with the principles of this Declaration should not be accepted for publication.

C.   Additional principles for medical research combined with medical care

1. The physician may combine medical research with medical care only to the extent that the research is justified by its potential preventive, diagnostic or therapeutic value and if the physician has good reason to believe that participation in the research study will not adversely affect the health of the patients who serve as research subjects.

2. The benefits, risks, burdens and effectiveness of a new intervention must be tested against those of the best current proven intervention, except in the following circumstances:

- The use of placebo, or no treatment, is acceptable in studies where no current proven intervention exists; or
- Where for compelling and scientifically sound methodological reasons the use of placebo is necessary to determine the efficacy or safety of an intervention and the patients who receive placebo or no treatment will not be subject to any risk of serious or irreversible harm. Extreme care must be taken to avoid abuse of this option.

3. At the conclusion of the study, patients entered into the study are entitled to be informed about the outcome of the study and to share any benefits that result from it, for example, access to interventions identified as beneficial in the study or to other appropriate care or benefits.

4. The physician must fully inform the patient which aspects of the care are related to the research. The refusal of a patient to participate in a study or the patient's decision to withdraw from the study must never interfere with the patient-physician relationship.

5. In the treatment of a patient, where proven interventions do not exist or have been ineffective, the physician, after seeking expert advice, with informed consent from the patient or a legally authorized representative, may use an unproven intervention if in the physician's judgement it offers hope of saving life, re-establishing health or alleviating suffering. Where possible, this intervention should be made the object of research, designed to evaluate its safety and efficacy. In all cases, new information should be recorded and, where appropriate, made publicly available.

22.10.2008

Final case studies

Final case study 1

IN THESE CASE studies, answer the questions for each section before moving on and reading the new section.

Part 1

You are in charge of the office of communicable disease surveillance in Menaland. The growing HIV epidemic worldwide and in your country has received substantial media attention recently. In your country, as in many other countries, people are advocating for youth education about HIV. A few HIV knowledge, attitude and practice (KAP) studies among high school and university students have shown that general knowledge regarding HIV prevention is weak, that almost half of all students who have had sex in the previous year had multiple partners, and that only 10% had used condoms.

Extramarital sexual relationships and homosexuality are highly stigmatized and people engaging in these practices may face legal prosecution in Menaland. There have been very limited surveillance activities among most-at-risk populations, such as injecting drug users, sex workers and men who have sex with men. Sporadic sentinel-based biological HIV surveillance activities have been conducted among prison populations that showed a wide range (2%–10%) of HIV prevalence.

The prevention office of the Ministry of Health has successfully lobbied for US$ 4 million for HIV prevention programmes in Menaland. There is a strong political movement to implement a national school-based programme to educate youth and develop a life-skills curriculum helping them to stay abstinent and avoid unsafe sex. The programme is going to be launched very soon, and your office receives a portion of the budget to collect baseline data to evaluate the programme.

Questions and discussion points:

1. Do you support or oppose the prevention programme proposed by the prevention office?

2. What specific programme(s) might you propose in response to these data instead?

3   What additional information do you need to make better programming decisions?

Part 2

The first round (baseline) of KAP studies among students aged 16–24 years old was consistent with what previous KAP studies had shown, including relatively weak HIV knowledge, misconceptions about HIV and negative attitudes toward HIV patients. However, closer inspection of the newly collected data has revealed that only 8% of all students had ever had sex at all, and only 4% had been sexually active in the past year. Also, 2% of all respondents had reported injecting drugs in the previous year.

Meanwhile, you receive a report from the directors of two nongovernmental organizations that offer shelter and basic medical treatment for 400 homeless people. The nongovernmental organizations are located near areas known for drug abuse and high crime rates in the capital city of Menaland. The report indicated that of the 400 clients of the shelter/medial treatment service, 90% are male, 45% are married, 60% have ever been in jail, 40% are drug injectors and 30% are HIV-infected. Also, based on classified data from the drug enforcement police, it is estimated that there are 50 000 injecting drug users in the capital city.

Questions and discussion points:

1. What might account for 30% HIV prevalence among the nongovernmental organizations' clients?

2. What does this information tell you about the current HIV/AIDS programming? Would you make any changes in your surveillance activities?

3. What would be your next step?

Part 3

You form a committee of experts on HIV to reassess in a systematic manner the current status of the HIV epidemic in Menaland. The committee pools all available data and concludes that the very high rates of HIV among injecting drug users might have been overlooked because health officials, the press and other agencies were looking only for evidence of unprotected sexual activity among youth.

Based on the committee's recommendation, you propose a comprehensive plan to assess the HIV epidemic among injecting drug users in Menaland. Harm reduction programmes are redesigned for people already injecting drugs, and injecting drug users are added as the top-priority group to the national surveillance system.

Questions and discussion points:

1. List the main components of your new comprehensive plan.

2. Would you consider including other most-at-risk populations, such as sex workers and men who have sex with men in your surveillance activities at this point? Why or why not?

3. What specific lessons can be drawn from this example?

Final case study 2

Part 1

Menaland is a traditional society with strong segregation of the sexes and engagement in homosexual activity has become a significant sexual outlet. Male-to-male sex is varied and often involves men who indulge in commercial sexual activities. This practice is tacitly accepted and involves various traditional categories of transgendered[1] and young male sex workers, many of whom are poor and unemployed.

Recently, a newspaper published a controversial article about the potential public health threat from transgendered sex workers selling sex. The article pointed out that in the evening they congregate around certain intersections in the wealthy areas of Menaland's capital city to solicit clients for sex. However, the newspaper neither discussed the size of the transgendered sex worker population nor to what extent their activities might be threatening the health of the general population.

Questions and discussion points:

1. As a public health official, would you take this newspaper article seriously? Why? Why not?

2. Would you respond to this situation? How?

Part 2

The Ministry of Health quickly orders a rapid assessment of the transgendered sex worker population of the Emrab district, the most populous district in the capital city, with 6 million inhabitants. The report estimates that there are 10 000 transgendered people who are now selling sex as their principal livelihood instead of using their traditional practices to get money, such as singing, dancing and begging. It is also noted that, although the prevalence of HIV is below 0.1% in the general population of the country, it is as high as 6% in the transgendered sex worker population.

---

[1] In Menaland, biologically male transgendered persons with ambiguous sexual identity and behaviour exist who are traditionally identified as a social group within society on the basis of these differences.

Questions and discussion points:

1. Would you immediately proceed to programming based on this report? Why? Why not?

2. Based on this new report, what further steps should be taken?

Part 3

The subnational AIDS Control Programme approves a detailed research study to identify the high-risk practices and HIV status of the transgendered sex worker population in the Emrab district. The district's population constitutes about 20% of the city's total population and 4% of the entire country. The district is divided into nine towns, with 150 local councils. The transgendered sex worker population lives in close quarters that are widely distributed throughout the district and are easily identified, so "natural" groupings with variable geographical distribution can be found in the population.

Question and discussion point:

1. Which sampling technique will be appropriate to recruit from this population?

Part 4

As part of a behavioural survey of sexual practices and knowledge about HIV/AIDS, 350 transgendered people were interviewed in the Emrab district. The survey found that almost 80% were engaged in other occupations along with sex work. Only 45% had adequate knowledge of HIV/AIDS. The level of education of respondents was significantly associated with their knowledge of HIV/AIDS. Fifty-one percent did not use condoms during their sexual encounters, 93% were sexually active with clients and friends, 34% had a history of an STI, 79% practised body piercing and 60% were willing to take an HIV test.

Questions and discussion points:

1. How will you use this data to respond to the situation?

2. What specific lessons can be drawn from this example?

Final case study 3

Part 1

You are the chief HIV surveillance officer of the Ministry of Health in Menaland. Recent reports indicate that HIV is spreading among injecting drug users in your city. You need to understand the nature of this epidemic, so you decide to do individually linked HIV biological and behavioural surveys among injecting drug users. Most injecting drug users are at higher risk of HIV acquisition and transmission due to sharing needles, but a formative assessment study shows that male-to-male sexual practices and selling sex for drugs are increasing HIV risk as well. Men who have sex with men are extremely stigmatized, and there has been no official report on this group so far.

Developing a protocol for this survey, you decide to use the standardized questionnaire designed by Family Health International (FHI) for behavioural surveillance among injecting drug users to save time and money. After translating the questionnaire, you realise that there are several questions that are less relevant to the culture of injecting drug users in Menaland. Your colleagues suggest that you should invite a small group of injecting drug users to pretest your questionnaire and find potential problems.

Questions and discussion points:

1. Based on your local common practice, discuss whether you must obtain an IRB approval for this survey.

2. If your answer is yes, discuss whether it would it be ethical to pretest the questionnaire before receiving IRB approval. (To refresh your memory about pretesting, you can refer to Unit 4.)

Part 2

Although according to your local policy this survey is considered a surveillance activity and is exempted from formal IRB review, your funding agency's policy requires you to get IRB approval in order to receive financial support. You send in your application, knowing that the IRB review always takes time.

The IRB sends back your application requesting that you attach a consent form to it. Given that the participants are a group of injecting drug users, many of whom are very poor,

and that you are considering giving them food stamps in return for their participation, you feel that the consent form may not be necessary. Your colleagues further comment that the questionnaire is already long enough and injecting drug users would not tolerate the additional time needed to obtain consent. To get around the issue, your team decides to put together a short consent form, and resends the application. Meanwhile, the team suggests that until you hear from the IRB, you can pilot test the questionnaire among a small group of injecting drug users.

Questions and discussion points:

1. Is it ever appropriate to conduct a survey with human subjects in which informed consent is not obtained?

2. Discuss whether it is ethical to pilot test the questionnaire while you are waiting to receive the IRB clearance.

3. Discuss ethical issues around giving monetary incentives to injecting drug users.

4. Given the nature of your survey, develop a verbal consent form.

Part 3

The IRB returns your consent form, this time asking for some revisions—specifically the inclusion of the two paragraphs below:

"Some of the questions I will ask are personal and include discussion of sexual behaviour and drug use. If you are embarrassed or uncomfortable about answering any questions, you may choose to skip them or you may end the interview at any point."

"To minimize any risk or discomfort that taking part in this interview may have for you, all information you give us will be anonymous (will not identify you personally). This means that you do not have to say your name or give any identifying information at any point. I will not collect any information that can later be used to identify or contact you. Only an assigned identification code will be attached to your responses."

Questions and discussion points:

1. For each paragraph, discuss which ethical principles they are referring to.

2. Have you already addressed these ethical principles in your consent form? If so, how?

3. Compare these suggested paragraphs with what you have already included in your consent form, and discuss why we should or should not add these paragraphs.

Module summary

Module summary

SURVEILLANCE IS THE SYSTEMATIC, regular and ongoing collection and use of data for public health action. Although they are often among the earliest data available in a country, one-time cross-sectional surveys should not be considered surveillance.

Behavioural surveillance involves regular, repeated, cross-sectional surveys collecting data that can be compared over time on HIV risk behaviours and other relevant issues. Biological surveillance also involves regular and repeated cross-sectional surveys but collects biological samples that are tested for HIV and other related illnesses, such as STIs and TB.

The uses of behavioural surveillance include the following:

- to provide an early warning of which groups and areas infection is likely to spread in and between
- to explain changes in HIV prevalence over time
- to provide information for developing prevention programmes
- to monitor and evaluate the effect of prevention programmes
- to reinforce the findings of biological surveillance
- to raise awareness of HIV among policy-makers.

Considerations when designing a behavioural surveillance system are:

- whom to include in surveillance
- where to access the surveillance populations
- how to link biological and behavioural surveillance data
- how to be certain that surveillance is appropriate for the context.

The two most difficult issues defining behavioural surveillance indicators are defining the behaviours themselves and defining the time period for which the indicator should refer.

The purpose of pre-surveillance assessment is to:

- identify surveillance populations and geographical areas
- determine feasibility of conducting surveillance in most-at-risk populations
- define most-at-risk populations and who can participate.

The methods that should be used for the pre-surveillance assessment are:

- assessing what is known about the national epidemic or subepidemic
- gathering information on risk behaviours and HIV levels in the general population
- conducting pre-surveillance assessment using qualitative assessment methods and mapping.

In surveys, measurement error can come from:

- questionnaires
- interviewers
- respondents.

Data collection methods that can be used in behavioural surveillance include:

- face-to-face interviews
- self-administered questionnaires
- computer- or audio-assisted methods.

General population survey instruments that are widely used, are considered to use solid sampling procedures and thorough statistical analysis, and are usually considered a reliable source of behavioural data include the following:

- Demographic and Health Survey (ICF MACRO)
- Multiple Indicator Cluster Survey (UNICEF)
- Behavioural Surveillance Survey (FHI).

The target population is the group that meets a survey's measurement objective, for example, all sex workers in a city. The survey population is the target population modified to take into account practical considerations, for example, all sex workers in a city over the age of 15 years, excluding those who are based at home, as they cannot be accessed.

There are several ways to gather information about people, through a complete count or through a sample survey. The accuracy of a sample survey depends on two major factors, including:

- the size of the sample
- the presence of systematic error in designing or collecting the sample.

Issues for sampling include:

- Consistency is required across survey rounds.
- General population surveys can rarely be used to access most-at-risk populations.
- Members of most-at-risk populations may be difficult to identify and access.
- Cluster sampling can be difficult when clusters are not stable.
- Cluster sampling is impossible if group members do not gather.

Behavioural surveillance among most-at-risk populations that are hard to reach includes:

- time-location sampling (TLS)
- respondent-driven sampling (RDS).

Data analysis is the process of summarizing, presenting and interpreting data. Behavioural surveillance analysis can either be:

- cross-sectional (one point in time)—analysing data from one surveillance round, or
- trend analysis (more than one point in time)—analysing data over several surveillance rounds.

Informed consent means informing the person enough about the nature of the surveillance for them to make a proper (informed) decision about whether to take part.

Potential harms caused by behavioural surveillance in most-at-risk populations include:

- physical (public attack, abuse, loss of health-care services)
- legal (arrest, prosecution)
- social (disclosure to family or workplace, discrimination, loss of employment, isolation).

Potential benefits from behavioural surveillance include:

- improving HIV prevention and care programmes
- raising public awareness of and sympathy for the burden of disease in the population
- reducing stigma and causing social change, especially around HIV infection
- providing information about the results to the community so the situation is better understood
- incentives.

References

References

1) MEASURE Evaluation. *The PLACE method: Priorities for Local AIDS Control Efforts. A manual for implementing the PLACE method.* Chapel Hill, NC, UNC Carolina Population Center, 2005 (http://www.cpc.unc.edu/measure/tools/hiv-aids/place, accessed 15 December 2009).

2) *Introduction to HIV, AIDS and sexually transmitted infection surveillance. Module 4: Introduction to respondent driven sampling.* Cairo, World Health Organization Regional Office for the Eastern Mediterranean, 2010.

3) Newman JC et al. The differential effects of face-to-face and computer interview modes. *American Journal of Public Health*, 2002, 92:294–7.

4) Hulley SB et al. *Designing clinical research: an epidemiological approach* (2nd edition). Philadelphia, PA, Lippincott Williams and Wilkins, 2001.

5) Sullivan LM. *Essentials of biostatistics in public health.* Sudbury, MA, Jones and Bartlett Publishers, 2008.

6) Magnani R et al. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS*, 2005, 19 (Suppl. 2):S67–S72.

7) Family Health International et al. *Behavioral surveillance surveys: guidelines for repeated behavioral surveys in populations at risk of HIV.* Arlington, VA, Family Health International, 2000.

8) *Reproductive health assessment toolkit for conflict affected women. Appendix B: Random numbers table and instructions.* Atlanta, GA, Centers for Disease Control and Prevention, 2007 (http://www.cdc.gov/reproductivehealth/refugee/pdf/appendixb.pdf, accessed 15 December 2009).

9) UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance. *Guidelines for the effective use of data from HIV surveillance systems.* Geneva, World Health Organization, 2004 (http://www.who.int/hiv/strategic/surveillance/en/useofdata.pdf, accessed 15 December 2009).

10) *World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects* (http://www.wma.net/en/30publications/10policies/b3/index.html, accessed 15 December 2009).

11) Council for International Organizations of Medical Sciences (cioms). *International ethical guidelines for biomedical research involving human subjects.* Geneva, cioms, 2002 (http://www.cioms.ch/frame_guidelines_nov_2002.htm, accessed 15 December 2009).

Annex I

Glossary

Glossary

ACASI	Acronym for "audio computerized assisted survey instruments".
Accuracy	Refers to how well the sample reflects (nearest to the truth) the study population.
Alliances	Partnerships created to assist with formative assessment. These partnerships differ based on the type of most-at-risk group being sampled, but usually include gatekeepers, governmental or nongovernmental organizations, influential members of the target group, advocates, and physicians and others who provide health care to the target group.
At-risk groups	Groups of people that are at increased risk for passing HIV on to others or for being infected by others.
Behavioural surveillance	Surveys of HIV-related behaviour that involve asking a sample of people about their risk behaviours, such as their sexual and drug-injecting behaviour.
Bias	A systematic error in the sample selection and the collection or interpretation of data.
Biological surveillance	Surveillance that involves regular and repeated cross-sectional surveys, but collects biological samples that are tested for HIV and other related illnesses, such as sexually transmitted diseases and TB.
Bridging populations	Bridging populations are populations at intermediate risk of exposure to HIV and provide links between the most-at-risk populations and the low-risk general population. Bridging populations are, for example, the sexual partners of injecting drug users, sex workers and men who have sex with men.
Chain referral sample	Any sampling method wherein participants refer other potential participants for inclusion in the sample. There are several types of chain referral sampling methods, most of which are non-probability samples. Examples of chain referrals include RDS, network sampling, random walk and snowball sampling.
Cluster	Any aggregate of the population of interest (for example, departments, villages or health facilities).
Cluster sampling	The population of interest is broken into groups or clusters and a sample of clusters is randomly selected.
Community-based surveys	Surveys that use samples that have been selected from non-clinical settings. They often include most-at-risk populations, such as sex workers or truck drivers, who are not included in clinic-based surveys. As with clinic-based surveys, the most common type of community-based survey is called "repeated cross-sectional community-based sentinel serosurveillance".

Community sites	Locations in the community, such as households or brothels.
Concentrated HIV epidemic	The epidemic state in which HIV has spread to a high level in a defined subpopulation but is not well established in the general population. HIV prevalence is consistently >5% in at least one defined subpopulation and is <1% in pregnant women in urban areas.
Confidence interval	The compound interval with a given probability (for example, 95%) that the true value of a variable such as mean, proportion or rate is contained within the limits. Also known as "confidence limits".
Continuous variable	Items that occur in a numerical order, such as height or age.
Convenience sampling	The selection of entities from a population based on accessibility and availability. Available participants may be people on the street, patients in a hospital or employees in an agency. This type of sampling does not generally represent the population of interest and is best used in the exploratory stage of research.
Cross-sectional survey	A survey that is conducted over a given period of time, such as during a single year, rather than over an extended period of time.
EMR/MENA	Acronym for the Eastern Mediterranean Region/Middle East and North Africa created explicitly for use in this module series. This area includes Afghanistan, Algeria, Bahrain, Djibouti, Egypt, Islamic Republic of Iran, Iraq, Jordan, Kuwait, Lebanon, Libyan Arab Jamahiriya, Morocco, Oman, Pakistan, occupied Palestinian territory, Qatar, Saudi Arabia, Somalia, Sudan, Syrian Arab Republic, Tunisia, United Arab Emirates and Yemen. EMR/MENA does not refer to an official UNAIDS or WHO region.
Equilibrium	In RDS, the point in the recruitment process where a variable is not expected to change by more than 2% with each successive wave.
Ethnographic mapping	Collecting information on the geographical location, temporal movement of and interactions among members of the study population.
Female sex workers	Females who engage in sex work, or the exchange of sex for money, which includes many practices and occurs in a variety of settings. These may include "direct" or "formal" sex workers, who are sometimes included in registries and often found in brothels, and "indirect" or "casual" sex workers, who do not engage in sex work full time and are unlikely to be included in registries.
Formative assessment (or research)	Research conducted before the study begins. Researchers use qualitative methods, such as focus groups, in-depth interviews, mapping or observations of the target population and the individuals who work with them to ensure that the research team sufficiently understands the community.
Gatekeepers	Persons who can provide access to a most-at-risk population. Examples are a brothel owner who can provide access to female sex workers, or a prison warden who can provide access to prisoners.
General population	The general population encompasses all living people living in a defined community, e.g. a country, a province or a city. If surveillance data on the general population exclude certain subpopulations, e.g. specific age-groups of the population, foreign workers, refugees or others, this should be stated clearly. The general population is at relatively low risk of HIV exposure in low-level and concentrated HIV epidemics.
General population surveillance	Surveillance that measures HIV risk behaviours in a sample of people selected to represent the people living in a region or nation. The surveillance can be restricted to certain ages (for example, young people aged 15–24) or genders.

Generalizability	The results from the sample are the same as the results we would have obtained had we tested every person in the study population (that is, the results from the sample are generalizable to the study population).
Generalized HIV epidemic	The epidemic state in which HIV is firmly established in the general population; HIV prevalence is consistently >1% in pregnant women.
Grey literature	Material that is not published in easily accessible journals or databases. Besides programme evaluations, government surveillance reports and programme planning documents, it includes the abstracts of research presented at conferences, and unpublished theses and dissertations.
High-risk behaviours	Behaviours that increase the risk that a person will contract a disease.
Indicators	Specific data that are gathered to measure how well a prevention or treatment programme is doing. They define an aspect of behaviour that is key to the spread of HIV. Indicators provide a way to track changes in behaviours over time and provide a way to compare levels of risk behaviours between different population groups.
Informed consent	The permission granted by a patient or a participant in a research study after he or she has received comprehensive information about a research study or medical procedure. Informed consent protects the person's freedom of choice and respects his or her autonomy with regard to decisions affecting his or her body and health.
Injecting drug users	Injecting drug users are persons who use needles or syringes to inject drugs. Injecting drug use is considered a high-risk behaviour.
Institutional review board (IRB)	The committee designated to approve, monitor and review biomedical and behavioural research involving humans with the aim of protecting the rights and welfare of research participants. Also known as an ethics committee.
Key informants	Members of the target group who can often become informal assistants.
Key populations	Key populations include most-at-risk, bridging and vulnerable populations, who, while being important to the dynamics of HIV transmission in a setting, are essential partners for an effective response to the epidemic, i.e. they are key to the epidemic and key to the response. Key populations will vary depending on the epidemic and community/country context.
Low-level epidemic	The epidemic state in which HIV has not spread to significant levels in any subpopulation, although HIV infection may have existed for many years; HIV prevalence has not consistently exceeded 5% in any defined subpopulation. This state suggests that networks of risk are rather diffuse or that the virus has only been recently introduced.
Male sex workers	Males who engage in sex work, or the exchange of sex for money, which includes many practices and occurs in a variety of settings.
Margin of error	An estimation of the extent to which a survey's reported percentages would vary if the same survey were made multiple times.
Most-at-risk population (MARP)	These are populations that experience the highest probability of being exposed to HIV and include injecting drug users, men who have sex with men (including male sex workers) and female sex workers.
Men who have sex with men (MSM)	Men who have sex with men are considered at increased risk of HIV infection if they engage in high risk behaviours such as unprotected anal intercourse. For the purposes of this module, we also consider male sex workers, transvestites and transgendered persons in the men who have sex with men category.

Multivariate analysis	One of the main types of analysis conducted in behavioural surveillance that is performed to look at the influence of at least two variables on another variable since relationships between variables are often complex and interwoven. Multivariate techniques can pinpoint the individual effects of several explanatory variables on an outcome variable, which may be related to each other.
Network	This sampling method may be used for groups whose members are socially linked. Ego-centred network sampling is based on random, representative or any other form of quota sampling. Full relational network sampling begins with identification of individuals ("seeds") who act as entry points to the network.
Non-probability sampling	When the sampling units are selected through a non-randomized process; therefore, the probability of selecting any sampling unit is not known.
Operational definitions of target populations	Definitions that are operationally useful for sampling and fieldwork purposes. For example, a definition that clearly identifies what constitutes a sex worker, in terms of duration of selling sex, form of payment, type of venue where they work, etc.
Point prevalence	Refers to prevalence at a single point in time. Also known as "point incidence".
Population-based serosurvey	A type of serosurvey that uses a probability sample of a population defined by geographical boundaries, such as villages or provinces, in order to obtain a direct measure of HIV prevalence in the general population.
Population subgroup	A group within a population that share certain characteristics or behaviours.
Precision	Refers to how well the results can be reproduced each time the survey is conducted.
Pre-surveillance assessment	Describes a set of activities that occur prior to beginning formal HIV and behavioural surveillance in most-at-risk populations. These activities include developing detailed plans and reviewing and collecting information that will help in planning and designing surveillance activities.
Prevalence	The proportion of persons in a given population with a disease or condition at a given point in time; a specific group infected. Prevalence is a direct measurement of the burden of disease in a population.
Primary sampling units	A sampling frame of a larger unit. When it is difficult or impossible to make a list/sampling frame of each individual in the target population, we can develop a sampling frame of some larger unit; that is, clusters or primary sampling units. We then sample in stages by first sampling clusters and then sampling people within the clusters.
Priorities for Local AIDS Control Efforts (PLACE)	A new, rapid assessment tool used to identify high transmission areas, which formalizes the collection of information on high transmission areas. PLACE uses key informants to identify sites where people meet new sex partners, then interviews people at the site in order to characterize the site in each area and map sites, and, finally, interviews individuals socializing at the site to describe the characteristics of the people at the site.
Probability proportional to size sampling	A class of unequal probability sampling in which the probability of a unit being sampled is proportional to the level of some known variable.
Probability sampling	All sampling units in the study population have a known, non-zero, probability of being selected in the sample, usually through a randomized process.

Protocol	The detailed plan for conducting a research study or other activities in which specific steps are required, including surveillance activities.
Qualitative research	Research that focuses on the characteristics or quality of things, rather than the quantity. The sample is much less used in qualitative research than in quantitative research.
Quantitative research	Research that focuses on the quantity of things, rather than the quality. Quantitative research has powerful tools for the analysis of numbers, but the things counted are often qualitative categories or definitions.
Random error	Also called non-systematic error. This is the type of error that results from chance and leads to imprecise results.
Rapid assessment and response (RAR)	A method that is used to assess the nature and extent of a public health problem and to suggest ways to address the problem. RAR is not designed as a surveillance tool, but as a way to assess a situation quickly, and bring in resources to address it.
RDS	See **respondent driven sampling**.
RDSAT	Acronym for "respondent driven sampling analysis tool" (a freeware software package for analysing RDS samples).
Representativeness	The degree to which the sample truly reflects the study population, i.e. whether it is representative of the study population.
Respondent driven sampling (RDS)	A sampling technique that does not require a sampling frame. It is an adaptation of a non-probability sampling method (snowball sampling) and is based on the assumption that members of the subpopulation themselves can most efficiently identify and encourage the participation in surveillance of other subgroup members. RDS starts with initial contacts or "seeds" who are surveyed and then become recruiters. Each of these recruiters is given coupons to use to invite up to three eligible people that he/she knows in the most-at-risk population to be interviewed. The new recruits bring their coupon to a central place where they are interviewed. The recruits then become recruiters. This occurs for five to six waves. Both the recruits and the recruiters are given incentives to encourage participation.
Sample	A selected subset of a population. There are specific types of samples used in surveillance and epidemiology such as convenience, systematic, population-based and random.
Sample size	The number of subjects to be used in a given study.
Sample frame	A list of units from which a sample may be selected. A sample frame is a fundamental part of probability sampling.
Sampling bias	Also called selection bias. This refers to errors in sampling that decrease accuracy and lead to incorrect estimates. We also use the term "biased samples" to mean that errors were made in choosing the people in the sample.
Sampling interval	Calculated by dividing the total target population by the sample size.
Sampling variation	Difference between the estimate you measure in a sample and the true value of the variable in the study population.
Selection bias	A systematic error in the process of respondent selection for a study or survey.
Sentinel sites	Sites at which sentinel surveillance activities take place, including clinics attended by individuals who may or may not be representative of the general population but are likely to represent groups initially infected or at higher risk for infection than the general population.

Sentinel surveillance	A surveillance system in which a pre-arranged sample of reporting sources at "watch post" or "sentinel" sites agree to report all cases of one or more notifiable conditions. Often designed to provide an early indication of changes in the level of disease. Depending on the nature of the population surveyed, these data may be representative of the general population, or they may simply give more detailed information about the populations tested.
Sexually transmitted infection (STI)	Diseases that are spread by the transfer of organisms from person to person during sexual contact.
Sex workers	Persons who engage in sex work or the exchange of sex for money, which includes many practices and occurs in a variety of settings. These may include "direct" or "formal" sex workers, who are sometimes included in registries and often found in brothels, and "indirect" or "casual" sex workers, who do not engage in sex work full time and are unlikely to be included in registries. The term "sex worker" can be used to refer to female, male and transgendered sex workers.
Simple random sampling (SRS)	Sampling where everyone has an equal chance of being randomly selected (a non-zero probability) and we know what that chance is.
Snowball sampling	Relies on informants to identify other relevant study participants in a chain referral pattern. Informants (seeds) who meet inclusion criteria are identified. The sampling design is based on chain referral and relies on the seed(s) to identify other relevant subjects for study inclusion. Those other subjects may identify other relevant subjects for inclusion. Snowball sampling is useful for studying populations that are difficult to identify or access. Representativeness is limited.
Stakeholders (or stakeholder's group)	Those with an interest in the results of surveillance activities. Includes public health practitioners, health-care providers, data providers and users, representatives of affected communities, governments at the district, province and national levels, and members of professional and private non-profit and donor organizations.
Standard error	Estimate of precision in probability sampling that can be used to construct a range of values within which the true population measure is likely to fall. We usually want to be 95% sure that the true population measure lies in our range.
STI	See **sexually transmitted infection**.
Stigma	A mark of disgrace or shame. For example, in some societies, being infected with HIV causes a person to be stigmatized.
Stratum	A subgroup in stratified sampling.
Stratification	The classification of a survey population into subgroups or strata on the basis of selected characteristics.
Stratified sampling	Stratified sampling is generally used to obtain a representative sample when the population is heterogeneous, or dissimilar, where certain homogeneous, or similar, subpopulations can be isolated (strata). A stratified sample is obtained by taking samples from each stratum or subgroup of a population.
Subpopulation	See **population subgroup**.
Surveillance	The systematic collection, analysis, interpretation and dissemination of health data on an ongoing basis, to gain knowledge of the pattern of disease occurrence and potential in a community, in order to control and prevent disease in the community.

Survey population	The target population modified to take into account practical considerations (for example, all commercial sex workers in a city over the age of 15, excluding those who are based at home, as they cannot be accessed).
Systematic sampling	A sampling method that consists of randomly selecting the initial patient who meets the inclusion criteria and then selecting every "*nth*" (for example, third or fifth) eligible patient thereafter until the predetermined sample size is reached or the survey period is over.
Target population	The group that meets a survey's measurement objective (for example, all commercial sex workers in a city).
Time-location sampling (TLS)	Similar to conventional cluster sampling, but gets around the problem of clusters that are not stable (that is, clusters where the number and type of people vary by, for example, time of day). Time-location sampling allows the same site to be included in the sample frame more than once (for example, at different times of the day or different days of the week).
Trend	A long-term movement or change in frequency, usually upwards or downwards.
UAT	See **unlinked anonymous testing**.
UNGASS	Acronym for United Nations General Assembly Special Session on HIV/AIDS.
Univariate analysis	The most basic, yet often the most important, type of behavioural surveillance analysis, because it shows the distribution of each variable. Most of the indicators defined for behavioural surveillance purposes are calculated through univariate analysis. They would include variables like the proportion of young men who have had sex with more than one partner during a given time period. When trends are analysed, statistical techniques are used to calculate how likely it is that changes in the proportions could have occurred by chance, or whether observed changes are likely to reflect real changes.
Unlinked anonymous testing (UAT)	Testing that occurs when a sample of blood originally collected for other purposes is tested for HIV after being anonymized. The person whose blood is taken does not know that his/her blood will be tested for HIV. All information that could identify the person is removed from the sample so that the results of the test cannot be linked back to them.
Voluntary counselling and testing (VCT)	A programme that provides both counselling and testing services to communities, allowing persons who are tested to obtain emotional and medical support before and after their HIV tests.
Vulnerable population	Vulnerable populations are found in all settings. Their risk of acquiring HIV depends on the prevalence of HIV in their networks and their ability to adopt safer behaviours and access prevention and treatment commodities and services. For that reason, populations such as people affected by humanitarian emergencies and migration, young people, women, prisoners and people living with disabilities may be at high risk in some places but not in others. Once these vulnerable populations adopt high-risk practices they become part of most-at-risk or bridging populations.
WHO EMRO	Acronym for the World Health Organization, Regional Office for the Eastern Mediterranean. The Eastern Mediterranean Region includes Afghanistan, Bahrain, Djibouti, Egypt, Islamic Republic of Iran, Iraq, Jordan, Kuwait, Lebanon, Libyan Arab Jamahiriya, Morocco, Oman, Pakistan, occupied Palestinian territory, Qatar, Saudi Arabia, Somalia, Sudan, Syrian Arab Republic, Tunisia, United Arab Emirates and Yemen.

Annex 2

Useful links

Useful links

The Body

An AIDS and HIV information resource based in New York City, USA. Provides information on various questions related to HIV/AIDS. Available at: http://www.thebody.com

Cochrane HIV/AIDS Group

An affiliate of the International AIDS Society and the University of California, San Francisco (UCSF) CSF AIDS Research Institute, the Cochrane Collaborative Review Group on HIV Infection and AIDS is an international network of health-care professionals, researchers and consumers working to prepare, maintain and disseminate systematic reviews on the prevention and treatment of HIV infection and AIDS. Available at: http://www.igh.org/cochrane

Council for International Organizations of Medical Sciences (CIOMS)

The Council for International Organizations of Medical Sciences (CIOMS) has prepared the *International ethical guidelines for biomedical research involving human subjects*, in collaboration with the WHO. Available at: http://www.cioms.ch/frame_guidelines_nov_2002.htm

Family Health International (FHI)

FHI has pioneered ways to curtail the spread of HIV/AIDS. Many of the HIV prevention "best practices" in use today have emerged from FHI's work in more than 60 countries. Available at: http://www.fhi.org/en/hivaids (English) or http://www.fhi.org/ar/fhiag.html (Arabic).

The Global Fund to Fight AIDS, Tuberculosis and Malaria

The Global Fund was created to finance a dramatic turnaround in the fight against AIDS, TB and malaria. These three diseases kill more than six million people a year. This massive scaling-up of resources is already supporting aggressive interventions against all three. Available at: http://www.theglobalfund.org

HIV INSITE

HIV InSite is developed by the Center for HIV Information (CHI) at UCSF. HIV InSite's mission is to be a source for comprehensive, in-depth HIV/AIDS information and knowledge. Available at: http://www.hivinsite.ucsf.edu

HIV/AIDS Rapid Assessment Guide

FHI has prepared an HIV/AIDS rapid assessment guideline. Available at: http://www.fhi.org/en/hivaids/pub/guide/rapidassessmentguide/index.htm

HIV/AIDS Survey Indicators Database

The HIV/AIDS Survey Indicators Database is overseen by a technical advisory committee that includes representatives from USAID, UNICEF, CDC, UNAIDS, WHO, US Census Bureau, FHI, MEASURE Evaluation, The Synergy Project and MEASURE DHS+ (the implementing organization). USAID is currently the primary funder for the initiative, with UNAIDS and UNICEF providing additional support. There are 180 surveys available in the database. Available at: http://www.measuredhs.com

Multiple Indicator Cluster Survey (MICS), UNICEF

The MICS is a household survey programme developed by UNICEF to assist countries in filling data gaps for monitoring the situation of children and women. It is capable of producing statistically sound, internationally comparable estimates of these indicators. Available at: http://www.childinfo.org

Respondent-Driven Sampling (Cornell)

Defines RDS and provides information on minimum data requirements, sampling references, intervention references and downloads. Available at: http://www.respondentdrivensampling.org

UNAIDS (Joint United Nations Programme on HIV/AIDS)

As the main advocate for global action on HIV/AIDS, UNAIDS leads, strengthens and supports an expanded response aimed at preventing the transmission of HIV, providing care and support, reducing the vulnerability of individuals and communities to HIV/AIDS and alleviating the impact of the epidemic. Available at: http://www.unaids.org

UNAIDS epidemiological information on HIV/AIDS

Available at: http://www.unaids.org/en/resources/epidemiology.asp

UNAIDS surveillance information on HIV/AIDS

Available at: http://www.unaids.org/en/in+focus/topic+areas/surveillance+and+reporting.asp

United Nations Children's Fund (UNICEF)

UNICEF is one of the United Nations' key agencies in the fight against HIV/AIDS, mobilizing financial resources and helping persuade governments to put HIV/AIDS at

the top of their agendas and to treat the epidemic as a national emergency. UNICEF is working in 160 countries around the world to combat the epidemic. Available at: http:// www.unicef.org/aids

United Nations General Assembly Special Session (UNGASS)

This site is dedicated to tracking compliance with the United Nations General Assembly Special Session on HIV/AIDS (UNGASS), which in 2001 concluded with a declaration of commitment signed by 189 Member States to take actions to reduce the spread and impact of HIV/AIDS. As part of this effort, UNAIDS reports on progress toward achieving this goal every two years. To measure progress, UNAIDS developed a set of 25 indicators called the UNGASS indicators. Available at: http://www.ua2010.org/index.php/en/ungass

United Nations Office on Drugs and Crime (UNODC)

UNODC is a global leader in the fight against illicit drugs and international crime. UNODC is involved in HIV/AIDS programming in regions, such as the Middle East and North Africa, where injecting drug use is known to drive the HIV/AIDS epidemic. Available at: http://www.unodc.org

U.S. Centers for Disease Control and Prevention (CDC)

CDC serves as the national focus for developing and applying disease prevention and control, environmental health, and health promotion and education activities designed to improve the health of the people of the United States. Available at: http://www.cdc.gov

The CDC Global AIDS Program (GAP) surveillance team has developed an interactive sampling selection tool for use in surveillance study sampling design. Proper sampling design is critical to the success of a study. The tool is availableat: http://globalhealthsciences.ucsf.edu/pphg/surveillance/sampling_tool.html

U.S. National Institutes of Health (NIH)

The National Institutes of Health is the federal focal point for medical research in the United States. The NIH, comprising 27 separate institutes and centres, is one of eight health agencies of the Public Health Service, which, in turn, is part of the U.S. Department of Health and Human Services. Simply described, the goal of NIH research is to acquire new knowledge to help prevent, detect, diagnose and treat disease and disability. Available at: http://www.nih.gov

World Bank, Global HIV/AIDS Program

The Global HIV/AIDS Program was created in 2002 to support the World Bank's efforts to address the HIV/AIDS pandemic from a cross-sectoral perspective. The programme offers global learning and knowledge sharing on approaches and best practices to addressing HIV/AIDS. Available at: http://www1.worldbank.org/hiv_aids/globalprogram.asp

World Health Organization (WHO)

The WHO is the United Nations specialized agency for health. WHO's objective, as set out

in its Constitution, is the attainment by all peoples of the highest possible level of health. WHO is governed by 192 Member States through the World Health Assembly. Available at: http://www.who.int

WHO Department of HIV/AIDS

The HIV/AIDS Department coordinates a strategic, organization-wide response to the HIV/AIDS epidemic and enables WHO to provide enhanced technical support in HIV/AIDS to countries and regional offices. Available at: http://www.who.int/hiv/en

WHO Regional Office for the Eastern Mediterranean

The Regional Office coordinates WHO activities for the Eastern Mediterranean Region. The Region includes: Afghanistan, Bahrain, Djibouti, Egypt, Islamic Republic of Iran, Iraq, Jordan, Kuwait, Lebanon, Libyan Arab Jamahiriya, Morocco, Oman, Pakistan, occupied Palestinian territory, Qatar, Saudi Arabia, Somalia, Sudan, Syrian Arab Republic, Tunisia, United Arab Emirates and Yemen. The web site provides a variety of information specific to the Eastern Mediterranean Region and links to partner and regional government websites. Available at: http://www.emro.who.int

WHO Test Kit Evaluation Programme

The WHO test kit evaluation programme aims to provide Member States, UN agencies and other partners with technical information and advice on the quality of currently available test kits and technologies. Additional information is available at: http://who.int/diagnostics_laboratory/evaluations/en/

Annex 3

Answers to warm-up questions and case studies

Answers to warm-up questions and case studies

ANSWERS ARE PROVIDED in italics for each unit's warm-up questions and case studies.

Answers to the questions within the unit are not included. Unit questions are designed to stimulate small group discussion among participants in the workshop or class.

Unit 1 answers

Warm-up questions

1. Which of the following can be a use of behavioural surveillance?
   a. To explain changes in HIV prevalence over time
   b. To provide information for prevention programmes
   c. To raise the awareness of HIV among policy-makers
   d. To provide an early warning of which groups and areas infection is likely to spread in and between
   e. All of the above

2. _____ surveillance involves regular, repeated, cross-sectional surveys to collect data on HIV risk behaviours and other relevant issues which can be compared over time.
   a. Behavioural
   b. Biological

3. True or false? Behavioural surveillance can be used to evaluate the effect of interventions or parts of programmes, if questions about them are included in behavioural surveys. Circle your answer below.
   True. In order to evaluate a specific component of the intervention programme, we need to include specific questions in the core behavioural surveillance surveys. We should also make sure that we include the population who were supposed to receive the intervention in our survey. Without adaptation, surveillance only provides evidence for the effect of HIV programmes as a whole and not of specific interventions or specific programme elements.

4.   True or false? In a generalized epidemic, everyone is at equal risk for infection.

False. Even in a generalized epidemic, not everyone in the population is at equal risk for HIV, nor has an equal role in the spread or continuance of the epidemic.

5.   _____ sites are facilities such as STI clinics, antenatal care clinics, blood donation centres, drug treatment programmes, prisons and needle-exchange programmes.

a. Sentinel

b. Community

6.   Which of the following is the definition of linking behavioural and biological data?

a. Collecting HIV, STI and behavioural data from the same people at the same time

b. Collecting HIV, STI and behavioural data from the same source population at different times

c. Analysing HIV, STI and behavioural data from a similar source population, using whatever data are available

d. Reporting behavioural and biological surveillance together

e. All of the above

7.   Collecting data at the _____ level provides more detailed information, but requires larger sample sizes and more time and money.

a. National

b. Subnational

8.   True or false? Regardless of the state of the HIV epidemic (low-level, concentrated or generalized), behavioural surveillance must be conducted among the general population every year.

False. Behavioural surveillance among the general population is recommended only in a generalized epidemic and in concentrated epidemics in areas where HIV prevalence is very high in most-at-risk populations.

Apply what you have learned/case study

1.   Name an HIV most-at-risk population in your country. Describe the group and explain why you think it is at high risk for HIV infection.

Local information must be used to determine what constitutes a most-at-risk population based on a population's importance or potential importance in the epidemic. For example, most people view sex workers as a most-at-risk population, but sex workers are not actually a single population; they can be subdivided into many subpopulations. For instance, individual women on the street who solicit sex do not always belong to or associate with sex workers who work in brothels.

2.   Look at the diagram on the next page which shows how the HIV epidemic works in a hypothetical community. The diagram helps us to understand what behaviours drive the HIV epidemic, the most-at-risk populations and their links to the general population. The ovals show the different population groups, and overlapped ovals mean some may belong to more than one group. The arrows show the links between the populations,

the thickness of arrows shows the strength of the links and dashed arrows represent potential links.

Most-at-risk populations are those who engage in risk behaviour involving sex or injecting drugs; in this case injecting drug users who share needles or clients of sex workers who are infected with STIs. Being infected with STIs indicates that the individuals have had unprotected sexual contact. Individuals from most-at-risk populations who engage in high-risk behaviour and have partners without high-risk behaviours are considered to constitute bridging populations. In this example, the regular partners of injecting drug users and the clients of sex workers, and their children, are considered low-risk populations.

Unit 2 answers

Warm-up questions

1. What are two characteristics of a good indicator?

    A good indicator measures something of relevance to the topic, measures the item of interest accurately, is easy to interpret and is defined in clear terms. It can be compared across different population groups and across time and is feasible to collect in terms of effort/cost.

2. True or false? Behavioural surveillance indicators should measure aspects of behaviour that are essential to the spread of HIV. Circle your answer below.

    True.

3. When should behavioural surveillance indicators be selected during behavioural surveillance?

    a. during planning
    b. during analysis

4. True or false? The two most difficult issues in defining indicators for behavioural surveillance are defining the behaviours and defining the time period to which the indicator should refer.

    True.

5. True or false? Behavioural surveillance indicators do not need to be consistent over time.

    False

6. Fill in the blank spaces using one of the answers given below. An indicator's reference period should be determined based on _____ of a risk behaviour that we want to measure. For instance, to accurately measure a very _____ risk behaviour, a _____ reference period should be used.

    a. frequency, frequent, shorter
    b. frequency, frequent, wider
    c. importance, important, shorter
    d. importance, important, wider

Apply what you have learned/case study

1. Compare the indicators listed below and discuss the advantages and disadvantages of each one. In general, which one is more useful?

   a. Proportion of men who have sex with men who had anal sex without condoms with other men in their last sexual contact

   b. Proportion of men who have reported unprotected sex with other men in the last month.

   The first indicator is more specific about high-risk sexual contact among men who have sex with men while the second one is not risk-group specific and can include bisexual or heterosexual men. Also by explicitly indicating "anal sex", we remove any confusion about type of sexual contact. "Unprotected sex" is sometimes ambiguous and clearly stating using "condoms" makes it more straightforward. Finally, the time reference for the first indicator makes it easier for respondents to recall the sexual act, while for the second one, the respondent has to recall all of his sexual contacts during one month.

   The negative aspect of the first indicator is that in some cultures it is not acceptable to say words that are sexually charged. For example "anal sex" as an act could be stigmatized and using this wording to describe it could make people feel embarrassed. Also, by limiting the time reference to the last sexual contact, we may get a distorted picture of the true frequency of this sexual behaviour in larger populations if by chance the last sexual act for people was different from their regular behaviour. Overall, the first indicator is more useful because the data collected will be more interpretable.

2. Discuss the differences between substance users and illicit drug users.

   "Substance" can include prescription drugs, such as diazepam, methadone, tobacco and sometimes alcohol. (Depending on the setting, sometimes alcohol is considered an illicit substance.) The term "illicit drugs" generally refers to drugs that are illegal to transport, sell or use, such as heroin, opium and cocaine.

Unit 3 answers

Warm-up questions

1. Which of the following is an example of a question you should answer to select appropriate geographical areas during a formative assessment?

   a. What are the regional differences in terms of HIV transmission and risk behaviour?

   b. Where are the interventions located?

   c. How much money and staff are available for surveillance?

   d. All of the above.

2. True or false? For some methods of formative assessment, you need to review existing surveillance data and published and unpublished literature. Also, you should talk to people who are knowledgeable about the epidemic. Circle your answer below.

   True.

3. Methods used in *quantitative* research include large sample size, random samples and shorter interviews.

Methods used in _qualitative_ research include unstructured questionnaires, lengthy interviews and fewer well-trained interviewers.

4. Suppose you want to determine areas of a city where street-based sex occurs. Which of the following is not an appropriate method to use during a formative assessment?

    a. literature/data review

    b. in-depth qualitative interviews

    c. general population survey

    d. all are appropriate

5. List three advantages to building alliances.

    a. increases resources

    b. results in more creative solutions

    c. spreads the workload

    d. creates a safer environment

    e. raises community awareness

Apply what you have learned/case study

You are the HIV surveillance officer in Menaland, a country in EMR/MENA that has a concentrated epidemic based on a consistent HIV prevalence of 5% in STI clinic patients. You decide that you would like to initiate surveillance among sex workers in the capital city, but first you must conduct a formative assessment to define and understand the characteristics of sex workers and ways of accessing the population.

To date, the National AIDS Programme has not implemented any surveillance or prevention activities in this population, so very little data exists. You decide to collaborate with community stakeholders who work closely with this population to aid in this task.

    a. Who would these stakeholders be?

    Sex workers, ex-sex workers, STI clinic staff, local nongovernmental organizations, HIV/AIDS providers, pimps.

    b. How would you identify and contact them?

    Talking to people who are knowledgeable about sex workers, who you can talk with easily, who understand the information you need and who are willing to give you information.

    c. How would you establish alliances?

    Establish a core working group, develop a general goal, identify and enlist interested and effective members, identify mutual benefits for alliance members, and develop a strategy and roles and responsibilities for each person in the group.

    d. Once you have established alliances, outline the steps that you would take to do a formative assessment before starting surveillance activities.

    Conduct key informant interviews, focus-group discussions, expert interviews and mapping.

    e. Describe the benefits of using PLACE for formative assessment among sex workers in Menaland.

Although it was designed as an HIV prevention tool, PLACE can also be useful to find where high-risk sex occurs and what most-at-risk groups may be involved. The PLACE method describes how to collect the information needed for formative assessment, such as key informant interviews, interviews with knowledgeable people at the sites, interviews with people socializing at the sites and maps of site locations. Because little is known about sex workers in Menaland and because no behavioural surveillance has been conducted, PLACE would provide valid guidance in the conduct of formative assessment.

Unit 4 answers

Warm-up questions

1. What is measurement error?

    When the data collected do not accurately measure the characteristics of interest. This affects the validity of the data.

2. In HIV behavioural survey studies, which of the following is a source of measurement error?

    a. interviewer

    b. respondent

    c. questionnaire

    d. all of the above

3. True or false? You can always control for bias during data analysis. Circle your answer below.

    False. In fact, in most cases it is impossible to control for bias unless we know the source and we have all data necessary to control for some types of bias, such as confounders. A confounder is a variable which is related to one or more of the variables in the study so it may mask an actual association or falsely demonstrate an apparent association between the study variables where no real association between them exists.

4. True or false? If we use face-to-face interviews to survey most at-risk populations, we can collect more accurate data compared with other survey methods, such as self-administrated paper-based or computer-assisted interviews.

    False. Each data collection method has advantages and disadvantages. For certain questions about high-risk behaviours that are socially stigmatized, self-administrated or computer-assisted interviews may help respondents to provide more-valid answers without feeling embarrassed compared to face-to-face interviews. However, there is no conclusive evidence that any of these methods can be used to collect more accurate data compared to the others, and this can only be evaluated on a case-by-case basis. We can never be 100% sure which answer is valid and which is not when we are dealing with self-reported data.

5. Which of the following is an advantage of using the self-administered paper-based questionnaire?

    a. inexpensive to administer

    b. no literacy requirement

    c. date entry step eliminated

6. Which of the following are advantages of adapting survey questions from surveys that have already been used successfully?
   a. builds on current best practice of how questions can be expressed
   b. saves time and money
   c. makes certain your data are consistent with other available data sources
   d. all of the above

7. What is a solution to the problem of interviewer safety when working with most-at-risk populations?
   a. selecting interviewers who know the area
   b. selecting interviewers who know the population
   c. arranging for interviewers to work in pairs
   d. All of the above

8. True or false? When you are having difficulty in getting members from most-at-risk populations to show up for interviews, one solution is to use incentives.
   True.

Apply what you have learned/case study

1. In a study of 1417 participants of syringe exchange programmes, two modes of interview were used: face-to-face and audio-computer assisted self-interviews (ACASI). The survey contained questions about several "stigmatized behaviours". After completion of the survey, researchers compared the frequency of responses to questions related to the stigmatized behaviours based on the interview methods. They found that participants who were in the ACASI group had reported more frequent stigmatized behaviours than participants in the face-to-face group [3].
   a. Discuss potential reason(s) that may explain this difference.

   Since participants were asked culturally-sensitive questions, those who participated in ACASI could freely report about their behaviours without feeling that they were being judged by an interviewer; therefore, they were more likely to express themselves without self-censoring and hesitation. Another explanation is that this difference is real and people who participated in ACASI in fact had higher rates of certain behaviours.

   b. Imagine that you want to do a survey among injecting drug users who go to drug treatment centres in your country. What would be the advantages or disadvantages of using each of the methods?

   Face-to-face interviews are preferable when we expect a substantial number of participants to be illiterate. Also we can improve completeness of answers because the interviewer can make sure all questions are asked. However, training interviewers and conducting surveys is very labour intensive, and participants would receive a low degree of privacy and anonymity.

   Like face-to-face interviews, audio-assisted interviews do not require that respondents be literate, but can provide a higher degree of privacy and anonymity and therefore more accurate data about sensitive questions. Audio-assisted interviews also remove the need for interviewer training; however, they require more equipment, which is difficult to use when surveys must be conducted outside of survey facilities.

2. What topics should be taught during field supervisor training? Outline what you would include in a field supervisor training course.

The entire interviewer training plan plus extra training, including:
- *introduction to the purpose of the study*
- *rapport building and communication skills*
- *cultural sensitivity/knowledge of the target population*
- *how to talk about sex with ease (how to be non-judgmental)*
- *survey protocol*
- *handling and tracking non-response from participants*
- *controlling the interview, handling interruptions and respondent fatigue*
- *use of survey instruments—skip patterns, coding, phrasing and meaning, note-taking*
- *how to respond to participants' questions*
- *ethics: respecting confidentiality, protecting respondents*
- *sampling methods*
- *importance of quality control and standardization*
- *awareness of falsified responses*
- *logistics in fieldwork*
- *safety of interviewers and respondents*
- *management skills*
- *ability to report to upper management.*

Unit 5 answers

Warm-up questions

1. The _____ is the group that meets a survey's measurement objective (for example, all sex workers in a city).

   a. target population

   b. survey population

2. Is choosing names randomly for sampling an example of probability sampling or non-probability sampling?

   a. probability sampling

   b. non-probability sampling

3. The accuracy of a sample survey depends on which factor?

   a. the size of the area included in the survey

   b. size of the sample

   c. the number of interviewers

   d. the number of months set aside for the survey

4. True or false? Non-probability sampling is prone to selection bias. Circle your answer below.

   True

5. You can increase the precision (that is, decrease standard error) of your estimate by increasing the_____.

    a. sample size

    b. quality of interview training

6. Point prevalence is the total number of times something occurs in the population at one point in time divided by the total size of the population.

    True

7. What is the estimate of precision that is used to construct a range of values in which the true population measure is likely to fall?

    a. standard error

    b. systematic sample

8. True or false? You should use non-probability sampling techniques whenever a sampling frame is not available or it is not feasible to develop one.

    True

Apply what you have learned/case study

1. Below, you will find three brief descriptions of surveys and sampling methods. Discuss each approach based on your experience and the information we have just covered on sampling.

    - Identify a problem or two in each approach.
    - Explain what you would do to adjust the sampling approach.

    a. You are participating in the planning for a population survey of all men and women in the country aged 16–55 years. Interview teams will visit homes in 50 clusters (areas chosen for the survey by the National Bureau of Statistics) from 10:00 to 14:00 daily, Monday–Friday. Men and women who consent to the survey will be interviewed on various health and economic topics.

    If you only interview in the hours and days selected, you will miss interviewing men and women who work outside the home during the day. Therefore, your conclusions will not be applicable to all men and women in the country.

    b. A nongovernmental organization conducted a survey among men who have sex with men in a large city to determine condom use.

    - In March, 100 men were interviewed. Condom use was 20%.
    - In September, interviewers returned to the same area to interview another 100 men. This time, condom use was 24%.

    The nongovernmental organization's report said that condom use was increasing in all men who have sex with men due to its outreach efforts.

    Sample size is insufficient to draw this conclusion. Sampling is biased because it was only conducted in one area, and data cannot be generalized to all men who have sex with men.

c. At a truck stop, a university group interviewed (and tested for HIV) a total of 150 truck drivers and their assistants. HIV prevalence was determined to be 8% (range of 5% to 11%). The university's final report stated that 8% (range of 5% to 11%) of truckers in the country are infected with HIV. The report included a plan to provide more care and treatment along all truck routes.

Further study is needed. The survey was conducted at only one truck stop and the sample is not generalizable to the whole country, which is necessary to make a determination and develop such a plan.

2. A survey of men who have sex with men using rds in your city found that 80% could correctly name the ways that HIV is transmitted and ways that HIV can be prevented.

Health education campaigns are planned for the coming year using men who have sex with men peer educators. You wish to show that HIV/AIDS knowledge has improved by at least 10% by the following year using the same method.

How many men who have sex with men do you need to include in this year's survey and next year's survey? Use Table 5.2. Assume that knowledge may increase or decrease and that you wish to have 80% power. Try to check your answer using the sample size formula to detect the change in proportion (Appendix A.5.2).

For the RDS survey, n = 398. From Table 5.2, go down to 0.80 in the first column and 0.90 in the second column. The third column of 199 is the sample size needed for each year's survey if you use simple random sampling (SRS) method. If RDS is used, the sample size should be increased because you should take into account the design effect. The sample size of 398 is the result of using a minimum design effect of 2 (n = 2 × 199 = 398). To refresh your memory about the design effect go to the section entitled "Additional issues in calculating sample size" in Unit 5.

3. You would like to study sex workers in the capital city of Menaland using RDS. Your main objective is to obtain a population-based estimate of health care utilization among sex workers in the city. You have defined health care utilization indicator as the proportion of sex workers who have ever sought treatment or consultation for reproductive or sexual health from public clinics, private clinics or pharmacies in the past year.

To the best of your knowledge, there is no data on the rate of STIs or STI clinic attendance among sex workers in Menaland. Furthermore, there is minimal data on health care utilization in Menaland and other similar countries. Therefore, the sample size calculation has to be done based on your projection considering related data on health care utilization from studies conducted elsewhere. You project that "ever" health care utilization in your proposed survey could be 20% among sex workers in Menaland.

a. Using the sample size formula for point prevalence, a two-sided alpha of 0.05 and a 4% margin of error, calculate how many subjects you must recruit to show that the population-based estimate of the prevalence of health care utilization is 20%.

The formula for calculating the sample size for a simple random sample is as follows:

$$n = \left(\frac{z}{m}\right)^2 p(1-p)$$

Where:

z is the z value (1.96 for 95% confidence level);

m is the margin of error (0.04 = ± 4%);

p is the estimated value for the proportion of a sample that will respond that have utilized health care in the last year (0.20 for 20%).

solving the sample size equation, we find:

$$n = \left(\frac{1.96}{0.04}\right)^2 0.2\,(1-0.2) = 385$$

b. Would you consider adjusting your estimated sample size because you used RDS? Demonstrate how.

We assume that the RDS method will result in a potential design effect of 2 based on similar surveys and on the recommendations.

$$\text{Adjusted } n = 2 \times 385 = 770$$

c. Would it make any difference if you knew that the estimated size of the sex worker population in Menaland is about 2000?

Finite Population Correction (FPC) factor. The sample size equation solving for n′ (new sample size) when taking the FPC into account is:

$$n' = \frac{n}{1 + \dfrac{n}{N}}$$

where

n′ is adjusted sample size for FPC;

n is the sample size based on the SRS calculated above (577);

N is population size (2000).

Calculating the new sample size for Aim 1 using the formula above, we find:

$$n' = \frac{770}{1 + \dfrac{770}{2000}} = 555$$

A sample size of 555 participants should show us the above indicators safely.

Unit 6 answers

Warm-up questions

1. True or false? It is better for staff members to share overall responsibility for data management. Circle your answer below.

   False. It is usually better for one person to have overall responsibility for the data than to share responsibility.

2. Data management does not include which of the following?

   a. data coding

b. data entry

c. data cleaning and checking

d. data framing

3. True or false? The data manager should not be involved during the questionnaire design.

   False. The data manager should be involved during the questionnaire design stage, not only to review the questions, but to review the way in which the questionnaire is organized.

4. Which type of behavioural surveillance analysis is performed to determine whether one variable is related to the distribution of another (for example, an association between a respondent's age and use of condoms)?

   a. univariate

   b. bivariate

   c. multivariate

5. Most of the indicators defined for behavioural surveillance purposes are calculated using _____ analysis.

   a. univariate

   b. bivariate

   c. multivariate

6. True or false? The surveillance cycle ends when the official report is published.

   False.

7. True or false? Data from other sources, including other countries, can be used to fill in any important gaps as long as the source is made very clear.

   True.

Apply what you have learned/case study

What would be the best way to present you data to a senior politician?

In communicating with policy-makers, keep your message to the bare essentials and tell them exactly what action they need to take. Formal pathways may need to be followed when communicating with policy-makers. You should also use all the informal networks you have because getting to see them can be difficult. Politicians and policy-makers, more than any other group, require the right messenger. Very senior policy-makers tend to have a small circle of senior advisers. These people may be easier to approach than the policy-makers, and they usually make very good messengers. Other channels that can be used include senior officials from international organizations or senior diplomats from donor countries.

Unit 7 answers

Warm-up questions

1. Match each ethical principle with its definition.

   <u> b </u>   Respect   for   a. Refers to maximizing benefit and
   persons   minimizing risk to individuals—not
   only physical risk, but also risk of
   psychological harm and stigmatization.

   <u> a </u>   Beneficence   b. Requires investigators to see study
   subjects not as passive sources of data,
   but as persons whose decisions and
   personal choices must be respected.

   <u> c </u>   Justice   c. Risks and benefits from studies should
   be distributed fairly and evenly in
   populations.

2. Complete this sentence. It is important that everyone in contact with the surveillance system, from community leaders and local officials to the people surveyed, is treated with _respect_.

3. What is informed consent?

   Informed consent means that you tell the person enough about the nature of the surveillance for them to make a proper (informed) decision about whether or not to take part.

4. Name two pieces of information that should be given before a person can make an informed decision to take part in a survey?

   - *the nature of the survey, for example, purpose of the survey, length of interview, type of question*
   - *who is conducting the survey*
   - *the potential risks and benefits*
   - *how the information will be used*
   - *how participants' privacy will be protected (names or addresses are not written)*
   - *that participation is voluntary*
   - *that participants have the right to refuse to answer any questions or stop the interview at any time, especially as they may find some of the questions sensitive*

5. True or false? Obtaining verbal consents to survey a most-at-risk population is sometimes better than written consents because we can protect the confidentiality of participants more rigorously. Circle your answer below.

   True. To ensure total confidentiality, it is usually best to obtain verbal consent because the name of the respondent does not need to be recorded.

6. What are two ways to ensure a participant's confidentiality?

   - *ensure names or other means of identification are not recorded on surveys*
   - *store data safely and appropriately*

- *train interviewers in the importance of confidentiality*
- *have clear disciplinary procedures for staff who breach confidentially*
- *identify problems and possible solutions related to confidentiality.*

7. Name two strategies to protect survey participants from loss of earnings.

   a.   Conduct the survey outside of work hours.

   b. Reimburse participants appropriately for their time and participation without being coercive or fostering false expectations of direct benefits to participants.

Apply what you have learned/case study

Design a consent form to be used with sex workers.

Compare your consent form with the two sample consent forms provided in Appendix A.7.1, and discuss the following questions:

1. Which consent form is more similar to what you have?

2. Which sample is more appropriate for the context of your country/region/district?

3. Does Sample 2 capture all information needed for participants to make an informed decision? Why or why not?

   The consent form Sample 2 is more comprehensive than Sample 1 and it addresses important issues that people should know before making their decision to participate. Some investigators may find Sample 2 very long and time consuming. Although having a comprehensive consent form can assure that you are protecting subjects' rights, you should also take into account practical considerations. For example, if the consent form is too long or there are repetitive sections, it may actually distract participants from listening to the entire document.

   It is always helpful to use such samples as a general framework to develop a consent form, but you should try to adapt them to your local ethical requirements and the specific cultural context of the target population.

Final case study answers

Final case study 1 Part 1

You are in charge of the office of communicable disease surveillance in Menaland. The growing HIV epidemic worldwide and in your country has received substantial media attention recently. In your country, as in many other countries, people are advocating for youth education about HIV. A few HIV knowledge, attitude and practice (KAP) studies among high school and university students have shown that general knowledge regarding HIV prevention is weak, that almost half of all students who had had sex in the previous year had multiple partners, and that only 10% had used condoms.

Extramarital sexual relationships and homosexuality are highly stigmatized and those engaged in these practices may face legal prosecution in Menaland. There have been limited surveillance activities among most-at-risk populations (such as injecting drug users, sex workers and men who have sex with men). Sporadic sentinel-based biological HIV surveillance activities were conducted among prison populations that showed a wide range (2%–10%) of HIV prevalence.

The prevention office of the Ministry of Health has successfully lobbied for an allocation of US$4 million for HIV prevention programmes in Menaland. There is strong political movement to implement a national school-based programme to educate youth and develop a life-skills curriculum helping them to stay abstinent and avoid unsafe sex. The programme is going to be launched very soon, and your office receives a portion of the budget to collect baseline data to evaluate the programme.

Questions and discussion points:

1. Do you support or oppose the prevention programme proposed by the prevention office?

2. What specific programme(s) might you propose in response to these data instead?

3. What additional information do you need to make better programming decisions?

Given the limitations of the provided data, it is premature to design a prevention programme based solely on these data. In real life situations there are always competing public health issues, and national prevention programmes should be prioritized. In this case, more evidence about the magnitude of the risk among youth is needed. You should find out about other most-at-risk populations before you decide which group should receive priority attention. You also need to know the size of the population at risk and what specific risk behaviours put them at risk for HIV infection.

Part 2

The first round (baseline) of KAP studies among students aged 16–24 years old was consistent with what previous KAP studies had shown, including relatively weak HIV knowledge and the presence of misconceptions about HIV and negative attitudes toward HIV patients. However, closer inspection of the newly collected data has revealed that only 8% of all students had ever had sex at all, and only 4% had been sexually active in the past year. Also, 2% of all respondents had reported injecting drugs in the previous year.

Meanwhile, you receive a report from the directors of two nongovernmental organizations providing shelter and basic medical treatment for 400 homeless individuals. The nongovernmental organizations are located near hotspot areas known for drug abuse and high crime rates in the capital city of Menaland. The report indicates that of the 400 clients of the shelter/medical treatment, 90% are male, 45% are married, 60% have ever been in jail, 40% are injecting drug users and 30% are HIV-infected. Also, based on classified data from the drug enforcement police, it is estimated that there are 50 000 injecting drug users in the capital city.

Questions and discussion points:

1. What might account for a 30% HIV prevalence among the nongovernmental organizations' clients?

A 30% HIV prevalence is very high. This might be explained by the fact that 40% of homeless people attending the services are injecting drug users and 60% have a history of being in jail. Homeless injecting drug users might have engaged in sharing needles, drugs and drug paraphernalia and/or exchanging sex for money due to a lack of financial resources to buy drugs. Other injecting drug users may also sell sex in exchange for drugs.

2. What does this information tell you about the current HIV/AIDS programming? Would you make any changes in your surveillance activities?

The HIV/AIDS programme has to be revisited, taking into account more accurate epidemiological evidence. Given that 2% of youth have reported injecting drug use, that 30% of homeless people in sentinel-based surveys are HIV infected, and that there is a sizable number of injecting drug users in the capital city, injecting drug users should receive urgent attention in HIV surveillance and prevention activities. However, there are still uncertainties about what specific changes must be made in HIV programming. For example, we must consider that injecting drug users visiting the clinics in the study are not representative of the entire population of drug users in the capital city, and at this stage we can not generalize the 30% HIV prevalence beyond clients of those clinics. Also, we do not know whether the high HIV prevalence is attributed to injection risk behaviour or sexual risk behaviour or both.

3. What would be your next step?

You should begin by forming a working group to focus on the problem and come up with a comprehensive plan.

Part 3

You form a committee comprising experts on HIV to reassess in a systematic manner the current status of the HIV epidemic in Menaland. The committee pools all available data and concludes that the very high rates of HIV among injecting drug users might have been overlooked because health officials, the press and other agencies were looking for evidence of unprotected sexual activity among youth.

Based on the committee's recommendation, you propose a comprehensive plan to assess the HIV epidemic among injecting drug users in Menaland. Harm reduction programmes are redesigned for those already injecting drugs, and injecting drug users are added as the top priority group in the national surveillance system.

Question and discussion points:

1. List the main components of your new comprehensive plan.

You should consider initiating local prevention programmes, such as needle exchange and condom distribution, as soon as possible to halt further spread of HIV among clients of these clinics and their sexual partners and potentially to the nearby communities.

Before implementation of any large-scale prevention programme, you need to gather more comprehensive and accurate data about HIV prevalence and related risk behaviours from individuals in the larger population of drug users in the capital city. Particularly, you should find out what subpopulations, such as injecting drug users, non-injecting drug users, age groups and genders, are at higher risk, and what specific behaviours put them at higher risk.

Using RDS or TLS sampling, you can get this information through linked HIV biological and behavioural surveys among drug users in the capital city. Also, you should set up sentinel-based biological surveys in prisons, rehabilitation centres and drug-treatment facilities. In addition, because youth who report injecting drugs have a greater lifetime risk for becoming infected with HIV, an in-depth investigation to explore injecting risk behaviour among this subpopulation is warranted.

2. Would you consider including other high risk groups, such as sex workers and men who have sex with men in your surveillance activities at this point? Why or why not?

Formative assessment to investigate other most-at-risk populations such as men who have sex with men and sex workers can help you get preliminary data for better HIV surveillance planning for the future and to identify potential overlap between the injecting drug user population and other most-at-risk populations.

3. What specific lessons can be drawn from this example?

Collection of HIV/AIDS epidemiological data is a priority for the development of effective prevention programmes and effective decision-making. In the absence of accurate HIV biological and behavioural epidemiological data, decisions can be misguided and limited public heath financial resources may be wasted on ineffective programmes.

Final case study 2 Part 1

Menaland is a traditional society that has a strong segregation of the sexes, and engagement in homosexual activity has become a significant sexual outlet. Male-to-male sex is varied and often involves men who indulge in commercial sexual activities. This practice is tacitly accepted and involves various traditional categories of transgendered[1] and young male sex workers, many of whom are poor and unemployed.

Recently, a newspaper published a controversial article highlighting the potential public health threat from transgendered sex workers selling sex in Menaland. The article pointed out that in the evening they congregate around certain intersections in the wealthy areas of Menaland's capital city to solicit clients for sex. However, the newspaper neither discussed the size of the transgendered sex worker population nor to what extent their activities might be threatening the health of the general population.

Questions and discussion points:

1. As a public health official, would you take this newspaper article seriously? Why or why not?

2. Would you respond to this situation? How?

Media can play an important role in confronting public health problems and this issue may draw the public's attention. However, before you immediately respond to this article, you should verify whether the claim being made by the newspaper is valid.

Transgendered sex worker sexual practices may be a controversial issue in Menaland. As public health officials we are more concerned about how much their sexual behaviour actually puts them and their clients at higher risk for HIV and other STIs. By blaming one group for the spread of HIV without any substantial scientific evidence, we may further marginalize the group from the society, and make it even harder to work with them to implement prevention and care programmes if they are needed.

Part 2

The Ministry of Health quickly orders a rapid assessment of the transgendered sex worker population of the Emrab district, the most populous district in the capital city,

---

[1]   In Menaland, biologically male transgendered persons with ambiguous sexual identity and behaviour exist who are traditionally identified as a social group within society on the basis of these differences.

with 6 million inhabitants. The report estimates that there are 10 000 transgendered people who are now selling sex as their principal livelihood instead of using their traditional practices to get money, such as singing, dancing and begging. It is also noted that, although the prevalence of HIV is below 0.1% in the general population of the country, it is as high as 6% in the transgendered sex worker population.

Questions and discussion points:

1. Would you immediately proceed to programming based on this report? Why? Why not?

2. Based on this new report, what further steps should be taken?

This preliminary assessment can verify that there is a potential risk, but you need to gather more detailed information to be able to act. You need to conduct a survey to accurately measure HIV sexual risk behaviours in this group.

Part 3

The subnational AIDS Control Programme approves a detailed research study to identify the high-risk practices and HIV status of the transgendered sex worker population in Emrab district. The district's population constitutes about 20% of the city's total population and 4% of the entire country. The district is administratively divided into nine towns, with 150 local councils. The transgendered sex worker population lives in close quarters that are widely distributed throughout the district and are easily identified, so "natural" groupings with variable geographical distribution can be found in the population.

Questions and discussion points:

1. Which sampling technique will be appropriate to recruit from this population?

Because transgendered sex workers are an identifiable population that forms natural groupings which are geographically distributed, you can conduct cluster sampling. Clusters would be areas in each town where groups of transgendered sex workers live. You first need to randomly select a certain number of clusters and within each cluster, to randomly select individuals. The other option is to conduct TLS. In TLS, you need to identify all places and times that transgendered sex workers congregate to sell sex in order to randomly select blocks of time and locations from which to sample individuals (refer to Unit 5). Either way, you need to conduct a preliminary formative assessment to define your clusters.

Part 4

As part of a behavioural survey of sexual practices and knowledge about HIV/AIDS, 350 transgendered sex workers were interviewed in the Emrab district. The survey found that almost 80% were engaged in other occupations along with sex work. Only 45% had adequate knowledge of HIV/AIDS. The level of education of respondents was significantly associated with their knowledge of HIV/AIDS. Fifty-one percent did not use condoms during their sexual encounters, 93% were sexually active with clients and friends, 34% had a history of an STI, 79% practised body piercing, and 60% were willing to take an HIV test.

Questions and discussion points:

1. How will you use this data to respond to the situation?

While 93% of the respondents were sexually active, only 51% used condoms and more than one third had a history of STIs. High STI prevalence is a good marker of high-risk sexual behaviour and potential risk for HIV transmission in this population. A comprehensive HIV prevention programmes to improve knowledge and promote safe sex practices and condom

use should be implemented near areas where transgendered sex workers congregate. Body piercing can be a route of HIV transmission if instruments are reused for multiple people and should be included in HIV prevention messages. HIV testing and counselling should be made available to transgendered sex workers who are willing to take the test so that they can be informed about their HIV status.

2. What specific lessons can be drawn from this example?

In this example, the media helped us to identify and be aware of an important public health issue. However, we are always obliged to collect scientific evidence before taking what may be impulsive and uninformed action.

Final case study 3 Part 1

You are the chief HIV surveillance officer of the Ministry of Health in Menaland. Recent reports indicate that HIV is spreading among injecting drug users in your city. You need to understand the nature of this epidemic, so you decide to do individually linked HIV biological and behavioural surveys among injecting drug users. Most injecting drug users are at higher risk of HIV acquisition and transmission due to sharing needles, but a formative assessment study shows that male-to-male sexual practices and selling sex for drugs are increasing HIV risk as well. Men who have sex with men are extremely stigmatized and there has been no official report on this group so far.

Developing a protocol for this survey, you decide to use the standardized questionnaire designed by Family Health International (FHI) for behavioural surveillance among injecting drug users to save time and money. After translating the questionnaire to your local language, you realise that there are several questions that are less relevant to the culture of injecting drug users in Menaland. Your colleagues suggest that you should invite a small group of injecting drug users to pretest your questionnaire and find potential problems.

Questions and discussion points:

1. Based on your local common practice, discuss whether you must obtain an IRB approval for this survey.

In some countries, surveillance activities are considered programmatic data collection and exempted from human subject review. You need to make sure that you follow local policy regarding surveillance. However, it would be beneficial to get IRB approval if you intend to publish the results of your survey in international peer-reviewed journals.

2. If your answer is yes, discuss whether it would it be ethical to pretest the questionnaire before receiving IRB approval. (To refresh your memory about pretesting, you can refer to Unit 4.)

In pretesting, we do not collect information about individuals' behaviours, rather we collect information about the overall fluency of questions, the order of questions and the appropriateness of the wording used. In most cases, pretesting can be done before securing IRB approval. However, if you think you may need to pretest and pilot test the questionnaire, you should get IRB approval in advance.

Part 2

Although according to your local policy this survey is considered a surveillance activity and is exempted from formal IRB review, your funding agency's policy requires you to get IRB approval in order to receive financial support. You send in your application, knowing that the IRB review always takes time.

The IRB sends back your application requesting that you attach a consent form to it. Given that the participants are a group of injecting drug users, many of whom are very poor, and that you are considering giving them food stamps in return for their participation, you feel that the consent form may not be necessary. Your colleagues further comment that the questionnaire is already long enough and injecting drug users would not tolerate the additional time needed to obtain consent. To get around the issue, your team decides to put together a short consent form, and resends the application. Meanwhile, the team suggests that until you hear from IRB, you can pilot test the questionnaire among a small group of injecting drug users.

Questions and discussion points:

1. Is it ever appropriate to conduct a survey with human subjects in which informed consent is not obtained?

No. You should always obtain informed consent to include human subjects in your survey. Obtaining informed consent assures us that people who participate are fully informed about the process and all potential positive or negative consequences of their participation.

2. Discuss whether it is ethical to pilot test the questionnaire while you are waiting to receive the IRB clearance.

No it is not. During the pilot testing you actually mimic the actual survey with real study instruments and participants, and you may or may not change the content of the questionnaire based on responses that you collect. In some cases, you can even include data collected during pilot testing in your final analysis.

3. Discuss ethical issues around giving monetary incentives to injecting drug users.

The best way to find out about the appropriateness of giving monetary incentives to injecting drug users is by talking to different stakeholders, including people from the injecting drug user community, policy-makers, other researchers, and outreach and social workers. Low amounts may not motivate individuals to participate in your survey. High amounts can be coercive and people may participate in your survey only to receive the money and ignore the fact that there might be negative consequences.

4. Given the nature of your survey, develop a verbal consent form.

Part 3

The IRB returns your consent form, this time asking for some revisions—specifically the inclusion of the two paragraphs below:

"Some of the questions I will ask are personal and include discussion of sexual behaviour and drug use. If you are embarrassed or uncomfortable about answering any questions, you may choose to skip them or you may end the interview at any point."

"To minimize any risk or discomfort that taking part in this interview may have for you, all information you give us will be anonymous (will not identify you personally). This means that you do not have to say your name or give any identifying information at any point. I will not collect any information that can later be used to identify or contact you. Only an assigned identification code will be attached to your responses."

Questions and discussion points:

1. For each paragraph, discuss which ethical principles they are referring to.

The first paragraph refers to "respect for the person". We should not force people in any way to answer all questions. Individuals may consent to participate in the survey, but their right to skip questions that they do not want to answer should be preserved.

The second paragraph refers to "beneficence" because there might be potential discomfort or physiological harm as a result of their participation in the survey. The investigators must protect participants against both physical risk and risk of psychological harm and stigmatization. Investigators should be working to minimize those potential risks to individuals.

2. Have you already addressed these ethical principles in your consent form? If so, how?

3. Compare these suggested paragraphs with what you have already included in your consent form, and discuss why we should or should not add these paragraphs.

Annex 4

Action plan for implementing HIV behavioural surveillance of most-at-risk populations in EMR/MENA countries

Action plan for implementing HIV behavioural surveillance of most-at-risk populations in EMR/MENA countries

Introduction

TO BEGIN THE process of operationalizing biobehavioural surveillance of most-at-risk populations, you will need to:

- develop country-specific operational guides for HIV surveillance
- develop country-specific implementation workplans

What is an action plan?

A well-developed action plan allows you to:

- establish clear goals and objectives
- present your ideas to achieve consensus among all persons involved
- establish a realistic budget
- ensure that the appropriate staff in each facility are trained on the surveillance activities
- determine activities
- determine responsible persons
- establish a timeline for completion of activities
- ensure ethical concerns are addressed
- establish a data analysis and dissemination plan.

Over the next couple of days, you will be developing a draft action plan and will discuss your action plan with the group.

You will use the worksheet for developing a national plan (Worksheet 1).

National action plan worksheet

Instructions

- If you are not doing so already, move now to sit with all the participants from your country.
- As a team, take some time to discuss and fill out the questions on the national action plan worksheet on the next page (the facilitator will hand out a clean version of this form). This worksheet will help you determine the activities needed to implement your country's surveillance activities. This worksheet will also help us to determine each country's plans and schedules, as we will be providing technical assistance where needed.
- We understand that you may not know all the final answers to the questions in the national action plan worksheet. Please fill in the information as completely and accurately as possible.
- If you have questions, facilitators will be available to discuss the worksheet process with you.

Timing

Choose a scribe to keep track of your team's comments on each area of the worksheet.

Spend about 30 minutes on this worksheet.

Worksheet for developing an action plan

(Worksheet 1)

---

1. What is the name of your country?

---

2. Which most-at-risk populations do you plan to include in your surveillance activity?

---

3. Are you planning biological or behavioural surveillance, or both?

---

4. Have funds been allocated for these activities? If so, how much?

---

5. If funds have not been allocated, what are your plans for securing funding?

---

6. When do you plan to begin your surveillance activity?

---

7. Which organization will act as the coordinating agency for the surveillance activity?

---

8. Who will be the focal point at this agency (name, title and contact details)?

---

9. Who are the stakeholders who will review your plan? Please provide names, if possible.

   Ministry of Health:

   WHO Regional Office for the Eastern Mediterranean:

   Nongovernmental organizations:

   Donors:

   Other:

---

10. List people who will be working to complete the actions in the action plan and list their positions.

    Develop a contact list with name, address, phone number, fax number, e-mail address and the role of each person.

    ● Finalize operational procedures manual:

    ● Define indicators:

    ● Develop data collect forms:

    ● Coordinate training (logistics, materials):

    ● Instructors:

---

11. List personnel in need of training.	Laboratory:
	Field staff (recruiters, interviewers, phlebotomists, etc):
	Project managers:
	Statistical support (data managers, data entry clerks, data analysts):

12. What is the estimated number of people in need of training?

13. What are the best dates to conduct trainings? List conflicting meetings/holidays during which the trainings cannot be held.

14. Are you aware of any sites where training can be conducted? If yes, please list the name and type of facility and how many people it can accommodate at one time.

15. Challenges in implementing your action plans can include:

- few or no designated trainers
- lack of or conflicting policies
- lack of needed materials
- schedule conflicts
- lack of money
- turnover/attrition of staff

List your possible challenges to the right.

16. Is there an ethics review committee in your country that can review your surveillance plans, survey instruments and consent documents? If yes, who is the contact person on the ethics committee?

17. List resources that you may be missing.

18. Is technical assistance needed? If so, what? How will you go about receiving it?

_____

Discuss your worksheet with a facilitator or go on to the next part of the action planning process.

Action plan development

Introduction

Now use the information from your national worksheet to develop an action plan. An action plan helps you keep track of all activities and helps you review your progress. We have provided an action plan template for your use.

Instructions

- Remain in your country group.
- Address the key elements listed below.
- Add additional elements if you wish.

Timing

- You will need a calendar for this activity.
- Take 30 minutes for this activity.

Activities

While completing your action plan, please address the key elements listed below (these activities and general timeframe have been added to the action plan template on the next page):

- identify stakeholders; debrief Ministry of Health and National AIDS Programme (within one month)
- finalize operational procedures manual (within two months)
- finalize test forms (within two months)
- train interviewers, phlebotomists (if being used) and laboratory personnel
- train data-entry persons (identify back-up)
- train data analysts
- develop a data dissemination plan.

You will want to consider other important areas and may add any of these to your action plan:

- determine budget
- determine final training dates
- select the appropriate audience for training
- adapt the training curriculum from existing materials
- organize the training(s) (facility, audiovisual equipment, supplies)
- evaluate the training
- conduct follow-up activities and site visits after the training to reinforce learning.

If you have other details you would like to add, please add them.

Due dates

Adding due dates to an action plan helps you establish a realistic schedule. Here is the sequence of events:

- list your activities
- put the activities in the order in which you (or your team) will do them
- add due dates.

Why have due dates?

Having due dates:

- provides the overall picture for planning your programme
- helps keep your project on schedule
- avoids assigning too many things to one person
- helps you to meet your programme goals and objectives
- helps you to remember critical steps so nothing is forgotten in the planning process.

How to choose due dates

When you are developing due dates, think about:

- the order of activities
- which activities are dependent on earlier activities
- the overall timeframe for completing the entire activity
- what factors might cause someone to miss a due date, such as existing schedules, commitments, holidays, vacation schedules or any other sources of delay.

It is important to remember to include the people who will be involved and who will be responsible for meeting the due dates. If they are involved in the decision-making about important issues, especially deadlines, the team that is involved will be more likely to meet those deadlines. Everyone involved should receive a copy of the agreed-to action plan.

Action plan template

An action plan template is provided on the next page.

- You may use the template or any other format you desire.
- Use the notes and discussion you had when you filled out your worksheet.
- Change the order of the activities or add additional activities.
- Check your calendar to assign realistic due dates for each activity. Some suggested timeframes have been added to the activities. You may change those if you wish.

Your presentation

- Develop your presentation as you complete the action plan or after it is done.
- Choose a member of your team to present the action plan to the group.
- The facilitators have a slide master, or you may use one of your own.

National action plan:
(your country and title of the plan)

(Worksheet 2)

Activities	Responsible person	Resources needed	Challenges/ solutions	Target due date	Actual completion date
1. Debrief Ministry of Health and National AIDS Programme (within one month)					
2. Finalize operational procedures manual (within two months)					
3. Finalize test forms (within one months)					
4. Conduct training of providers and laboratories (within one month of finalizing forms and operational manual)					
5. Talk with statistics office to obtain death records (within two months)					
6. Check national database to make sure it is set up appropriately					
7. Train data-entry staff and back-up staff.					

Activities	Responsible person	Resources needed	Challenges/ solutions	Target due date	Actual completion date
8.					
9.					
10.					
11.					
12.					
13.					
14.					
15.					
16.					

Surveillance is the systematic, regular collection of information on the occurrence, distribution and trends of a specific infection, disease or other health-related event.

Second generation HIV surveillance is designed to collect and integrate data reported from a variety of sources including behavioural surveillance, HIV case reporting, HIV seroprevalence surveillance, and sexually transmitted infections surveillance. The goals of second generation HIV surveillance are to help countries better understand the HIV epidemic trends over time, to better understand the behaviours driving the epidemic, to focus on subpopulations at highest risk for infection and to better use surveillance data for planning the response to HIV epidemic.

HIV surveillance in the Eastern Mediterranean Region needs to be strengthened in order to fill the gaps in our understanding of the dynamics of the epidemic and to be in a better situation to plan appropriately for an effective response.

This training module is one of four selected from a series originally developed by University of California, in San Francisco. The selected modules were adapted to the regional context through a long process involving consultations with HIV regional experts. These modules are expected to provide a good tool for training staff working in HIV surveillance at country level. Countries are free to further adapt these modules or to translate them into local use.