

GUIDELINES

CONDUCTING HIV SURVEILLANCE BASED ON ROUTINE PROGRAMME DATA

AMONG PREGNANT WOMEN
ATTENDING ANTENATAL CLINICS

UNAIDS/WHO working group
on global HIV/AIDS and STI surveillance

AUGUST 2015



HIV STRATEGIC INFORMATION FOR IMPACT

WHO Library Cataloguing-in-Publication Data:

Guidelines for conducting HIV surveillance among pregnant women attending antenatal clinics based on routine programme data.

- 1.HIV Infections – epidemiology.
- 2.Maternal Health Services.
- 3.Pregnant Women.
- 4.Data Collection.
- 5.Epidemiologic Methods. I.World Health Organization.

ISBN 978 92 4 150972 5 (NLM classification: WC 503.4)

© **World Health Organization 2015**

All rights reserved. Publications of the World Health Organization are available on the WHO website (www.who.int) or can be purchased from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; e-mail: bookorders@who.int).

Requests for permission to reproduce or translate WHO publications –whether for sale or for non-commercial distribution– should be addressed to WHO Press through the WHO website (www.who.int/about/licensing/copyright_form/en/index.html).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

CONDUCTING HIV SURVEILLANCE BASED ON ROUTINE PROGRAMME DATA

**AMONG PREGNANT WOMEN
ATTENDING ANTENATAL CLINICS**

UNAIDS/WHO working group
on global HIV/AIDS and STI surveillance

AUGUST 2015

ACKNOWLEDGEMENTS

Global surveillance of HIV and sexually transmitted infections is a joint effort of WHO and the Joint United Nations Programme on HIV/AIDS (UNAIDS).

This document is based on experience in countries and expert reviews from WHO/UNAIDS and the United States Centers for Disease Control and Prevention (CDC). WHO would like to thank the individuals who contributed to this document.

The development of this document was coordinated by Jesus M. Garcia Calleja of WHO and Jacob Dee of CDC.

Major technical input was provided by Lori Newman of WHO; Steve Gutreuter, Chris Murrill and Irum Zaidi of CDC; Simon Gregson of Imperial College London; and Basia Zaba of London School of Hygiene and Tropical Medicine.

Additional technical review was provided by Chika Hayashi of WHO; Jamie Houston, Andrea Kim, Bharat Parekh, Ray Shiraishi and Peter Young of CDC; Webster Kasongo of the Zambia Tropical Diseases Research Centre; Davies O. Kimanga of the Ministry of Public Health and Sanitation, National AIDS and Sexually Transmitted Infection Control Programme, Kenya; Wegene Tamene Mekasha of the Ethiopian Public Health Institute; and Issa B. Kawu of the Federal Ministry of Health, Nigeria.

The findings, conclusions and recommendations in this report are those of the WHO and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Funding to support this work was provided by the US President's Emergency Plan for AIDS Relief UNAIDS and WHO.

ABBREVIATIONS

AIDS	acquired immunodeficiency syndrome
ANC	antenatal clinic
ART	antiretroviral therapy
CDC	United States Centers for Disease Control and Prevention
EIA	enzyme immunoassay
EMR	electronic medical record
EMTCT	elimination of mother-to-child transmission
EQA	external quality assessment
HIV	human immunodeficiency virus
HMIS	health management information system
HTC	HIV testing and counselling
MCH	maternal and child health
PDA	personal digital assistant
PLHIV	people living with HIV
PMTCT	prevention of mother-to-child transmission
STI	sexually transmitted infections
UAT	unlinked anonymous testing
UNAIDS	Joint United Nations Programme on HIV/AIDS
WHO	World Health Organization

CONTENTS

1. Introduction	8
1.1 Purpose of this guidance document	8
1.2 HIV surveillance among pregnant women in the context of second-generation surveillance: Know your epidemic	9
1.3 Use of routine programme data for surveillance among pregnant women attending ANC	9
1.4 Engaging partners to strengthen surveillance	10
1.5 Organization of this guidance	11
2. Ethical considerations	12
3. Surveillance design	15
4. Variables for ANC surveillance	17
4.1 Selecting variables for surveillance	17
4.2 Surveillance of syphilis among pregnant women attending ANC	19
5. Census of ANC	20
5.1 Approaches to a census of ANC	20
5.1.1 Census based on individual-level data	20
5.1.2 Census based on aggregate data	21
5.2 Site selection for a census of ANC	22
5.3 Duration of surveillance period for a census of ANC	23
5.4 Variable considerations for a census of ANC	23
5.5 Eligibility criteria for a census of ANC	23
5.6 Data collection for a census of ANC	24
5.6.1 Data transfer	24
5.6.2 Data standards	25
5.6.3 Staffing for data collection and management	26

5.6.4 Collecting routine records with missing data	27
5.7 Data management for a census of ANC	27
5.7.1 Data security	27
5.7.2 Data cleaning	28
5.8 Analysis of data from a census of ANC	29
5.8.1 Calculating prevalence	29
5.8.2 Examining trends in HIV prevalence over time	30
5.8.3 Estimating national prevalence	30
5.8.4 Analysis of routine records with missing data	30
6. Sentinel surveillance design	32
6.1 Selection of ANC sites in sentinel surveillance	32
6.2 Duration of the sentinel surveillance period	34
6.3 Frequency of sentinel surveillance rounds	34
6.4 Sample-size for sentinel surveillance	34
6.4.1 Sample-size for a single prevalence proportion	34
6.4.2 Sample-size for detecting changes in HIV prevalence	36
6.4.3 Sample-size considerations for young pregnant women	36
6.5 Eligibility criteria for sentinel surveillance	37
6.6 Data collection for sentinel surveillance	37
6.6.1 Defining the routine data source	39
6.6.2 Defining a surveillance code	39
6.6.3 Retrospective data collection methods	39
6.6.3.1 Staffing and training for retrospective data collection	39
6.6.3.2 Sampling records of pregnant women in retrospective data collection	41
6.6.3.3 Methods to collect data in retrospective data collection	41
6.6.3.4 Quality assurance of retrospective data collection	43
6.6.4 Real time data-collection methods	44
6.6.4.1 Staffing and training for real time data collection	44
6.6.4.2 Sampling pregnant women for real time data collection	44



6.6.4.3 Methods to collect data in real time data collection	45
6.6.4.4 Quality assurance of real time data collection	47
6.6.5 Collecting routine records with missing data	49
6.7 Data management for sentinel surveillance	49
6.7.1 Data security	49
6.7.2 Data entry and cleaning	50
6.8 Data analysis for sentinel surveillance	50
6.8.1 Calculating site prevalence	50
6.8.2 Comparing HIV prevalence from two surveillance rounds	51
6.8.3 Examining trends in HIV prevalence over time	51
6.8.4 Aggregating data and reporting summary prevalence	52
6.8.5 Estimating national prevalence	52
6.8.6 Analysis of routine records with missing data	52
7. Monitoring of surveillance	54
7.1 Monitoring individual-level data quality: data completeness and validity	55
7.2 Monitoring data quality: routine site data tools	55
7.3 Monitoring aggregate data quality: HMIS/routine reporting systems	56
7.4 Monitoring routine HIV testing	57
7.5 Addressing underperforming sites	58
8. Interpretation of results	59
8.1 Issues of representativeness and generalizability	59
8.2 Considerations affecting the interpretation of trend results	62
8.2.1 Incidence and mortality	63
8.2.2 Trends among young pregnant women	64
8.2.3 Interpretation of surveillance results across systems	64
8.3 Other potential biases	65
8.4 Interpretation of syphilis surveillance results	66

9. Dissemination of surveillance results	69
10. Appendices	71
A1 Sample-size for a single prevalence proportion in sentinel surveillance	71
A2 Sample-size for detection of changes in HIV prevalence in sentinel surveillance	72
A3 Sample-size considerations for young pregnant women in sentinel surveillance	73
A4 Modelling analysis of trends in HIV prevalence from sentinel surveillance	74
A5 Example of a sentinel surveillance data-collection form	76
11. References	78



1. INTRODUCTION

In this section

- 1.1 Purpose of this guidance document
- 1.2 HIV surveillance among pregnant women in the context of second-generation surveillance: know your epidemic
- 1.3 Use of routine programme data for surveillance among pregnant women attending ANC
- 1.4 Engaging partners to strengthen surveillance
- 1.5 Organization of this guidance

1.1 Purpose of this guidance document

This document is written for national HIV surveillance programme staff responsible for monitoring trends in country HIV epidemics. Its purpose is to describe how routine prevention of mother-to-child transmission of HIV (PMTCT) programme data can be used to conduct HIV surveillance among pregnant women attending antenatal clinics (ANC). **These guidelines assume that surveillance programmes have already assessed the readiness of routine programme data to be used for surveillance. WHO's 2013 *Guidelines for assessing the utility of prevention of mother-to-child transmission (PMTCT) programme data for HIV sentinel surveillance among pregnant women* describes these assessment methods (1).**

Technical guidance in this document covers:

- surveillance design and implementation
- site selection
- sample-size calculation
- data collection
- analysis and interpretation
- monitoring of surveillance
- dissemination of results.

The guidance is based on the input of surveillance experts and surveillance programmes, and lessons learnt from the field. Countries may adopt or adapt this guidance as appropriate.

1.2 HIV surveillance among pregnant women in the context of second-generation surveillance: know your epidemic

Countries require information on HIV prevalence to monitor the course of their HIV epidemics and to allocate resources, and plan and evaluate interventions, for HIV control and prevention (2). Over the past two decades, HIV surveillance among pregnant women attending ANC has provided valuable information about the burden of HIV and trends in HIV prevalence among pregnant women attending sentinel ANC sites. Historically, ANC sentinel surveillance has been conducted using serosurveillance based on unlinked anonymous testing (UAT) (3).

ANC provide an accessible cross-section of healthy, sexually active women in the general population. Although the limitations and biases inherent in surveillance among pregnant women attending ANC are recognized, HIV surveillance estimates from ANC can serve as a useful proxy for HIV prevalence trends in the general female population (4). Epidemiologic information provided by HIV surveillance among pregnant women attending ANC, combined with other sources of surveillance data (e.g. household surveys, key population surveys, HIV case surveillance and surveillance of sexually transmitted infections [STIs]), allow surveillance programmes to address key aspects of the “know your epidemic” approach of second-generation surveillance of HIV. These aspects include understanding changes in the direction of the epidemic, understanding subnational variations in an epidemic and identifying localized geographical areas with higher burdens of HIV (5).

Finally, data from HIV surveillance among pregnant women attending ANC are an important input for statistical models that estimate the national and subnational burden and incidence of HIV using estimation and projections tools developed by the Joint United Nations Programme on HIV/AIDS (UNAIDS), including Spectrum (6).

1.3 Use of routine programme data for surveillance among pregnant women attending ANC

HIV prevention and treatment programmes have improved dramatically in the past decade. Improvements include more widespread access to diagnostic HIV testing in ANC settings, higher coverage of prevention of PMTCT and antiretroviral therapy (ART) programmes, and ongoing transition to initiating immediate, lifelong ART (“option B+”) for HIV-positive pregnant women (7). In 2014, an estimated 50% of pregnant women in low- and middle-income countries had an HIV test and received their results. In addition, an estimated 1,079,713 (73% [68-79%]) pregnant women living with HIV in low- and middle-income countries received ARTs to prevent mother-to-child transmission of HIV (8).

As programmes expand and strengthen, routine programme data are increasingly able to provide most of the demographic, HIV and syphilis data previously captured by UAT-based serosurveillance. **For the purpose of these guidelines, routine programme data are**

defined as data that are routinely generated through ANC and PMTCT service delivery, and routinely recorded in standard site data tools (e.g. registers).

HIV surveillance among pregnant women attending ANC based on routine programme data offers multiple advantages over UAT serosurveillance in so far as it:

- adheres to ethical standards by ensuring that HIV serostatus data used for surveillance comes from HIV testing that includes, as per routine clinical practice:
 - the informed and free choice to accept or decline (i.e. opt out of) testing;
 - pre- and post-test counselling;
 - return of test results;
 - referral to PMTCT, HIV care and treatment services if test results are positive;
- requires fewer human and financial resources than UAT serosurveillance, by eliminating the collection and testing of additional biological specimens in real time;
- provides the opportunity to expand coverage of the surveillance system; and
- supports the sustainability of surveillance efforts by contributing to a system and culture of using routine data for surveillance.

1.4 Engaging partners to strengthen surveillance

HIV surveillance among pregnant women attending ANC based on routine programme data relies on effective functioning, monitoring, quality assurance and strengthening of inputs and activities owned and managed by Multiple programmes, including those whose main function is not related to surveillance. These inputs and activities include:

- *routine programme data tools (e.g. registers)* – data-collection standards and ANC personnel are typically managed by maternal and child health (MCH) or PMTCT programmes;
- *routine HIV and syphilis testing* – typically managed by PMTCT and MCH programmes;
- *HIV and syphilis testing strategies and validated testing algorithms and quality assurance* – typically overseen by the national reference laboratory;
- *HIV and syphilis test kit procurement and supply chain logistics* – typically managed by central medical stores and the programmes themselves;
- *health management information systems (HMIS)/routine reporting systems* – typically managed by an informatics or information systems unit; and
- *electronic medical record (EMR) systems* – typically managed by service programmes, an informatics unit or health partners.

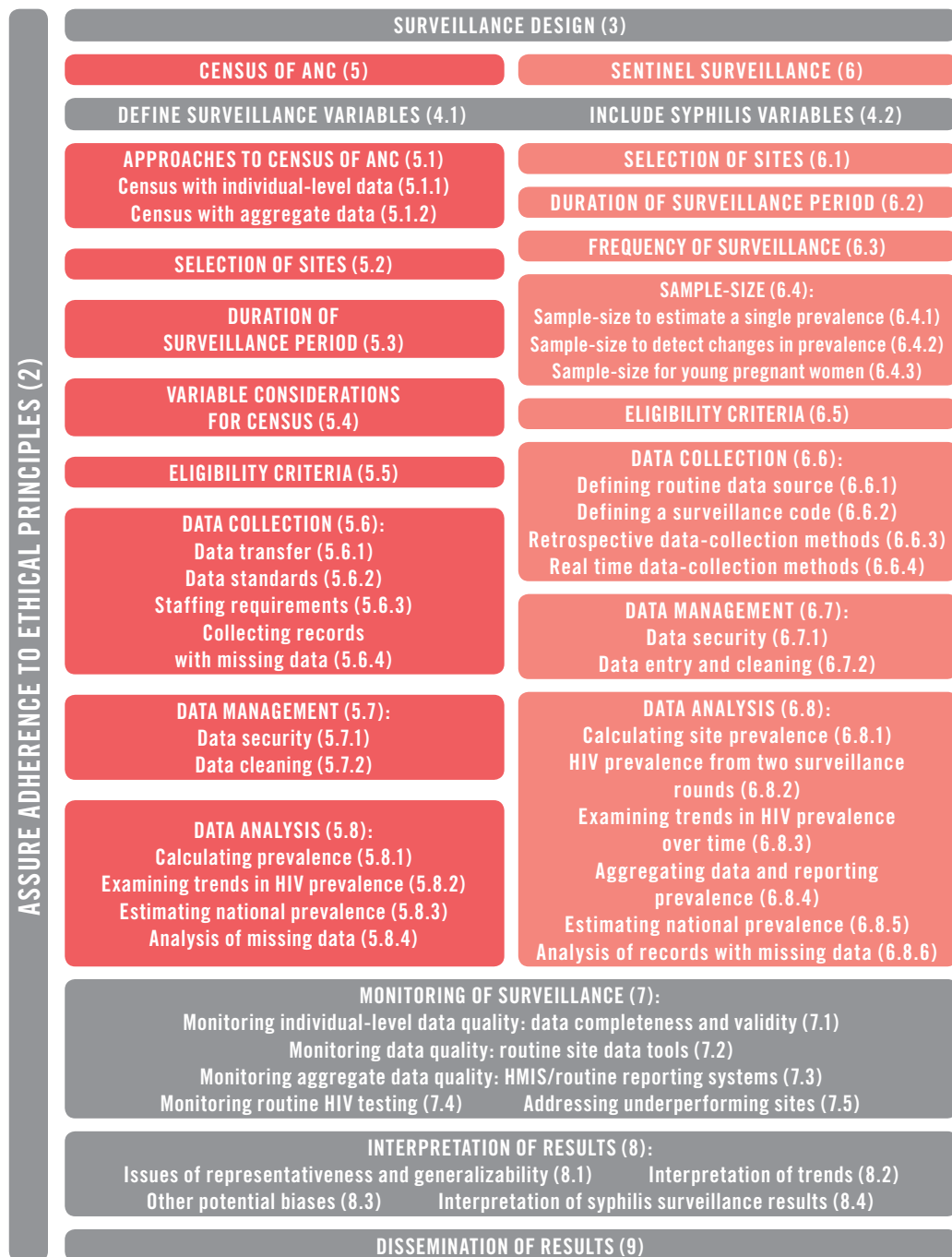
As a result, surveillance based on routine programme data should be a collaborative endeavour that engages programmatic partners early and often. Where surveillance based on routine programme data identifies gaps or challenges in the quality of routine data or HIV testing, these results should be used to inform collaborative system-strengthening measures. Surveillance based on routine data should be an opportunity to improve routine activities to benefit service delivery, programme monitoring and surveillance.

1.5 Organization of this guidance

This document is organized by surveillance design; that is, by sentinel surveillance and census. Within each design, the guidance is organized to accompany surveillance programmes through each step of designing and implementing HIV surveillance among pregnant women attending ANC based on routine programme data. The guidance covers sample-size, data collection (real time or retrospective), data management, data analysis, surveillance monitoring, and results interpretation and dissemination. Fig. 1 broadly illustrates these steps.



Fig. 1. Surveillance planning, implementation and organization of this guidance (by section no.)



2. ETHICAL CONSIDERATIONS

All surveillance activities, including HIV surveillance among pregnant women attending ANC based on routine programme data, should adhere to ethical principles of biomedical research and surveillance. These principles include (9, 10):

- respect for persons (enable and respect the choice of competent individuals, and protect vulnerable persons);
- beneficence (maximize possible benefits and minimize possible harms to participants); and
- justice (the burdens and benefits of surveillance are shared equally by all populations)

Protocols for HIV surveillance among pregnant women attending ANC based on routine programme data should be reviewed and approved by all relevant ethical review boards before implementation.

Informed consent of pregnant women is not required for HIV surveillance among pregnant women attending ANC based on routine programme data. All surveillance data (including biomarkers) come exclusively from information routinely generated and recorded by ANC and PMTCT service delivery through standard clinical practice. No data or biological specimens are collected from pregnant women for the purpose of surveillance.

HIV surveillance among pregnant women attending ANC based on routine programme data should include safeguards to protect the rights, welfare and confidentiality of pregnant women whose data are collected for surveillance. These safeguards should include human, physical and electronic protections to ensure confidentiality and data security at every stage of surveillance, including data collection, transfer, storage, analysis and dissemination (Sections 5.7.1 and 6.7.1). Additionally, all surveillance staff collecting or handling data should be trained in data security and confidentiality procedures, and should sign a confidentiality agreement in accordance with national data security policies.

Routine programme records used for HIV surveillance among pregnant women attending ANC contain personal identifying information (e.g. name and ANC service number) and medical information of a sensitive nature (e.g. HIV status). To protect the confidentiality of pregnant women whose routine data are collected for surveillance, HIV surveillance among pregnant women attending ANC should collect as little personal identifying information as possible, and preferably should not collect such information. Personal identifying information is generally not needed for surveillance data collection or analysis. Any surveillance codes or unique identification numbers should not be linked to personal identifying information.

In some contexts, personal identifying information or linked codes may be needed to ensure that data-collection processes are valid. Codes or personal identifying information may be required to exclude duplicate records or to link records from multiple routine data sources

(e.g. an ANC register and an HIV-testing register) to make up the complete surveillance dataset. In such cases, personal identifying information should be permanently deleted from the dataset once the information has served its data-collection purpose.

Further guidance on this subject is provided in WHO's 2013 *Guiding principles on ethical issues in HIV surveillance* (9).



3. SURVEILLANCE DESIGN

The surveillance design should be formulated such that **HIV surveillance among pregnant women attending ANC provides reliable epidemiological data and is resource efficient.**

These guidelines present two designs for HIV surveillance among pregnant women attending ANC, which can be briefly summarized as follows:

- A *census of ANC* (Section 5) is the gathering of individual-level or aggregate data from all pregnant women attending all (or nearly all) ANC sites providing PMTCT services, by leveraging above-site (e.g. regional and national) data repositories. Census can collect individual-level data (using existing EMR systems) or aggregate data (using HMIS/routine reporting systems).
- A *sentinel surveillance design* (Section 6) is a convenience sample of ANC sites that are chosen to represent geographical areas or populations of interest for HIV surveillance. This is the design historically used for HIV surveillance among pregnant women attending ANC.

Census is the preferred design because it is highly resource efficient and provides complete geographical coverage of surveillance of pregnant women attending ANC. Census based on individual-level data (from EMR systems that have a verified high degree of accuracy, completeness and integrity) is preferable because it allows richer analyses, greater modeling opportunities and increased ability to monitor data quality. However, census based on aggregate data can also provide core surveillance data if HMIS/routine reporting systems have a verified high degree of accuracy, completeness and integrity.

Countries are encouraged to strengthen and expand EMR systems and HMIS/routine reporting systems to progress, over time, towards a census approach.

At present, many countries may not be ready to use the census approach. EMR systems are not widely present at ANC sites, and many HMIS/routine reporting systems face quality challenges. Hence, this document also provides guidance on the sentinel surveillance design approach.

Table 1 summarizes the advantages, limitations and additional requirements of each surveillance design.

Table 1. Comparison of designs for HIV surveillance among pregnant women attending ANC

	Census: collection of individual-level or aggregate data from all women attending all (or nearly all) ANC sites that provide PMTCT services	Sentinel surveillance: convenience sample of ANC sites chosen to represent different population characteristics
Advantages	<ul style="list-style-type: none"> • Involves no surveillance sampling • Requires minimal surveillance fieldwork • Produces direct measures (not estimates) of HIV prevalence and changes in prevalence among ANC attendees • Provides comprehensive coverage of surveillance system • Provides HIV prevalence measures at all subnational levels 	<ul style="list-style-type: none"> • Because sites are selected by convenience sampling, surveillance can pre-screen or specially capacitate (or both) ANC sites so that only high-performing ANC sites participate in surveillance
Limitations	<ul style="list-style-type: none"> • Aggregate data may limit ability to conduct subanalyses (e.g. by age) or monitor data quality • It is not possible to pre-screen ANC sites for readiness to participate in surveillance – underperforming sites will be included in surveillance 	<ul style="list-style-type: none"> • Does not provide representative, statistically valid estimates of HIV prevalence among ANC attendees at national or subnational levels
Additional requirements	<ul style="list-style-type: none"> • Availability of HIV testing in all, or nearly all, ANC sites • To use individual-level data: requires high coverage of EMR systems at ANC sites, and EMR data should have a verified high degree of data accuracy, completeness, integrity and accessibility • To use aggregate data: requires an HMIS/routine reporting system with a verified high degree of data accuracy, completeness and integrity • Most ANC sites should be ready to participate in surveillance 	<ul style="list-style-type: none"> • Requires fieldwork to collect routine data from ANC sites

ANC, antenatal clinics; EMR, electronic medical record; HIV, human immunodeficiency virus; HMIS, health management information system; PMTCT, preventing mother-to-child transmission

4. VARIABLES FOR ANC SURVEILLANCE

In this section

- 4.1 Selecting variables for surveillance
- 4.2 Surveillance of syphilis among pregnant women attending ANC

4.1 Selecting variables for surveillance

Surveillance programmes should define which variables will be collected for surveillance of pregnant women attending ANC. All variables needed to conduct appropriate epidemiological analyses should be selected. The following core variables should be included as a minimum:

- *Surveillance code*: All eligible pregnant women whose records are sampled by surveillance should be assigned a surveillance code (Section 6.6.2).
- *Date of visit*: This variable is needed to ensure that the pregnant women's visit falls within the surveillance period, and to maintain quality assurance of data collection.
- *Age*: Trends in HIV prevalence should be monitored by age, to understand the distribution of HIV infection across age groups.
- *Uptake of routine HIV testing*: Surveillance programmes should collect information on whether pregnant women received HIV testing, declined HIV testing, did not receive testing because they were previously known to be HIV-positive, or did not receive HIV testing for other reasons (e.g. test kit stock out, operation problems with rapid tests or counsellor not present). Information on uptake of routine HIV testing can serve as a measure of the coverage of HIV-testing programmes in ANC settings, and is important for monitoring potential selection bias in surveillance based on routine programme data.
- *HIV status*: This includes:
 - final HIV-testing results (i.e. results of the full HIV-testing algorithm) for pregnant women receiving an HIV test at their first ANC visit; if possible, surveillance programmes should also collect the results of each individual test in the HIV-testing algorithm;
 - pregnant women who were not HIV tested but who already know their status to be HIV-positive when presenting at their first ANC visit; and
 - HIV status not documented: non-documentation of HIV status may reflect missing data, decline of HIV testing or the unavailability of HIV testing; this lack of information should be recorded.

- *Syphilis test done, syphilis test result and syphilis treatment*: Surveillance should collect data on whether pregnant women received syphilis testing, what type of assay was done (treponemal or non-treponemal), the results, and whether treatment was provided (see Section 4.2 for further information on the inclusion of syphilis in HIV surveillance among pregnant women attending ANC).
- *ART status*: If possible, surveillance should collect data on whether HIV-positive pregnant women are already on ART at their first ANC visit.

Surveillance programmes can collect additional, non-core variables, if these variables are available in routine records and are considered important for HIV surveillance and control efforts. These data can also assist in characterizing the catchment population of the site. Examples of non-core variables include:

- *gravidity or parity (or both)*: gravidity (the total number of pregnancies) and parity (the total number of children borne by a woman, excluding miscarriages or abortions in early pregnancy but including stillbirths) are used to examine the association between HIV infection and exposure to unprotected sex; in contexts when barrier contraceptive use is high, gravidity or parity may be a better measure of sexual exposure than a pregnant woman's age (11);
- marital status;
- educational level;
- occupation; and
- other biomarkers (if hepatitis B or C testing, or other biomarkers of interest, are routinely performed at ANC, these can be considered for surveillance variables).

In selecting variables to be collected by surveillance, it is important to take into account the following considerations:

- In general, personal identifying information is not needed for surveillance purposes and should not be collected for surveillance. In some contexts, personal identifying information or linked codes may be needed to ensure valid data-collection processes. Codes or personal identifying information may be required to exclude duplicate records or to link records from multiple routine data sources (e.g. an ANC register and an HIV-testing register) to make up the complete surveillance dataset. In such cases, personal identifying information should be permanently deleted from the dataset after its data-collection purpose has been fulfilled.
- All variables selected for surveillance should be available in routine records. Because pregnant women have not provided informed consent to participate in a survey, no data should be collected from pregnant women specifically for the purposes of surveillance.

- Variables selected for surveillance should be reasonably accessible. Variables requiring substantial additional time or effort to collect (e.g. consulting an additional register to collect a variable not available in the standard ANC or PMTCT register) should provide epidemiological or public health value that justifies the additional effort.
- Surveillance should collect only variables that are needed for epidemiological analysis and public health action. Data that will not be analysed, disseminated and acted upon should not be collected.

4.2 Surveillance of syphilis among pregnant women attending ANC

Historically, HIV serosurveillance of ANC attendees frequently included surveillance for maternal syphilis. This integrated approach to surveillance was both practical and epidemiologically appropriate. Syphilis and HIV share many transmission pathways; therefore, syphilis infection may be a marker of increased HIV risk. Additionally, syphilis infection in HIV-positive women is an important marker of increased risk of ongoing HIV transmission.

Routine syphilis testing in pregnancy is considered an essential antenatal care intervention by WHO, and is supported by national policy in most countries (12). Syphilis screening for reactive syphilis serology using rapid diagnostic tests, and treatment of seropositive women with at least one dose of injectable penicillin, is highly effective and cost-efficient in preventing adverse outcomes such as stillbirth, low birth weight, prematurity, neonatal death and congenital disease in the newborn. Without intervention, syphilis in pregnancy will lead to an adverse outcome in at least half of the affected pregnancies. Given the similar programme platform, mode of transmission and interventions, initiatives for elimination of mother-to-child transmission (EMTCT) of syphilis have been launched as dual EMTCT efforts with HIV in Africa, the Americas and the Asia Pacific (13-16).

For these reasons, surveillance of syphilis should be incorporated into HIV surveillance among pregnant women attending ANC based on routine programme data. Core indicators for EMTCT of syphilis include proportion of pregnant women tested, syphilis treponemal or non-treponemal seropositivity in ANC attendees, and proportion of positive pregnant women treated appropriately (17). These variables are typically available in the same routine site data tools (e.g. ANC registers) that may serve as the data source for HIV surveillance among pregnant women attending ANC based on routine programme data.

Further information on syphilis surveillance in ANC is given in WHO's 2011 *Methods for surveillance and monitoring of congenital syphilis within existing systems* (17).

5. CENSUS OF ANC

In this section

- 5.1 Approaches to a census of ANC
- 5.2 Site selection for a census of ANC
- 5.3 Duration of surveillance period for a census of ANC
- 5.4 Variable considerations for a census of ANC
- 5.5 Eligibility criteria for a census of ANC
- 5.6 Data collection for a census of ANC
- 5.7 Data management for a census of ANC
- 5.8 Analysis of data from a census of ANC

In a census approach, routinely generated individual-level or aggregate data from all pregnant women attending all (or nearly all) ANC are accessed at an above-site (e.g. regional or national) data repository. The census approach has several advantages:

- census results are direct measures of parameters in the population; no sampling or sample-size calculations are used – results are not estimates and contain no statistical uncertainty;
- a census requires a low amount of time and resources (because it involves no site-level fieldwork) and is methodologically simple (because it involves no sampling); and
- the large volume of data collected by a census can enable more granular subanalyses, such as direct measures of HIV at lower geographical levels (e.g. districts) or within demographic categories (e.g. age) at those levels.

5.1 Approaches to a census of ANC

A census can be based on either individual-level or aggregate data, as described in the following sections.

5.1.1 Census based on individual-level data

In a census approach based on individual-level data, individual-level data from all (or nearly all) ANC are accessed retrospectively at an above-site data repository. Such an approach is only feasible in the presence of EMR systems to capture individual-level ANC

data, and electronic data systems to move large volumes of EMR data to a data repository and store them in an accessible fashion. Although the coverage of EMR systems is currently limited in ANC settings, these systems have the potential to efficiently, accurately and inexpensively provide routine individual-level data for surveillance.

A census based on individual-level data is the preferred approach because it allows richer analyses, greater modeling opportunities and increased ability to monitor data quality. However, this method has some important additional requirements:

- HIV testing is available in all, or nearly all, ANC;
- there is high coverage of EMR systems in ANC;
- individual-level EMR data are amassed and accessible at a data repository (e.g. regional or national), and such data have a verified high degree of accuracy, completeness and integrity;
- there are adequate security and confidentiality protections for individual-level electronic data accessed for surveillance; and
- there is informatics capacity to manage and analyse large quantities of individual-level electronic data and, potentially, multiple EMR systems.

5.1.2 Census based on aggregate data

In a census approach based on aggregate data, aggregate service counts (e.g. number of pregnant women making their first ANC visit, number HIV tested, number tested HIV-positive) routinely reported from all (or nearly all) ANC sites providing PMTCT services to the national HIV programme are accessed at an above-site data repository. This approach typically leverages data aggregated and reported by an HMIS/routine reporting system, and does not require EMR systems.

This method has some limitations and additional requirements:

- HIV testing must be available in all, or nearly all, ANC.
- **Census based on aggregate data requires an HMIS/routine reporting system with a verified high degree of data accuracy, completeness and integrity.** International standards and tools for data-quality assessment are available from the Global Fund to Fight AIDS, Tuberculosis and Malaria and The President's Emergency Plan for AIDS Relief (18, 19).
- Aggregate data may preclude subanalysis of HIV or syphilis prevalence by core surveillance variables (e.g. by age) if aggregate data are not routinely reported by these subcategories.
- If possible, aggregate data on HIV status should be broken down into 5-year age categories to allow for more granular surveillance analyses and modelling. If 5-year

age categories are not available, aggregate data on HIV status should, if possible, be available in age categories of ≤ 20 and > 20 years.

- The use of aggregate programme data may pose challenges to monitoring the quality of routine programme data, because routine HIV testing and data-quality challenges are more difficult to identify when examining aggregate data. Hence, surveillance programmes should take extra measures to monitor the quality of surveillance based on aggregate census data (Section 7.3).

5.2 Site selection for a census of ANC

By definition, HIV surveillance based on a census of ANC should collect data on all pregnant women from all ANC in a country. However, some ANC sites may not be ready to participate in census, because of the following issues:

- challenges specific to a census based on individual-level electronic data could include:
 - lack of EMR systems;
 - the quality (completeness and integrity) of EMR data being inadequate for surveillance;
 - EMR data being of high quality but not accessible;
- challenges specific to a census based on aggregate data from HMIS/routine reporting systems could include:
 - low quality or reliability of aggregate data due to challenges in site-level data aggregation or the communication and management of data at district, regional or national levels;
 - inaccessibility of aggregate data (in particular, private or other nongovernmental ANC sites may not report to the national system);
- challenges common to a census based on either individual-level or aggregate data include:
 - routine HIV testing being not continuously available or uptake of HIV testing among pregnant women being substandard; and
 - routine HIV or syphilis testing quality being substandard, or quality assurance of routine HIV or syphilis testing being substandard.

It is acceptable to exclude a small proportion of ANC sites from a census based on the reasons above. The interpretation and dissemination of surveillance results should take into account, and present information about, the proportion and distribution of ANC sites that are not included in the census, and the limitations this imposes on surveillance results. However, if a substantial minority of ANC sites are not ready to participate in a census, the surveillance programme should consider a sentinel surveillance design.

5.3 Duration of surveillance period for a census of ANC

For a census of ANC, the duration of the surveillance period (i.e. the period from which ANC data are collected) should be a full calendar year.

5.4 Variable considerations for a census of ANC

A census of ANC should consider the same core and optional surveillance variables described in Section 4, including syphilis variables. At a minimum, a census should collect:

- date of first ANC visit;
- age (or age categories);
- uptake of HIV testing (received HIV testing, declined HIV testing or did not receive testing because pregnant woman was previously known to be HIV-positive);
- HIV status (newly tested positive, newly tested negative, previously known positive, result not documented);
- syphilis test done, syphilis test result and syphilis treatment; and
- ART status.

If a census is based on aggregate data:

- **aggregate data should also be broken down by HIV status options (newly tested positive, newly tested negative, previously known positive, result not documented), rather than just positive or negative; and**
- **if possible, aggregate data on HIV status should be available in 5-year age categories, to allow for granular surveillance analyses and modelling using UNAIDS' Spectrum. If 5-year age categories are not available then if possible to collect information for < 20, 20-24 and > 25 and if not possible, collect data in age categories of <20 and >20 years.**

If aggregate data are not available by these age and HIV status categories, surveillance programmes should consider a census based on individual-level data or a sentinel surveillance design.

5.5 Eligibility criteria for a census of ANC

Surveillance programmes should clearly define eligibility criteria for pregnant women attending ANC whose routine data will be collected for surveillance. Eligibility should be restricted to the recommended criteria outlined below, which surveillance programmes should modify according to local context.

Inclusion criteria:

- women with a confirmed pregnancy;
- aged 15–49 years; and
- making their first ANC visit during their current pregnancy (this avoids duplicate sampling).

Exclusion criteria:

- pregnant women who have made a previous ANC visit during their current pregnancy to any ANC; and
- pregnant women aged ≤ 14 years or ≥ 50 years.

Pregnant women meeting eligibility criteria are considered eligible irrespective of their HIV status, ART status and whether or not they received an HIV test at their first ANC visit.

The format of census data may affect the application of these eligibility criteria:

- *Pregnant women aged 15–49 years:* If a census of ANC is based on aggregate data, HMIS/routine reporting systems may use age categories that do not align with the age criteria historically used for surveillance. In such cases, surveillance programmes should take into account any changes in age eligibility criteria when surveillance data are analysed and results are disseminated.
- *Pregnant women making their first ANC visit during their current pregnancy:* Maintaining this criterion is essential to exclude double counting. Census data must be able to exclude pregnant women making their second, third or fourth ANC visit. Surveillance programmes should work with the managers of EMR and HMIS/routine reporting systems to ensure that surveillance data only include pregnant women who meet this eligibility criterion. If existing EMR or HMIS/routine reporting systems at ANC sites are not able to reliably identify pregnant women who meet this criterion, the programme may not be ready to use the census design.

5.6 Data collection for a census of ANC

This section describes methods, requirements and roles for surveillance data collection based on a census of ANC.

5.6.1 Data transfer

Data collection for a census involves no site-level fieldwork, but requires close collaboration with partners to identify and access national PMTCT data. Surveillance programmes will need to negotiate the transfer of individual-level records or aggregate counts from the Managers of EMR or HMIS/routine reporting systems.

The surveillance data manager should work with the Managers of EMR or HMIS to identify and extract appropriate data from those systems, including:

- understanding the variable definitions and formats in EMR or HMIS databases (see Section 5.6.2 for further information on data standards);
- identifying and extracting data relevant to the surveillance period;
- identifying and extracting only those variables needed for surveillance; and
- if possible, identifying and extracting only records of pregnant women who meet surveillance eligibility criteria; further work to identify eligible pregnant women may need to be done by the surveillance data manager and statistician.

In some settings, personal identifying information may be temporarily required to exclude duplicate records or to link records from multiple routine data systems to make up the complete surveillance dataset. In such cases, this information should be permanently deleted from the dataset after its data-collection purpose has been fulfilled.

5.6.2 Data standards

Data standards provide a common understanding of the core data elements and data formats needed to facilitate the collection, exchange, integration and retrieval of electronic health information. They define how information is packaged and communicated from one party to another. Adoption of national data standards (which is typically achieved via adoption of a national data dictionary) should be based on internationally recognized health data standards, such as Health Level 7. Adoption of such standards is needed to support individual-level EMR or aggregate HMIS/routine reporting system data for surveillance (20). The goal of data standards is to ensure that all census data are consistently understood and can be exchanged, in a harmonized format, regardless of the source.

Data standards should clearly define:

- the variables;
- the format in which variables are encoded (e.g. date is often stored in a DD/MM/YY format); and
- data transmission (i.e. computer-mediated communication among electronic system users, and between electronic systems); to ensure and optimize information transmission from one computer to another, data must be encoded into certain formats using data transmission standards.

There are two ways in which data standards can help to facilitate the collection and use of census data:

- *Programme use versus surveillance use:* EMR and HMIS/routine reporting systems are designed for programme and programme monitoring purposes, and store data in a way that meets programme and programme monitoring needs. However, surveillance data needs may be different in terms of variable definitions and data formats.

- **Multiple data sources:** Different ANC settings may use different EMR or HMIS/routine reporting systems, and variable definitions and formats may differ between them. Collaboration may be needed to harmonize EMR datasets to allow them to be combined and analysed together. Data standards can assist with this mapping process.

For both these reasons, it is important to collaborate with all EMR and HMIS Managers to adopt data standards that meet the needs of all partners, including surveillance.

Further information about electronic data standards is available in *Framework and standards for country health information systems*, published in 2008 by the WHO Health Metrics Network, and in the *Joint Inter-Ministerial Policy Dialogue on eHealth Standardization and Second WHO Forum on eHealth Standardization and Interoperability*, published in 2014 by WHO (21, 22).

5.6.3 Staffing for data collection and management

Required personnel for census data collection include data managers, statisticians and surveillance coordinators. Responsibilities for all surveillance staff members should be clearly defined in the surveillance protocol. Table 2 outlines responsibilities for surveillance staff for surveillance based on a census of ANC.

Table 2. Responsibilities for surveillance staff for a census of ANC

Staff	Responsibilities
Data manager	<ul style="list-style-type: none"> • Work with partners to arrange transfer of EMR or HMIS/routine reporting system data • Ensure security of surveillance data during transfer, storage and analysis • Ensure confidentiality measures • Clean surveillance data • Manage surveillance database
Statistician	<ul style="list-style-type: none"> • Support the development and implementation of surveillance protocol, especially surveillance design, data analysis and surveillance monitoring • Ensure security of surveillance data during transfer, storage and analysis • Clean and analyse surveillance data • Support the development of dissemination materials
Surveillance investigator or surveillance technical working group	<ul style="list-style-type: none"> • Ensure adequate funding for surveillance • Liaise with partners and build organizational and political support for surveillance • Develop surveillance protocol • Provide oversight of surveillance implementation, data analysis and surveillance monitoring • Provide oversight of data security and confidentiality measures • Interpret surveillance findings and develop dissemination materials

ANC, antenatal clinics; EMR, electronic medical record; HMIS, health management information system

5.6.4 Collecting routine records with missing data

HIV surveillance based on a census of ANC is likely to include routine records of eligible pregnant women with missing data. Such data may include core surveillance variables (e.g. age and HIV status) or non-core surveillance variables (e.g. gravidity and occupation).

Routine records of eligible pregnant women that include missing data should be collected. These records provide important information for monitoring routine data quality, and may be useful for certain analyses (Sections 7.1–7.3).

5.7 Data management for a census of ANC

A robust and comprehensive set of data management practices is essential to ensure the integrity of surveillance data, the reliability of surveillance measures and the confidentiality of pregnant women whose routine data are collected for surveillance.

Surveillance based on a census of ANC data may require significant informatics capacity to manage, merge, clean and manipulate large quantities of individual-level electronic data and, potentially, multiple EMR systems.

5.7.1 Data security

Comprehensive data security measures include human, physical and electronic procedures and protections at every stage of the surveillance activity, including data collection, transfer, storage and analysis:

- All electronic devices (e.g. computers, personal digital assistants [PDAs], tablets and USB drives) should be physically secured when not in use.
- All electronic devices and databases (including EMR or HMIS databases) should be encrypted, password protected and accessible only by appropriate surveillance staff.
- The transfer of data from EMR or electronic HMIS to the surveillance programme should be accomplished through a secure medium.
- The transfer of individual-level EMR data, or aggregate HMIS data, should be limited to variables required for surveillance, excluding any directly identifiable personal information.
- All surveillance databases should be regularly backed up on a secure, external system and the backups must be secured.
- Electronic surveillance data should only be shared with non-surveillance staff where appropriate and in accordance with national data-sharing policies.
- Surveillance databases should be securely preserved for future analysis.

- All staff with access to surveillance data must receive training in data security and confidentiality, and must individually affirm that they will abide by data security and confidentiality principles and procedures.

Further guidance on data security is provided in the 2006 UNAIDS publication *Guidelines on protecting the confidentiality and security of HIV information* (23).

5.7.2 Data cleaning

A census of ANC poses unique challenges for assuring a clean and reliable surveillance dataset.

For a census based on individual-level data, standard data-cleaning procedures should be followed:

- Frequency tables should be regularly produced for each ANC site and surveillance variable, to identify data quality issues.
- Missing, inconsistent or invalid values should be investigated. These gaps could be inherent in the source data (the EMR dataset) or have been produced during the extraction, transfer or merging of EMR datasets.
- Sometimes a single country will use multiple EMR systems. If datasets from different EMR systems are collected and merged, it is essential to identify instances where variable formats and definitions are not harmonized across EMR systems. For example, one EMR system may code the date as MM/DD/YYYY, whereas another may code it as MM/DD/YY. Such instances should be resolved to ensure a consistent and analysable dataset. In some cases, it may not be possible to completely harmonize variables. For example, one EMR system may code HIV status as "Known positive", "R" and "NR", whereas another may code it as simply "Pos" and "Neg". In this example, the second EMR system has no value that is equivalent to the "Known positive" value in the first system. Such instances should be noted as a limitation of the analysis. Establishing data standards (Section 5.6.2) can facilitate the harmonization of variable formats and definitions.
- All corrections to the dataset should be documented.

For a census based on aggregate data, data cleaning should attempt to follow the same procedures as described above. However, the use of aggregate data (i.e. counts) may make it more difficult to identify and correct data quality issues. Additional measures and considerations for cleaning aggregate census data include the following:

- Examine site-level count totals to look for inconsistencies between variables. All variable counts should add up to the same number of pregnant women making their first ANC visit; if they do not, this suggests a data-quality problem. For example:

- If a site reported that 100 pregnant women made their first ANC visit, the HIV status count (known positive, newly tested negative or newly tested positive) only adds up to 90, this suggests that 10 women were not tested or that data are missing.
- If a site reported that 100 pregnant women made their first ANC visit, but the total of the age categories count is 110, this suggests that one of the two figures is incorrect.
- Examine site-level data counts to look for implausible values. For example, if the number of pregnant women making their first ANC visit at a site fluctuates dramatically from month to month, this may indicate a data-quality problem.
- All corrections to the dataset should be documented.

Where data-quality issues are identified, surveillance programmes will have to decide whether the problem is acceptably minor that the site can be retained, or sufficiently large that the site should be excluded from surveillance. All sites with data-quality issues, whether retained or excluded, should be documented as a limitation of surveillance.

5.8 Analysis of data from a census of ANC

This section presents methods for analysing data from a census of ANC. Although it presents methods for calculating HIV measures, these same methods can also be applied to calculating syphilis measures.

5.8.1 Calculating prevalence

To obtain a prevalence estimate using data from a census of ANC, the calculation is as follows:

$$\rho = \frac{\chi}{\eta} \times 100$$

where:

ρ = prevalence; this is a direct measure of the prevalence in the population, not an estimate;

χ = the number of pregnant women who are HIV-positive, including pregnant women who newly tested positive or were previously known to be HIV-positive; and

η = the number of pregnant women with known HIV status, including pregnant women who newly tested positive, newly tested negative, or were previously known to be HIV-positive.

This calculation is valid for national and all subnational levels. Because census data are not a sample, and are direct measurements rather than estimates, they require no weighting or calculation of variance (i.e. confidence intervals). In addition, census data from different sites or geographical areas can be combined to provide estimates of larger geographical areas.

5.8.2 Examining trends in HIV prevalence over time

HIV prevalence trends from a census of ANC can be examined at the national and all subnational levels if surveillance has maintained the same methods and sites. Plotting HIV prevalence over time for specific demographic groups (especially age groups) will help in identifying specific changes in levels of HIV infection. The most effective way to examine trend data is graphically using scatter plots or line graphs, where prevalence measures are plotted on the vertical axis and time is plotted on the horizontal axis. It is preferable to have at least three rounds of data when presenting trends graphically.

5.8.3 Estimating national prevalence

Data from HIV surveillance based on a census of ANC should not be used to estimate HIV prevalence beyond the population sampled by surveillance (pregnant women attending selected ANC sites) (Section 8.1). Estimates of HIV prevalence and trends in the general population using data from HIV surveillance among pregnant women attending ANC should be made using UNAIDS' Spectrum analysis tools (6).

5.8.4 Analysis of routine records with missing data

HIV surveillance based on a census of ANC is likely to collect some routine records of eligible pregnant women with missing data. Such data may include core surveillance variables or non-core surveillance variables (Section 4). The surveillance programme should define a consistent approach to handling records with missing data for each facet of analysis.

A census of ANC based on individual-level data provides the value of each variable for each pregnant woman; this facilitates the identification of missing data, so that decisions about the inclusion or exclusion of these data can be made. It is suggested that:

- records with missing core surveillance variables cannot provide basic epidemiological data, and therefore should be excluded from all analyses; core variables are listed in Section 4; and
- records with all core variables available but with missing non-core variables (e.g. marital status and ART status) can still provide basic epidemiological data, and therefore should be retained for surveillance analyses; however, these records may need to be excluded when conducting surveillance analyses specific to non-core variables.

Census of ANC based on aggregate data only provides counts for each variable, which makes identifying and quantifying missing data more challenging. However, aggregate data can be examined to identify inconsistencies between variable counts, as described in Section 5.7.2. It is suggested that:

- surveillance programmes consider excluding ANC sites where there are substantial inconsistencies in counts among core variables; and

- ANC sites with substantial inconsistencies in counts among only non-core variables be retained for core surveillance analyses, but may need to be excluded when conducting surveillance analyses specific to non-core variables.

The interpretation and dissemination of surveillance results should take into account and present information about missing data, and the limitations such data impose on surveillance results.



6. SENTINEL SURVEILLANCE DESIGN

In this section

- 6.1 Selection of ANC sites in sentinel surveillance
- 6.2 Duration of the sentinel surveillance period
- 6.3 Frequency of sentinel surveillance rounds
- 6.4 Sample-size for sentinel surveillance
- 6.5 Eligibility criteria for sentinel surveillance
- 6.6 Data collection for sentinel surveillance
- 6.7 Data management for sentinel surveillance
- 6.8 Data analysis for sentinel surveillance

This section presents methods for sentinel surveillance, an approach that historically has been widely used for HIV surveillance among pregnant women attending ANC.

6.1 Selection of ANC sites in sentinel surveillance

In sentinel surveillance, ANC sites are selected for surveillance based on convenience sampling. Because convenience sampling is a nonprobability design, it cannot provide statistically representative estimates of HIV among ANC attendees at national or subnational levels. Sentinel surveillance can only produce valid site-level HIV estimates, which can be summarized across sites using summary measures (e.g. medians of site estimates). Results obtained from sentinel ANC sites represent only the populations of pregnant women attending those sites.

Sentinel surveillance selects ANC sites based on two factors. First, it selects on the practicality of including the site in surveillance. Practicality considerations include:

- physical accessibility of the site;
- capacity of (and potential for collaboration with) site staff and supervisors;
- Sufficient volume of ANC clients to each sample-size requirements; and
- site performance: the demographic and HIV serostatus data needed for surveillance are accurate, complete and accessible in routine records and the quality of routine HIV testing is assured (Sections 7.1 and 7.4).

Second, sentinel surveillance selects ANC sites based on characteristics that the national surveillance programme would like represented in surveillance. These characteristics could include:

- setting: urban, periurban (if relevant) and rural populations;
- geographical areas of the country (e.g. region, province or district);
- geographical areas of special interest for HIV surveillance (e.g. border areas, transport corridors, areas where certain industries are concentrated and areas containing displaced persons); and
- populations of different socioeconomic status; clinics in the private or nongovernmental sector should be considered for selection, especially if they cater to a significantly different socioeconomic population of pregnant women.

The number of ANC sites selected by sentinel surveillance will depend largely on surveillance needs, the size and heterogeneity of the country, the surveillance programme's human and financial resources, the number of sites that the surveillance programme can effectively manage, and the characteristics the surveillance programme would like represented. At a minimum, one to two ANC sites (one urban and one rural) should be selected per region, province or district, depending on surveillance needs and geographical coverage. If too many ANC sites are selected and adequate monitoring cannot be ensured, the quality of surveillance operations and data may be compromised. Selecting too few sites will result in limited epidemiologic information.

In rural or other settings where population density is low, client volume at ANC sites may be too low to reach sample-size requirements during a reasonable period of time. In this context, surveillance programmes may need to select multiple ANC sites (sometimes called "satellites") to represent a locality. These sites can then be grouped together to meet sample-size requirements if the populations served by the different sites are similar. These sites should be considered as one unit for purposes of analysis.

ANC sites from previous surveillance rounds ("historical sites") should be retained, to produce reliable trend data. In this context, a historical site should only be dropped if:

- changes at the site significantly reduce the practicality of including the site;
- challenges at the site significantly reduce the quality and reliability of the site's routine programme data or HIV testing (Section 7); or
- the characteristics of the ANC site's catchment population change so substantially that the site no longer represents the characteristics it was originally selected to represent (e.g. a transport corridor stops being a transport corridor).

6.2 Duration of the sentinel surveillance period

The surveillance period is the defined period of time from which data will be collected. If data collection is undertaken after the surveillance period, this should happen as soon as possible after the surveillance period, to provide up-to-date epidemiologic information. In sentinel surveillance, the duration of the surveillance period will be determined by the time required to reach the desired sample-size, and therefore by the volume of ANC attendees at surveillance sites.

6.3 Frequency of sentinel surveillance rounds

The frequency of surveillance among pregnant women should take into account the data needs served by surveillance as well as resource considerations. HIV prevalence trend data from HIV surveillance among pregnant women attending ANC is important for countries to monitor the course of their HIV epidemics and to allocate resources for, and plan and evaluate, HIV control and prevention interventions. To meet these data needs, surveillance among pregnant women should be conducted at regular intervals (about every 2 years). Because it is unlikely that a measurable change in the epidemic could be observed in 1 year, yearly or continuous surveillance among pregnant women may not produce additional value that justifies the required resources.

6.4 Sample-size for sentinel surveillance

This section provides guidance on site-level sample-size calculations for sentinel surveillance. The sample-size is determined by the desired precision of prevalence estimates and the need to detect changes in prevalence over time. Good prevalence estimates maximize precision for a given cost. Sample-sizes that are too small or too large waste resources. Samples that are too small provide estimates that are too imprecise to be useful, whereas those that are too large provide more precision than is actually needed.

6.4.1 Sample-size for a single prevalence proportion

Sample-size requirements to obtain a single prevalence estimate for a specified level of precision at the site level can be obtained from published tables or by using statistical software. Table 3 presents sample-size requirements – expressed as the number of individual pregnant women – for random sampling based on observed prevalence and desired precision of the prevalence estimate. Precision is commonly measured by the margin of error of an estimate, which is defined as half the width of a 95% confidence interval. The narrower the confidence interval, the more precise and reliable the estimated prevalence will be in describing the true prevalence. Appendix A1 provides detailed instructions on calculating sample-size requirements for a single prevalence.

Table 3. Sample size requirements by observed HIV prevalence and relative half-width of 95% confidence interval, assuming simple random sampling

Observed HIV prevalence (%)	Relative half-width ^a of 95% confidence interval (CI)									
	15%	20%	25%	30%	35%	40%	50%	60%	70%	80%
1	16 903	9508	6085	4226	3105	2377	1521	1057	777	595
2	8366	4706	3012	2092	1537	1177	753	523	385	295
3	5521	3105	1988	1380	1014	777	497	346	254	195
4	4098	2305	1475	1025	753	576	369	257	189	145
5	3244	1825	1168	811	596	456	292	203	149	115
6	2675	1505	963	669	492	376	241	168	123	95
7	2269	1276	817	567	417	319	204	142	105	80
8	1964	1105	707	491	361	276	177	123	91	70
9	1727	971	622	432	317	243	156	108	80	61
10	1537	865	553	384	282	216	139	97	71	55
11	1382	777	498	346	254	195	125	87	64	49
12	1252	705	451	313	230	176	113	79	58	45
13	1143	643	412	286	210	161	103	72	53	41
14	1049	590	378	262	193	148	95	66	49	37
15	968	544	349	242	178	136	87	61	45	35
16	897	504	323	224	165	126	81	57	42	32
17	834	469	300	209	153	117	75	53	39	30
18	778	438	280	195	143	110	70	49	36	28
19	728	410	262	182	134	103	66	46	34	26
20	683	384	246	171	126	96	62	43	32	25
25	512	288	185	128	94	72	46	33	24	19
30	399	224	144	100	73	56	36	25	19	15
40	256	144	92	64	47	36	23	17	12	10

^a The relative half-width is half the width of the confidence interval as a percentage of the observed HIV prevalence. For example, if the prevalence is 10% and the relative half-width is 20%, then the half-width of the confidence interval will be $10\% \times 20\% = 2\%$. The 95% confidence interval will therefore be 8–12%.

Accounting for expected missing data

Sample-size calculations should be adjusted for expected missing data. It may be anticipated that a certain small proportion of routine records will have core surveillance variables missing, rendering these records unusable (Section 4.1). For this reason, sample-size calculations should be adjusted upwards to account for potential missing data. To do this, the unadjusted sample-size is divided by the proportion of records expected to have complete data (i.e. to have no missing data).

$$\frac{\text{unadjusted sample-size}}{\text{proportion of records expected to be complete}} = \text{sample-size adjusted for expected missing data}$$

For example, if the unadjusted sample-size is 500, and the proportion of records expected to be complete is 95%, then the adjusted sample-size is 526:

$$\frac{500}{0.95} = 526$$

6.4.2 Sample-size for detecting changes in HIV prevalence

Detecting statistically significant changes in HIV prevalence over time requires larger sample-sizes than estimating prevalence at a single point in time. The smaller the change in prevalence over time, the larger the sample-size required to detect a change. Therefore, surveillance programmes that wish to detect changes in prevalence should decide on the magnitude of change in prevalence that will be detectable (e.g. 10% change or 50% change), the level at which a change in prevalence will be measured (e.g. national or regional), and whether resources allow for the increased sample-size that may be needed. In concentrated or lower level epidemics, sample-size to detect changes in prevalence at subnational levels could be very large. Calculations for sample-size to detect changes in HIV prevalence are further described in Appendix A2.

6.4.3 Sample-size considerations for young pregnant women

The “know your epidemic” approach of second-generation HIV surveillance recommends that countries track HIV incidence to understand the direction of the epidemic and where new infections are arising, both nationally and for subnational geographical areas (4). Because direct measurement of HIV incidence through prospective cohort methods is prohibitively costly, alternative and proxy methods have been developed to estimate trends in HIV incidence.

Pregnant women aged 15–24 years attending ANC are likely to have only recently experienced sexual debut; thus, infections in this population are likely to be recent. Trends in HIV prevalence among young pregnant women do not directly mirror trends in HIV incidence because there are other factors that complicate how HIV incidence is reflected in ANC prevalence amount young women (Section 8.2.2). However, HIV prevalence trends

among young pregnant women can be valuable as one of multiple data points used to triangulate trends in HIV incidence.

Because sentinel surveillance is a nonprobability design, data from different sites cannot be pooled to provide estimates at the national or subnational level. Sentinel surveillance can only produce valid site-level HIV estimates among young pregnant women, which can be summarized across sites using summary measures (e.g. medians of site estimates). If surveillance programmes want HIV prevalence estimates of a specified precision among young pregnant women at the site level, sample-size calculations should be adjusted to ensure that sufficient young pregnant women will be sampled. There are two options to adjust sample-size: inflation and oversampling. These methods are described in detail in Appendix A3.

6.5 Eligibility criteria for sentinel surveillance

Surveillance programmes should clearly define eligibility criteria for pregnant women attending ANC whose routine data will be collected for surveillance. Eligibility should be restricted to the recommended criteria given below, which surveillance programmes should modify according to local context.

Inclusion criteria:

- women with a confirmed pregnancy;
- aged 15–49 years; and
- making their first ANC visit during their current pregnancy (this avoids duplicate sampling).

Exclusion criteria:

- pregnant women who have made a previous ANC visit during their current pregnancy to any ANC; and
- pregnant women aged ≤ 14 years or ≥ 50 years.

Pregnant women meeting eligibility criteria are considered eligible irrespective of their HIV status, ART status and whether or not they received an HIV test at their first ANC visit.

6.6 Data collection for sentinel surveillance

When developing data-collection methods for HIV surveillance among pregnant women attending ANC based on routine programme data, surveillance programmes should consider approaches that are:

- economical given available resources;
- appropriate for routine site data tools and data quality at ANC;

- sustainable and replicable over time; and
- of sufficiently high methodological rigour to produce reliable HIV estimates with a minimum of bias.

There are two potential methods for data collection:

- *Retrospective*, in which surveillance staff identify eligible pregnant women and abstract individual-level surveillance data from routine site data tools (e.g. registers) after the defined surveillance period. This approach to data collection requires a lower level of resources, but offers little control over the quality of routine programme data or HIV testing.
- *Real time*, in which ANC site staff identify eligible pregnant women and collect individual-level surveillance data from routine site data tools (e.g. registers or maternal cards) during the surveillance period as pregnant women visit the site. This is a more resource-intensive approach, but it can help to support the quality of routine programme data or HIV testing where performance of ANC sites is uneven.

These methods are compared in Table 4.

Table 4. Comparison of data-collection methods for HIV sentinel surveillance among pregnant women attending ANC

Method	Advantages	Disadvantages
Retrospective	<ul style="list-style-type: none"> • Small number of data-collection teams, each visiting multiple sites: <ul style="list-style-type: none"> – lower level of human and financial resources needed – fewer personnel need training, monitoring and quality assurance 	<ul style="list-style-type: none"> • Limited control over quality of routine inputs: routine programme data have already been recorded and routine HIV testing has already been done before surveillance involvement
Real time	<ul style="list-style-type: none"> • Increased control over quality of surveillance inputs: surveillance staff may monitor and support the quality of routine programme data and HIV testing in real time 	<ul style="list-style-type: none"> • Higher level of human and financial resources needed: training and monitoring of ANC site staff at all surveillance sites • Need to monitor and assure quality of data-collection methods at every site over the course of the surveillance period

ANC, antenatal clinic; HIV, human immunodeficiency virus

6.6.1 Defining the routine data source

Variables needed for surveillance (Section 4) are often found in multiple routine data tools at ANC (e.g. ANC registers, PMTCT registers, laboratory registers, and HIV counselling and testing registers). Before starting data-collection activities, surveillance programmes should define which routine data tool or tools will be the source of surveillance data. The programmes should select a routine site data tool that is accessible, comprehensive (i.e. contains all of the variables needed for surveillance), reliable and complete. In some cases, programmes may need to define two routine site data tools (that can be linked through a unique identifier) as the source for surveillance data, if no single data tool contains all the variables needed for surveillance. For example, the ANC register may serve as the source for demographic data and the laboratory register may serve as the source for HIV-testing data. However, use of multiple routine site data sources should be avoided if possible because it increases the duration and complexity of surveillance field operations. To ensure standardized methods and comparability of data across sites, the same routine data sources should be used at all sites.

WHO guidance to countries on standardized, comprehensive formats for ANC and PMTCT data tools can be found in the 2012 publication *Three interlinked patient monitoring systems for HIV care/ART, MCH/PMTCT (including malaria prevention during pregnancy), and TB/HIV* (24).

6.6.2 Defining a surveillance code

All eligible pregnant women whose records are sampled by surveillance should be assigned a surveillance code to uniquely identify the record, and facilitate data cleaning and management. This code should contain information about the surveillance record and sampling. A simple format for a surveillance code is a site code followed by an ordinal number. Sites codes should be unique; for example:

“MES-155”

The first three letters indicate that the record is from Messina site, which has the site code “MES”, and the “155” indicates that it is the 155th record sampled from that site.

The surveillance code should not contain any personal identifying information (e.g. ANC patient code) that could be used to connect surveillance data to individual pregnant women.

6.6.3 Retrospective data-collection methods

6.6.3.1 Staffing and training for retrospective data collection

Required personnel for retrospective data collection include data collectors, supervisory staff, data managers or statisticians, and surveillance coordinators. Responsibilities for all surveillance staff members should be clearly defined in the surveillance protocol. Table 5 outlines responsibilities for surveillance staff for retrospective data collection.

Table 5. Responsibilities for surveillance staff for retrospective data collection

Staff	Responsibilities
Data collectors	<ul style="list-style-type: none"> • Collect surveillance data according to field protocol • Ensure confidentiality of pregnant women whose routine data are collected for surveillance • Ensure security of surveillance data during collection
Field coordinators	<ul style="list-style-type: none"> • Supervise and support data collectors • Ensure that data collection is implemented according to field protocol • Conduct quality assurance of data-collection procedures • Ensure confidentiality of pregnant women whose routine data are collected for surveillance, including the permanent delinking of surveillance data from any personal identifying information • Ensure security of surveillance data during collection and transit • Liaise with surveillance coordinator
Surveillance coordinator	<ul style="list-style-type: none"> • Coordinate implementation, quality assurance and monitoring of surveillance field activities • Supervise and support field data-collection teams • Organize staff trainings • Develop field protocols or standard operating procedures • Ensure provision of equipment, supplies and data tools • Ensure confidentiality of pregnant women whose data are collected for surveillance • Ensure security of surveillance data during transit, storage, analysis and dissemination
Data manager	<ul style="list-style-type: none"> • Ensure security of surveillance data and metadata (e.g. sampling frames, survey weights) during transit, storage and analysis • Arrange for quality-assured data entry • Clean surveillance data • Manage surveillance database
Statistician	<ul style="list-style-type: none"> • Support the development and implementation of surveillance protocol, especially surveillance design, sample-size calculation, data analysis and surveillance monitoring • Ensure security of surveillance data during storage and analysis • Support the development of dissemination materials

Staff	Responsibilities
Surveillance investigator or surveillance technical working group	<ul style="list-style-type: none"> • Ensure adequate funding for surveillance • Liaise with partners and build organizational and political support for surveillance • Develop surveillance protocol • Provide oversight of surveillance implementation, data analysis and surveillance monitoring • Provide oversight of data security and confidentiality measures • Interpret surveillance findings and develop dissemination materials

To conduct high-quality surveillance, all personnel involved should be trained. Training should occur before every surveillance round, and should include a review of operational procedures, staff responsibilities, field protocol or standard operating procedures, and data security and confidentiality measures. Surveillance staff should be provided with a written field protocol, standard operating procedures or other training materials for future reference. If the surveillance round involves new or updated procedures or a significant number of new personnel, it is a good idea to conduct a field test of surveillance operational procedures. A field test can identify challenges to surveillance operational procedures, and provide hands-on training to surveillance staff.

6.6.3.2 Sampling records of pregnant women in retrospective data collection

In retrospective data collection, a surveillance data-collection team visits ANC sites after the defined surveillance period has ended. The team identifies the routine site data tool or tools defined as the source for surveillance data (Section 6.6.1). The data collectors then identify records of eligible pregnant women in the designated routine site data tools and sample them for surveillance.

Consecutive sampling is a simple, effective technique for selecting records of eligible pregnant women. In this approach, the data-collection team starts with the first day of the surveillance period and consecutively samples records of all eligible pregnant women until the sample-size has been reached, at which point collection stops. The disadvantage of this strategy is that samples from higher volume clinics are more concentrated in time than those from lower volume clinics.

6.6.3.3 Methods to collect data in retrospective data collection

Surveillance programmes should define the specific method used to collect surveillance data from routine site data tools. The approach should be:

- methodologically rigorous to ensure the quality and integrity of surveillance data;
- suitable given routine programme data tools at ANC;

- easy to implement; and
- economical and sustainable given available resources.

Common data-collection options include:

- *Manual abstraction*, in which a data collector manually copies data from routine site data tools (e.g. registers) onto surveillance data-collection forms. An example of such a data-collection form is provided in Appendix A5. The method has the advantage of simplicity, but the disadvantage of the possibility of errors being introduced in the process.
- *Manual abstraction into electronic format*, in which a data collector manually enters data from routine site data tools into an electronic device, such as a PDA, tablet or laptop database (e.g. Access, Excel or Epi Info). If routine programme data are entered directly into an electronic device at the point of data collection, best practices regarding electronic data entry (Section 6.7.2) and security (Section 6.7.1) should be followed to ensure the accuracy and integrity of surveillance data. This method has the advantage of combining data abstraction and data entry into one process, as well as the ability to do real time data-quality checks (e.g. checking for missing or invalid values) and corrections (referring back to the source data) of electronic data. However, this method has the disadvantage of displacing data entry to a less controlled field environment.
- *Electronic data capture*, in which a data collector takes a digital image (photograph or scan) of routine site data tools, and returns the digital image to the central level for manual data entry.

Whatever approach is used, personal identifying information of pregnant women (e.g. names or ANC patient identifiers) should not be captured unless this information is needed for data collection methods. If routine data are captured as a digital image, personal identifying information can be excluded by physically covering this information before the digital image is captured.

Data-collection methods for retrospective data collection with multiple routine data sources

In some contexts, patient codes or personal identifying information may be temporarily required to exclude duplicate records or to link records from multiple routine data to make up the complete surveillance dataset. In such cases, personal identifying information should be permanently deleted from the dataset after its data-collection purpose has been fulfilled. For example, the primary source of surveillance data may be the ANC register. However, the test results for each individual test in the HIV diagnostic algorithm may only be available in a laboratory or HIV testing and counselling (HTC) register. The surveillance programme would need to match data from the two routine data sources to make a full record for each individual eligible pregnant woman. To do this, the surveillance programme could collect the ANC number for each record abstracted from the ANC register. This number could then be used to find the same women in the laboratory or HTC register, and

abstract the results for each individual test in the HIV diagnostic algorithm. Once this has been done, the ANC number would be removed from the surveillance data collection form.

Data collection from EMR

Retrospective data collection based on EMR systems seeks to take advantage of existing, accessible, individual-level electronic data on pregnant women attending ANC. In an EMR-based approach, the surveillance team would work with the EMR data manager to identify and extract records of eligible pregnant women visiting surveillance sites during the surveillance period. Collection of data from EMR systems implies additional considerations:

- working with electronic data may require significant informatics capacity to manage and manipulate individual-level electronic data and, potentially, multiple EMR systems;
- the extraction of individual-level EMR data should be limited to variables required for surveillance, excluding any personally identifiable information (Section 2);
- electronic data should be extracted with all due measures taken to ensure electronic data security and confidentiality (Section 6.7.1); and
- different health partners may use different EMR systems, and variable definitions and formats may differ between them; hence, significant work may be necessary to harmonize EMR datasets to allow them to be combined and analysed together (Section 5.6.2).

6.6.3.4 Quality assurance of retrospective data collection

Quality assurance during surveillance data collection is essential to ensure the reliability of surveillance estimates. It is the responsibility of data collectors and field coordinators to conduct quality assurance as part of surveillance data-collection procedures. In particular, it is important to routinely validate that:

- surveillance data are collected from appropriate routine site data tools (e.g. registers), as defined by the protocol;
- surveillance data are only collected from eligible women during the surveillance period, as defined by the protocol;
- surveillance data are accurately and completely collected from routine programme data tools;
- sampling methods and sample-size requirements for records of eligible pregnant women are adhered to, as defined by the protocol; and
- surveillance data are managed and stored in a secure fashion, as defined by the protocol.

After data collection at a site, surveillance field staff and field coordinators should validate a subset of collected records (e.g. 5–10%) against the source data. Data collection should be repeated if errors in collected data are found to exceed a certain protocol-defined threshold (e.g. 1–3% of records have errors). Errors could include incompleteness, disagreement between collected and source data, and inclusion of ineligible pregnant women.

6.6.4 Real time data-collection methods

Real time data collection may provide advantages in certain settings. In a real time approach, ANC site staff identify eligible pregnant women and collect individual-level surveillance data from routine site data tools (e.g. registers or maternal cards) in real time as pregnant women visit the site. This approach is substantially similar to the method used by UAT serosurveillance, except that no biological specimens are aliquoted for surveillance. This method has the advantage of affording increased control over surveillance inputs: routine programme data, routine HIV testing and routine syphilis testing. In a real time approach, surveillance staff may monitor and support the quality of routine programme data, routine HIV testing and routine syphilis testing in real time. Hence, this method may be appropriate for countries where the quality of routine programme data, routine HIV testing and routine syphilis testing is uneven at ANC sites. However, this method requires substantially more time and resources than retrospective data collection for the training and monitoring of ANC site staff over the duration of the surveillance period.

6.6.4.1 Staffing and training for real time data collection

The staffing and training needs for real time data collection are largely similar to those described for retrospective data collection (Section 6.6.3.1). The principal differences in the real time method are as follows:

- Data collectors are ANC staff themselves (not a surveillance field team), who routinely record data in routine site data tools (registers, maternal cards) and abstract these data for surveillance in real time as pregnant women visit ANC sites. This substantially increases the number of data collectors involved in surveillance. Implicated ANC staff at all surveillance sites should be trained and monitored with regards to identifying eligible pregnant women, collecting data from routine records of eligible pregnant women according to the protocol, and ensuring the security of surveillance data during collection and storage at the site until data are retrieved by field coordinators.
- Field coordinators are responsible for overseeing data collection by a large set of ANC staff at all surveillance sites. Supervision by field coordinators will primarily take the form of site visits.

6.6.4.2 Sampling pregnant women for real time data collection

In real time data collection, consecutive sampling of eligible pregnant women is preferred for the simplicity of field implementation. In this approach, ANC site staff sample records of eligible pregnant women in real time as they attend ANC. Sampling starts on the first day

of the surveillance period and continues until a defined sample-size is reached, at which point sampling stops.

6.6.4.3 Methods to collect data in real time data collection

As in retrospective data collection, surveillance programmes should define a real time data-collection method that is:

- methodologically rigorous to ensure the quality and integrity of surveillance data;
- suitable given routine programme data tools at ANC;
- easy to implement; and
- economical and sustainable given available resources.

In real time data collection, surveillance is conducted as pregnant women make their ANC visit. ANC staff record demographic and clinical data in site data tools as per routine practice. The staff then abstract these routine data from site data tools onto a surveillance data-collection form. An example of such a data-collection form is provided in Appendix A5.

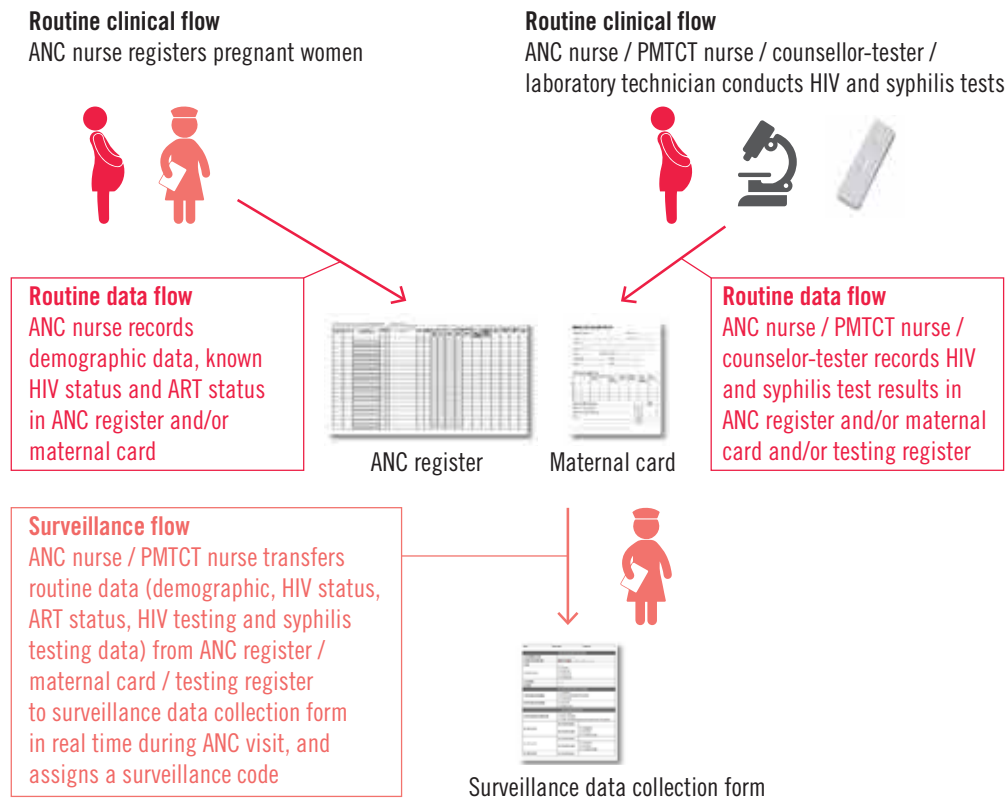
An example of a simple data-collection flow for real time data collection is shown in Fig. 2. Real time data-collection flow should:

- be designed around routine site patient flow;
- not disturb routine site patient flow; and
- enable the collection of all required surveillance variables.

Again, as in retrospective data collection, real time data collection should not capture personal identifying information of pregnant women (e.g. names or ANC patient identifiers) unless this information is temporarily needed for data collection methods. Personal identifying information collected to exclude duplicate records or to link records from multiple routine data sources should be permanently deleted after its data-collection purpose has been fulfilled.



Fig. 2. Example of a simple data-collection flow for real time data-collection



ANC, antenatal clinic; ART, antiretroviral therapy; PMTCT, preventing mother-to-child transmission

Data-collection methods for real time data collection with multiple routine data sources

If all surveillance variables are not available in a single routine data tool, the data-collection flow should be designed to collect variables from multiple routine data tools. For example, the ANC register may be the primary routine data source for surveillance data. The ANC nurse captures most surveillance data by abstracting it from the ANC register onto the surveillance data-collection form. However, results for each individual test in the HIV diagnostic algorithm may only be available in a laboratory or HTC register. To capture these data, the surveillance programme would need to devise a data-collection flow that allows the data-collection form to collect information from both the ANC register and the laboratory or HTC register for each eligible pregnant woman. In this example, there are

two options for collecting data from two sources. After most surveillance data has been abstracted from the ANC register onto the surveillance data-collection form during ANC registration:

- the surveillance data-collection form could physically follow the pregnant women or their biological specimen to the laboratory or HTC room; once there, the results for each individual test in the HIV diagnostic algorithm could be abstracted from the laboratory or HTC register onto the surveillance data-collection form; or
- a site patient code (e.g. an ANC number) could be abstracted from the ANC register onto the surveillance data-collection form; this number could then be used to find the same women in the laboratory or HTC register, and abstract the results for each individual test in the HIV diagnostic algorithm, after which the ANC number would be removed from the surveillance data-collection forms.

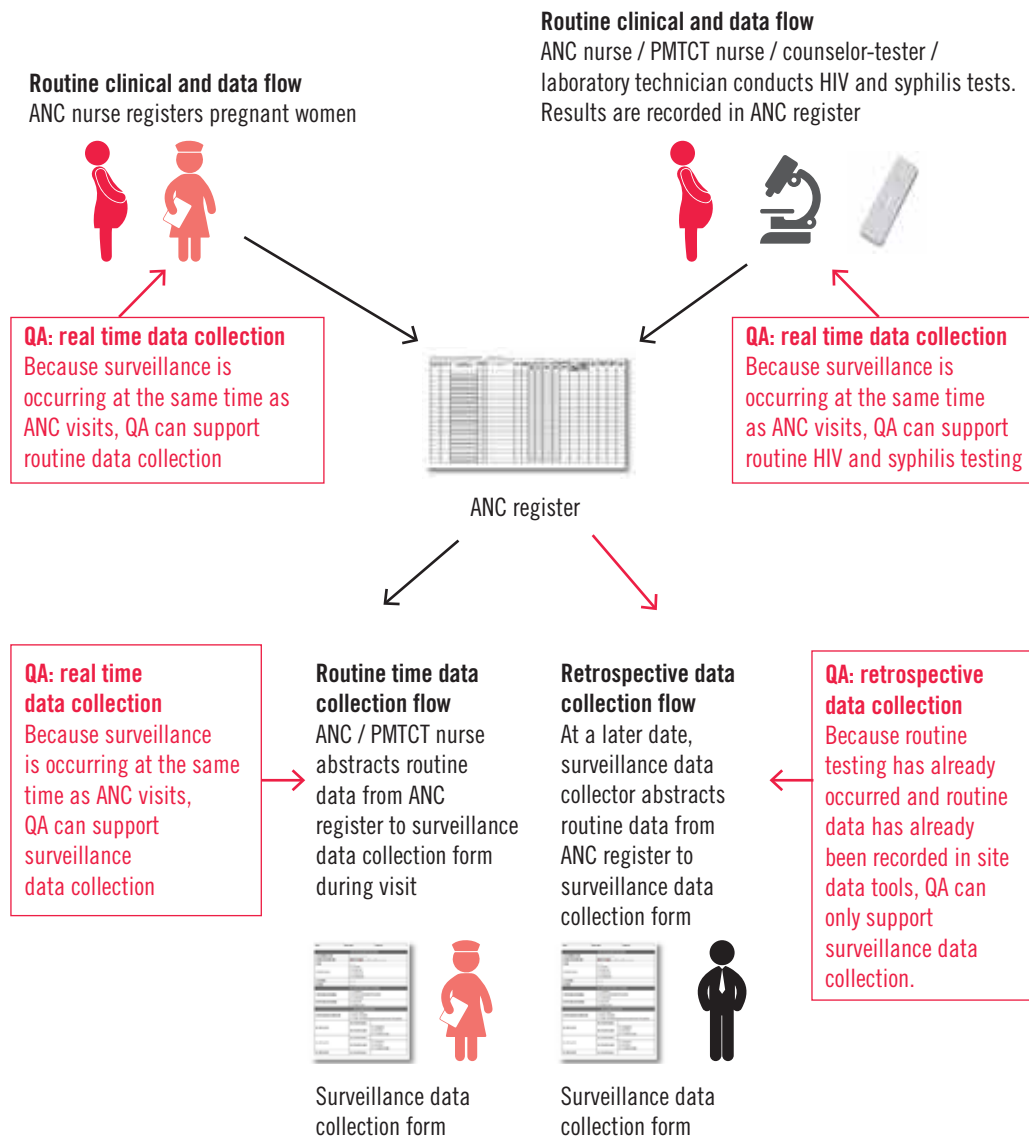
6.6.4.4 Quality assurance of real time data collection

Quality assurance methods for real time data collection are substantially similar to those described for retrospective data collection (Section 6.6.3.4), with the following important difference. For retrospective data collection, the surveillance period is in the past. There is no opportunity for the surveillance programme to support the quality of routine programme data, routine HIV testing and routine syphilis testing.

By contrast, for real time data collection, the surveillance period is ongoing. Therefore, the surveillance programme can assure the quality of (a) the surveillance team's own methods for collecting existing routine data (demographic and serostatus) from site data tools and (b) routine programme data, routine HIV testing and routine syphilis testing. This helps to ensure that routine programme data, routine HIV testing and routine syphilis testing are of high quality. Methods for monitoring the quality of routine programme data, routine HIV testing and routine syphilis testing are described in Section 7.

Fig. 3 shows the difference in quality assurance methods for real time and retrospective data collection.

Fig. 3. Quality assurance methods for real time and retrospective data collection



ANC, antenatal clinic; ART, antiretroviral therapy; PMTCT, preventing mother-to-child transmission; QA, quality assurance

6.6.5 Collecting routine records with missing data

Regardless of the surveillance design or data-collection method, HIV surveillance among pregnant women attending ANC based on routine data is likely to encounter routine records of eligible pregnant women with missing data. Missing data may include core surveillance variables (e.g. age or HIV status) or non-core surveillance variables (e.g. gravidity or occupation) (Section 4.1).

Routine records of eligible pregnant women that include missing data should be collected. These records provide important information for monitoring routine data quality (Section 7.1). In addition, sample-size calculations should have been adjusted to account for missing data (Section 6.4).

6.7 Data management for sentinel surveillance

A robust and comprehensive set of data management practices is essential to ensure the integrity of surveillance data, the reliability of surveillance estimates and the confidentiality of pregnant women whose routine data are collected for surveillance (Section 2).

6.7.1 Data security

Comprehensive data security measures include human, physical and electronic procedures and protections at every stage of the surveillance activity, including data collection, transfer, storage and analysis:

- All surveillance forms and electronic devices (e.g. computers, PDAs, tablets and USB drives) should be physically secured when not in use, both in the field and in an office setting.
- All routine data sources (e.g. ANC register and laboratory register) should be handled and maintained in a secure fashion during data collection, after which they should be returned to the appropriate secure location when they are routinely stored.
- All electronic devices, digital images and databases should be encrypted, password protected and accessible only by appropriate surveillance staff.
- All transfers of surveillance data though the internet should be secure and encrypted.”
- All surveillance databases should be regularly backed up on a secure, external system.
- Electronic surveillance data should only be shared with non-surveillance staff where appropriate and in accordance with national data-sharing policies.
- Surveillance databases should be securely preserved for future analysis.

Further guidance on data security is provided in the 2007 UNAIDS publication *Guidelines on protecting the confidentiality and security of HIV information* (23).

6.7.2 Data entry and cleaning

Data entry

Several best practices can help ensure accurate data entry when inputting data from paper forms or digital images into an electronic database:

- Ensure that data entry is done by trained clerks under supervision of a data manager.
- Design electronic data entry screens to match the format of data-collection tools. Epi Info and other programmes have the functionality to generate custom data entry screens.
- Define legal data values in the data entry screen that require confirmation of invalid data values (e.g. a pregnant woman aged 7 years). Verify invalid values against paper forms or digital images.
- Enter all data twice (“double data entry”) to allow checks for accurate entry (available software can automate this verification process). Compare the two datasets and resolve all discrepancies by checking paper forms or digital images.
- During data entry, regularly back up data files.

Data cleaning

As data are entered, frequency tables should be regularly produced for each ANC site and surveillance variable to identify data-quality issues that may have originated in data collection or data entry. Missing, inconsistent or invalid values should be checked against data-collection forms or digital images and corrections made. All corrections should be documented.

6.8 Data analysis for sentinel surveillance

This section presents methods for analysing data from HIV surveillance among pregnant women attending ANC. Although it presents methods for calculating HIV measures, these same analysis methods can also be applied to calculating syphilis measures.

6.8.1 Calculating site prevalence

Site prevalence \hat{p} is calculated by dividing the number of pregnant women who are HIV-positive (χ) (including pregnant women who newly tested positive or were previously known HIV-positive) by the number of pregnant women with known HIV status (η) (including pregnant women who newly tested positive, newly tested negative, or were previously known to be HIV-positive). Multiplying this proportion by 100% will express HIV prevalence as percentage positive.

The confidence interval for prevalence estimates can be calculated using the following formula, based on the normal approximation to the binomial distribution:

$$\hat{p} \pm 1.96 \sqrt{\frac{(1-\hat{p})\hat{p}}{\eta}} \times 100$$

where \hat{p} is the prevalence estimate and η is the total number of pregnant women with known HIV status.

6.8.2 Comparing site-level HIV prevalence from two surveillance rounds

Surveillance programmes may be able to test whether site-level HIV estimates from two sentinel surveillance rounds are different. Because sentinel surveillance only produces valid HIV prevalence estimates at the site level, this test can only be conducted at the site level, not at the national level (Section 8.1).

An approximate 95% confidence interval for the difference between two prevalence proportions \hat{p}_1 and \hat{p}_2 is given by:

$$\hat{p}_1 - \hat{p}_2 \pm 1.96 \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{\eta_1+2} + \frac{\hat{p}_2(1-\hat{p}_2)}{\eta_2+2}}$$

where:

$$\hat{p}_1 - \hat{p}_2 = \frac{\chi_1+1}{\eta_1+2} - \frac{\chi_2+1}{\eta_2+2}$$

and χ_1 and χ_2 are the total numbers of women who were HIV-positive in samples 1 and 2, respectively, and η_1 and η_2 are the corresponding sample-sizes.

For this test of difference to be valid, it is important that sentinel surveillance methods remain consistent across surveillance rounds.

6.8.3 Examining trends in HIV prevalence over time

Surveillance programmes can examine trends in site-level HIV prevalence or trends in median HIV prevalence across sites at national or subnational levels, if surveillance methods have remained the same.

Plotting HIV prevalence over time for specific demographic groups (especially age groups) will help in identifying specific changes in levels of HIV infection. The most effective way to examine trend data is graphically using scatter plots or line graphs, where prevalence estimates (and possibly confidence intervals) are plotted on the vertical axis and time is

plotted on the horizontal axis. It is preferable to have at least four rounds of data when presenting trends graphically.

HIV prevalence trends from sentinel surveillance can also be estimated and tested using statistical modelling. This is further described in Appendix A4.

6.8.4 Aggregating data and reporting summary prevalence

In sentinel surveillance, a surveillance estimate is only truly representative of the population that accesses the particular ANC site for which the estimate is made. Therefore, aggregating data across surveillance sites and using these aggregated data to calculate subnational (e.g. regional) or national prevalence estimates is not appropriate because sentinel ANC sites chosen by convenience are not representative of national or regional ANC sites. Similarly, surveillance programmes should not calculate the average (mean) of site-level prevalence estimates.

To report a single summary statistic of HIV prevalence among all sentinel ANC sites, surveillance programmes should use the median of site-level prevalence estimates. This will minimize the effects of outlying site-level prevalence estimates. The range for site-specific prevalence should also be reported to demonstrate the differences that may exist between sites. However, it should be noted that the median of site-level prevalence estimates is only a summary of site-level results—it is not an estimate of HIV prevalence among ANC attendees at the national or regional level.

6.8.5 Estimating national prevalence

Data from HIV surveillance among pregnant women attending ANC should not be used to estimate HIV prevalence beyond the population sampled by surveillance (pregnant women attending selected ANC sites) (Section 8.1). Estimates of HIV prevalence and trends in the general population using data from HIV surveillance among pregnant women attending ANC should be made using UNAIDS' Spectrum analysis tools (6).

6.8.6 Analysis of routine records with missing data

HIV surveillance among pregnant women attending ANC based on routine data is likely to collect some number of routine records of eligible pregnant women with missing data. Such data may include core surveillance variables or non-core surveillance variables (Section 4). The surveillance programme should define a consistent approach to handling records with missing data for each facet of analysis. It is suggested that:

- records with missing core surveillance variables cannot provide basic epidemiological data, and should therefore be excluded from all analyses; and
- records with missing non-core variables (e.g. marital status, ART status) can provide basic epidemiological data, and therefore should be retained for core surveillance

analyses; however, these records may not be able to contribute to non-core surveillance analyses.

The interpretation and dissemination of surveillance results should take into account and present information about missing data, and the limitations they impose on surveillance results.



7. MONITORING OF SURVEILLANCE

In this section

- 7.1 Monitoring individual-level data quality: data completeness and validity
- 7.2 Monitoring data quality: routine site data tools
- 7.3 Monitoring aggregate data quality: HMIS/routine reporting systems
- 7.4 Monitoring routine HIV testing
- 7.5 Addressing underperforming sites

All surveillance systems should be continuously monitored to ensure that surveillance methods, implementation and data provide reliable information for public health action. All surveillance systems should also be regularly evaluated to assess how well the surveillance system is meeting its objectives. Guidelines for the evaluation of surveillance systems can be found in WHO's 2013 publication *Evaluating a national surveillance system* (25). Monitoring HIV surveillance among pregnant women attending ANC based on routine programme data should focus on elements that could affect the reliability of surveillance estimates, including routine programme data quality, HIV-testing quality and other concerns.

Monitoring of surveillance can produce important information for interpreting surveillance results and understanding the quality of the surveillance system. Documenting the strengths of a surveillance system will generate confidence in surveillance among data consumers and civil society, whereas identifying limitations will inform measures to strengthen the system. To accomplish this, findings from surveillance monitoring should be freely disseminated and acted upon. Monitoring that does not inform action will not contribute to the reliability or sustainability of the surveillance system.

Because HIV surveillance among pregnant women attending ANC based on routine programme data relies on data and activities owned and managed by non-surveillance programmes (e.g. MCH and PMTCT programmes, and the national reference laboratory), monitoring HIV surveillance among pregnant women attending ANC should be a collaborative activity that engages these partners. Results of surveillance monitoring will likely provide valuable information to strengthen routine ANC data and HIV testing. Surveillance monitoring, and the translation of monitoring data into strengthening measures, will only be successful if surveillance programmes ensure the involvement and buy-in of relevant partners.

7.1 Monitoring individual-level data quality: data completeness and validity

When routine programme data collected for surveillance are analysed, the data should be monitored for completeness and validity at the site level to identify data-quality gaps. Individual-level surveillance data – whether from sentinel surveillance or a census of ANC based on individual-level data – can directly assess and report measures of data completeness and validity:

- Data are considered complete if they are present and legible in the designated field. If the field is blank or illegible, the value is considered incomplete.
- Data are considered valid if they are within the expected range of values for that field and invalid if they are not within that range. For example, a country may have standard operating procedures that dictate that the variable “HIV test result” should be recorded as “positive” or “negative”. Either of these is a valid value. However, the values “—” or “x” are considered invalid.

Data-quality monitoring measures should be the natural product of data-cleaning efforts (Sections 5.7.2 and 6.7.2).

7.2 Monitoring data quality: routine site data tools

The availability of routine programme data needed for surveillance is dependent on the routine site data tools (e.g. registers or EMR systems) used by ANC. If such tools do not contain a variable needed for surveillance, or capture the variable in a way that does not meet the data needs of surveillance, surveillance data will be incomplete. Sites using outdated routine data tools (e.g. an outdated version of a register or EMR system) or nonstandard routine data tools (e.g. a handmade register) may not be able to provide data needed for surveillance. Surveillance programmes should monitor the types of routine data tools used by ANC, and understand and document their effect on surveillance data.

Routine data tools used in ANC are regularly updated by the national MCH or PMTCT programmes, particularly when programme approaches shift, as in the case of PMTCT option B+. Changes in routine programme data tools can positively or negatively affect the quality of routine programme data. Surveillance programmes should be aware of planned changes to routine site data tools in ANC, and work with MCH or PMTCT programmes to ensure that any revised tools capture variables needed for surveillance.

7.3 Monitoring aggregate data quality: HMIS/routine reporting systems

The quality of data collected by a census based on aggregate data depends on routine site-level activities; that is, on the primary recording of data in site-level data tools. In addition, aggregate census data quality depends on above-site activities; that is, on the transfer and storage of aggregate data from the ANC sites to an above-site data repository through HMIS/routine reporting systems. Because the strength of the HMIS/routine reporting system will affect the quality of aggregate census data, surveillance programmes should work with programme partners to monitor the quality of these systems.

Comprehensive data-quality assessments can provide valuable information about the functioning of HMIS/routine reporting systems. However, more limited, routine activities should also regularly assess and support system-wide data quality:

- supervision to support good data practices at sites;
- routine data validation exercises; and
- regular data quality checks and feedback from districts to sites.

These routine activities may address the many facets of a reporting system, including:

- the primary recording of data in ANC site data tools;
- the aggregation and reporting of data from the ANC site through all levels of the reporting system to a centralized data repository;
- the capacity, roles and responsibilities and training of all staff involved in reporting at all levels of the system;
- clearly defined and documented data management and reporting requirements and processes at all levels of the system (e.g. standard operating procedures);
- standardized data collection and reporting forms and tools; and
- data-quality mechanisms and controls at all levels of the system.

International standards and tools for data-quality assurance and assessment are found in the WHO publications *Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies* and *Data quality assurance standards and tools for PMTCT programmes* and elsewhere (26-29).

Surveillance programmes should also work with programme partners to monitor changes in the HMIS/routine reporting systems, standards and tools used to transfer, manage and store aggregate ANC data at an above-site data repository. Such changes could affect the quality, accuracy, format or integrity of aggregate census data.

7.4 Monitoring routine HIV testing

HIV surveillance based on routine programme data relies on results from routine HIV testing (usually rapid diagnostic testing) at ANC to provide serostatus information on ANC attendees. To produce reliable surveillance estimates, the quality and accuracy of routine HIV rapid diagnostic testing should be high and supported by an appropriate system of quality assurance (30). Monitoring the following quality assurance elements will help to identify possible changes or limitations in the quality of routine HIV testing for diagnosis.

- *Training of staff:* Nurses, laboratory technicians or other ANC personnel conducting HIV testing of pregnant women should be well-trained and certified before performing HIV rapid diagnostic tests, and should undergo re-certification at least every 2 years. Comprehensive training should include theoretical basis of HIV diagnosis, HIV testing (specimen collection, test procedure, results interpretation), the validated national testing algorithm, quality assurance elements, and hands-on training with known and coded specimens.
- *Testing practices:* HIV rapid diagnostic tests should be performed following standard operating procedures developed for each assay of the testing algorithm. Critical elements include (a) application of right volume of specimen; (b) correct use of buffer(s), if indicated; (c) reading the results within the designated time period; and (d) following the national testing algorithm. Agreement between the individual tests in the algorithm, invalid results or other spurious results should be routinely monitored and investigated. Further guidance on developing and validating a national HIV testing algorithm is available in other WHO publications (31, 32).
- *Standardized record keeping:* Personnel conducting diagnostic testing should record and track the product names of test kits, lot numbers, expiration dates and individual test results in a standardized format and register as part of routine practice.
- *Participation in external quality assessment (EQA) programmes:* Testing sites and personnel performing diagnostic testing should participate in one or more EQA programmes. The EQA can involve multiple formats: (a) site visits by experienced laboratory technicians for review of testing and data management practices using a standardized checklist; (b) retesting of routinely collected specimens at another site or by different personnel; and (c) periodic testing of a standard panel of blinded or coded specimens with known HIV status by site staff ("proficiency testing").
- *Routine use of quality control specimens:* Testing personnel should routinely test known HIV-positive and negative specimens as quality control specimens. The frequency should be (a) once per day or week depending on the volume of testing; (b) when a new kit is opened; and (c) when a new kit lot arrives. Incorrect quality control results should be investigated before client specimens are tested.

- *Test kit stock management*: Inventory management is an essential element of a quality management system to ensure that (a) test kits are used within their manufacturer-defined shelf lives; (b) test kits are not compromised by inappropriate storage or handling; and (c) the correct testing algorithm is not compromised by stock-outs of any of the tests in the algorithm.

Further guidance on quality assurance for HIV rapid diagnostic testing is available from WHO (33).

It is also critical that syphilis testing quality assurance and quality control measures be robust and routine, to ensure the accuracy and reliability of routine syphilis testing data. In most settings, quality assurance and quality control measures for syphilis testing can be readily integrated with those for HIV, because the quality assurance considerations noted above also apply to routine syphilis testing.

7.5 Addressing underperforming sites

Whether the surveillance design is sentinel or census, surveillance may encounter some sites with underperforming PMTCT systems. Underperformance may be characterized by:

- substandard routine programme data quality for core surveillance variables: site routine programme data tools (e.g. registers or EMR systems) not formatted to appropriately collect the variables needed for surveillance, or recording of these variables being incomplete;
- routine HIV testing not being continuously available, or uptake of HIV testing among pregnant women being substandard; or
- routine HIV or syphilis testing quality being substandard, or quality assurance of routine HIV or syphilis testing being substandard.

If underperforming sites are identified, those sites should be targeted for support and strengthening to ensure high-quality programme services, data for programme monitoring and data for surveillance. If a substantial percentage of surveillance sites are found to be underperforming, this may indicate that a retrospective approach is not appropriate at the present time. In such a case, the surveillance programme may consider sentinel surveillance with real time data collection until routine PMTCT systems are strengthened.

8. INTERPRETATION OF RESULTS

In this section

- 8.1 Issues of representativeness and generalizability
- 8.2 Considerations affecting the interpretation of trend results
- 8.3 Other potential biases
- 8.4 Interpretation of syphilis surveillance results

This section provides guidance on the interpretation of the results of HIV surveillance among pregnant women attending ANC, including issues of representativeness, generalizability and bias pertaining to both prevalence estimates and trends in prevalence.

8.1 Issues of representativeness and generalizability

HIV surveillance among pregnant women attending ANC provides estimates of HIV among pregnant women attending selected ANC sites. Certain factors limit the representativeness or generalizability of surveillance results to additional populations. These factors are discussed below.

Generalizability to all pregnant women attending ANC

HIV surveillance among pregnant women attending ANC provides estimates of HIV prevalence among pregnant women attending ANC sites that are involved in surveillance. However, generalizing surveillance results to pregnant women attending all ANC should take into account the following considerations:

- *Surveillance design*: The generalizability of surveillance results to pregnant women attending ANC not involved in surveillance depends on the surveillance design.
 - *Sentinel surveillance design*: A convenience surveillance design does not provide a statistically representative sample of pregnant women attending ANC. Because sentinel ANC sites are not selected probabilistically, each site's HIV prevalence estimate is only representative of pregnant women attending ANC at that particular site. Further, site-level prevalence estimates from sentinel ANC sites cannot be combined to generate national or subnational prevalence estimates—only summary measures (e.g. medians of site estimates) are valid. These summary measures could be biased in so far as sentinel ANC sites may not accurately represent pregnant women attending non-sentinel ANC sites.
 - *Census*: A census of ANC will provide direct measures (not estimates) of HIV among all pregnant women attending all (or nearly all) ANC in a country.

- *Attendance at private ANC:* In many low- and middle-income countries, the overwhelming majority of pregnant women attend public ANC. However, in some countries the number of pregnant women accessing private ANC is large or growing. In these contexts, underrepresentation of private ANC in surveillance could introduce bias if the populations of pregnant women attending private and public ANC differ with regards to HIV exposure.
- *Rural pregnant women:* Because HIV prevalence tends to vary between urban and rural areas, it is essential that rural ANC attendees be appropriately represented in surveillance. Because ANC serving pregnant women in rural areas are often remote, smaller facilities with low client volume, it can be challenging to include them in surveillance. However, underrepresentation of rural ANC attendees will limit the generalizability of surveillance results.

Generalizability to all pregnant women

By definition, pregnant women who do not attend ANC are not represented in surveillance. If a substantial proportion of pregnant women do not attend ANC, caution should be employed in generalizing the findings of surveillance to all pregnant women. Non-attendance at ANC may be linked to demographic characteristics (e.g. place of residence, educational level, socioeconomic status or parity) that could also be related to HIV exposure. Rates of ANC attendance in a country may not be uniform and could differ by geographical region. Rates of ANC attendance could also change over time in a way that differs by geographical area. Surveillance programmes should carefully consider the impact of current and changing ANC attendance rates in the interpretation of surveillance results, and include information on trends in ANC attendance in materials used to disseminate surveillance results. Information on MCH indicators is generally available from population-based surveys, such as demographic and health surveys (34).

Generalizability to all women

Because only pregnant women who attend ANC are eligible to be included in HIV surveillance among pregnant women attending ANC, by definition women who do not become pregnant are not represented in surveillance. Age of sexual debut, fertility and family planning patterns and ART could significantly complicate the generalization of surveillance results to the general female population.

- *Age of sexual debut:*
 - HIV prevalence among young pregnant women will not be representative of prevalence among similarly aged women in the general population, which includes women who have never had sexual intercourse. This effect will be most pronounced among women younger than the typical age of sexual debut in a country. In these women the proportion of women who have never had sex is high, so exposure to HIV would be significantly higher among pregnant women (who have all experienced sexual debut) compared to similarly aged women in the general population.

- Increasing age of sexual debut may reduce HIV exposure and HIV prevalence among young women. This decline may be reflected among older (≥ 25 years) ANC attendees, among whom a later sexual debut might translate into reduced lifelong sexual encounters and HIV exposure. However, HIV prevalence trends among younger (< 25 years) ANC attendees may not reflect this decline because women who delay sexual debut (and thus reduce their risk of pregnancy and HIV exposure) will be less represented among young ANC attendees.
- *Effect of HIV on fertility:* HIV prevalence among older pregnant women (≥ 25 years) is likely to be lower than HIV prevalence among similarly aged, non-pregnant women because HIV infection is associated with decreased fertility (35). Biological reasons for lower fertility in HIV-positive women include higher rates of amenorrhea, coinfection with other STIs that cause subfertility, and increased rates of miscarriage and still births. Social reasons for lower fertility in infected women are higher rates of widowhood and divorce, lower rates of re-marriage or sexual activity, and fear of transmitting infection to the child (36). The fertility lowering effects of HIV generally increase with duration since infection. Thus, older HIV-positive women are less likely to become pregnant.
- *Effect of ART on fertility:* There are two ways in which ART could increase fertility among HIV-positive women, making ANC attendees more representative of the general female reproductive-aged population:
 - *Biological fertility:* ART may mitigate the negative biological effect of HIV. As a result, HIV-infected women may increasingly have the ability to conceive and carry a pregnancy to term. Research in this area is ongoing (37).
 - *Fertility intentions:* Dramatic expansion in the coverage of ART and PMTCT services in recent years has increased the coverage of HIV testing, reduced the likelihood of vertical transmission and improved the survival outcomes of HIV-positive women. These changes have the potential to alter the fertility preferences of HIV-positive women who know their status. Increased survival of HIV-positive women and men and reduced chance of vertical transmission may reduce widowhood and increase the fertility intentions of HIV-positive women, and increase their presence among ANC attendees. However, increased awareness of HIV status among HIV-positive women who know their status may lower their fertility intentions and reduce their presence among ANC attendees.
- *Family planning:*
 - There is a complex interaction between condom use and the relationship of the HIV prevalence in the ANC population and prevalence in the general female reproductive-aged population. Condoms protect against HIV and pregnancy; hence, HIV prevalence and fertility should be lower among women who consistently use condoms. However, ANC attendees have, by definition, not consistently used condoms. Therefore, frequent condom use in the general female population would result in higher HIV prevalence among ANC attendees compared to similarly aged women in the general population. On the other hand, condom use may become frequent among HIV-positive women (who seek to avoid pregnancy or transmission to an uninfected

partner). Consistent condom use among HIV-positive women who know their status would reduce their fertility. This would reduce HIV prevalence among ANC attendees compared to similarly aged women in the general female reproductive-aged population. These two effects could also vary by age group.

- Use of hormonal methods of contraception (pills, injectables, implants) may also differ between HIV-negative and HIV-positive women. For example, HIV-positive women who know their status may decide not to become pregnant because they are concerned with the possibility of vertical transmission or their children becoming orphans. If hormonal contraceptive use is higher in HIV-positive women, then HIV prevalence will be lower among pregnant women than among similarly aged, non-pregnant women.
- In the context of UNAIDS' global plan for elimination of new HIV infections among children and keeping their mothers alive, HIV-positive women are increasingly benefiting from integrated HIV treatment and family planning services (38). These advances could enhance the ability of HIV-positive women to avoid unintended pregnancies and reduce their presence in ANC populations. This would reduce HIV prevalence among ANC attendees relative to the general female reproductive-aged population.

Generalizability to the general population

By definition, men are not represented in HIV surveillance among pregnant women attending ANC, complicating the generalization of surveillance results to the general population. Because HIV prevalence tends to be higher among women than among men in generalized epidemics (as is the case in sub-Saharan Africa), surveillance among pregnant women attending ANC may overestimate HIV prevalence in the general population. In addition, because men in generalized epidemics tend to become infected at an older age than women, the extrapolation of HIV prevalence estimates among ANC attendees aged <35 years to similarly aged men may overestimate HIV prevalence (39).

Conversely, in concentrated epidemics where key populations (including men who have sex with men and people who inject drugs) bear a heavy burden of HIV, the HIV epidemic may be disproportionately male (40). In these contexts, HIV surveillance among pregnant women attending ANC results may underestimate HIV prevalence in the general population.

Finally, because HIV surveillance among pregnant women attending ANC excludes women beyond reproductive age, surveillance results will not represent either men or women aged ≥ 50 years. Surveillance may also poorly represent the population aged ≥ 35 years, as women aged ≥ 35 years typically comprise a smaller proportion of ANC attendees compared to their proportion of the population as a whole. These age limitations may become increasingly significant as epidemics age due to increased survival of people living with HIV (PLHIV) on ART.

8.2 Considerations affecting the interpretation of trend results

Trend analysis that provides information on the direction of HIV epidemics is a key part of the "know your epidemic" approach of second-generation surveillance of HIV. However,

certain considerations should be taken into account when interpreting trends in HIV prevalence among ANC attendees.

8.2.1 Incidence and mortality

The size of the HIV-positive population has two main determinants: incidence and mortality. Incidence is the rate of new HIV infections in a susceptible population over a given period of time. HIV-associated mortality is the rate at which PLHIV die from HIV disease. Thus, in a stable subpopulation under surveillance, incidence measures the inflow of persons into the HIV-positive subpopulation, and mortality measures the outflow of persons exiting the HIV-positive subpopulation. In- and out-migration may also have a limited effect on the size of the HIV-positive population or underlying general population. This effect may be particularly important in small subnational populations.

Before the advent and scale-up of ART, trends in HIV prevalence could be interpreted as more closely following the evolution of the epidemic. Early in HIV epidemics, HIV-associated mortality was low and prevalence was highly sensitive to incidence, as new infections were added to the HIV-positive subpopulation but few exited that subpopulation. As epidemics matured and HIV-associated mortality increased, the relationship between incidence and prevalence was moderated but still strong—especially in younger age groups where duration of infection was short and few deaths occurred. Because HIV-associated mortality was high, PLHIV spent a relatively short and well-defined time in the HIV-positive subpopulation (approximately 8–12 years on average) before exiting at death (41). Hence, prevalence was dependent on new infections entering the subpopulation (incidence) to replace those exiting the subpopulation (mortality).

However, the scale-up of ART has complicated the relationship between incidence and prevalence. Increases in ART coverage have dramatically improved the survival of PLHIV and reduced HIV-associated mortality. In this context, the effect of changes in incidence on prevalence is less evident. Coverage of ART and related changes in HIV-associated mortality should be carefully considered when interpreting trends in HIV prevalence among pregnant women attending ANC.

In addition, the ANC population may have special characteristics that could affect the relationship between incidence, mortality and prevalence represented in this population. Age structure of the ANC population is different from the general population. Generally (depending on the country context), approximately 30–35% of the ANC population is aged 15–24 years, 25–30% is aged 25–29 years, 10–15% is 30–34 years, and some 10–15% is aged ≥ 35 years. Although there are a limited number of ways to exit the general population (death, out-migration), there are many ways HIV-positive women could exit the ANC population: death, out-migration, ageing, HIV-related reduced fertility (including widowhood), reduced sexual activity, or contraception use. In all cases, an HIV-positive woman would no longer be counted as a prevalent case in the ANC population. Thus, ageing and reduced childbearing may moderate the above described effects of lower mortality on prevalence, making ANC prevalence more sensitive to incidence.

8.2.2 Trends among young pregnant women

HIV infections among young pregnant women aged 15–24 years are assumed to be recent as women in this group have likely only recently experienced sexual debut. Further, HIV-associated mortality in this group is assumed to be low, so trends in prevalence are more closely related to incident infections, and less influenced by mortality. Hence, trends in HIV prevalence among this population have been considered a general proxy for trends in HIV incidence (42, 43). However, this logic would be challenged if female children infected through mother-to-child transmission survive to reproductive age (43-47). Survival rates among this population would be expected to increase with the expansion of access to ART. As they age and become sexually active, these female children could increasingly contribute to the population of HIV-positive 15–24 year old pregnant women as non-recent infections. The scale-up of PMTCT interventions would be expected to counteract this effect in due course. The population of young female children infected with HIV through vertical transmission would diminish over time as coverage and quality of PMTCT programmes improve.

For these reasons, the contribution of women infected through vertical transmission to the population of young HIV-positive pregnant women may increase, diminish or remain stable, reflecting the history of the HIV epidemic and scale-up of ART and PMTCT in country. Epidemic and programme history should be carefully considered when interpreting trends in HIV prevalence among young pregnant women.

When interpreting trends in HIV prevalence among young pregnant women, it is also important to consider changes in family planning practices and age of sexual debut among the underlying 15–24 year old female population, because prevalence patterns in this population are particularly sensitive to changes in age of sexual debut and barrier contraceptive use (Section 8.1) (11).

8.2.3 Interpretation of surveillance results across systems

Changes in trends in HIV prevalence across the evolution of surveillance systems should be interpreted with caution, taking into account the following considerations.

Impact of testing strategies on HIV trends

Transition from UAT-based serosurveillance to using routine programme data for surveillance may imply a change in HIV testing strategy and technology. HIV-testing strategies and algorithms for diagnosis of HIV among pregnant women attending ANC may differ from HIV-testing strategies and algorithms historically used for UAT-based serosurveillance (48).

Recent UAT-based ANC serosurveys have often been conducted using a fourth-generation enzyme immunoassay (EIA) that detects both HIV antigen and antibodies to HIV-1/2. However, most, but not all, HIV rapid diagnostic tests used in ANC settings detect only antibodies to HIV-1/2. Hence, fourth-generation HIV EIAs are expected to be marginally more sensitive, and produce slightly more positives, than second- and third-generation HIV

rapid diagnostic tests that detect antibodies only. Hence, HIV prevalence estimates based on second- and third-generation HIV antibody rapid diagnostic tests could be marginally lower than estimates based on a fourth-generation EIA only.

Impact of changes in surveillance design

Transition from sentinel surveillance to surveillance based on a census of ANC may have implications for the interpretation of surveillance trends. Results of sentinel surveillance are (a) prevalence estimates of a specified precision, among (b) a sample of pregnant women attending (c) a subset of ANC sites sampled by convenience. By contrast, results of a census are (a) direct measures of prevalence, among (b) all pregnant women attending (c) all (or nearly all) ANC sites providing PMTCT services. Hence, surveillance programmes transitioning from sentinel surveillance to census may see a correction in ANC surveillance trend lines as geographical coverage of surveillance greatly expands and the uncertainty around surveillance measures is removed.

8.3 Other potential biases

Biases arising from sampling or data-collection methods may affect the validity of a measured HIV prevalence. Surveillance programmes should take steps to identify and minimize the potential for bias in surveillance estimates. Three common biases that may affect HIV surveillance among pregnant women attending ANC based on routine programme data are selection, information and measurement biases, as discussed below.

Selection bias

Selection bias occurs when pregnant women attending ANC whose routine data are collected for surveillance differ in an important way, such as demographic characteristics or HIV exposure, from pregnant women whose data are not collected. Selection bias can be introduced by changes in:

- ANC attendance rates, including changes in policies affecting access to antenatal services (e.g. service fees);
- uptake of routine HIV testing among pregnant women; and
- the services offered at ANC.

The rollout of option B+ also has the potential to introduce selection bias. Depending on the service delivery model, ANC and ART services could become increasingly integrated in the context of B+. In this scenario, known HIV-positive women who are already on ART and become pregnant could receive ANC services at a site other than their traditional ANC. These women might, instead, receive ANC services at an ART site or enhanced ANC that provides integrated ANC/ART services. If known HIV-positive pregnant women who are already on ART are systematically directed to select ANC/ART sites, bias could be

introduced into the surveillance sample. HIV prevalence estimates at traditional ANC could decline, while prevalence estimates at integrated ANC/ART sites could rise.

The surveillance programme should monitor changes in ANC attendance, and ANC or PMTCT service delivery or uptake that could affect the validity of surveillance estimates.

Information bias

Information bias is a systematic error in the method of surveillance data collection. Information bias is related to the instruments used to collect the data and how the data are recorded. Information bias can be introduced by changes in:

- surveillance design or methods for selecting surveillance sites;
- surveillance data-collection methods;
- the structure, quality or availability of routine programme data collected for surveillance, in particular information on pregnant women who already know their status to be HIV-positive at ANC booking; and
- (where EMRs are used for surveillance) the quality, accuracy or integrity of EMR data, or changes in the format or management of EMR data.

To minimize information bias, surveillance programmes should work to ensure that surveillance methods are rigorous and consistent across surveillance rounds, and carefully monitor routine data quality (Section 7).

Measurement bias

Measurement bias is a systematic error in the measurement or classification of surveillance parameters or variables. To minimize measurement bias, surveillance should verify that variables of interest have been correctly measured or classified in routine programme data. This is particularly important with regards to HIV status. If there are systematic errors in routine HIV-testing results, surveillance estimates based on these results will be biased. The monitoring of routine HIV-testing quality is further described in Section 7.4.

8.4 Interpretation of syphilis surveillance results

To appropriately interpret routine syphilis data, it is important to understand the impact of different testing strategies for detection of syphilis (Table 6).

Table 6. Comparison of routine syphilis diagnostic strategies

Type of test	Characteristics	
Laboratory-based	Non-treponemal (e.g. RPR, VDRL)	<ul style="list-style-type: none"> • Can distinguish between active and past treated infection • Rare false-positive results (inflammation from other causes) • Rare false-negative results (prozone effect)
	Treponemal (e.g. TPPA, TPHA, FTA-ABS, EIA)	<ul style="list-style-type: none"> • Cannot distinguish between active and past treated infection
Rapid diagnostic test	Single treponemal	<ul style="list-style-type: none"> • Cannot distinguish between active and past treated infection
	Combination treponemal/non-treponemal	<ul style="list-style-type: none"> • Can distinguish between active and past treated infection
	Combination treponemal/HIV	<ul style="list-style-type: none"> • Cannot distinguish between active and past treated infection

EIA, enzyme immunoassay; FTA-ABS, fluorescent treponemal antibody absorption; RPR, rapid plasma regain; TPHA, *Treponema pallidum* haemagglutination assay; TPPA, *Treponema pallidum* particle agglutination assay; VDRL, venereal disease research laboratory

Traditionally, syphilis testing in pregnancy has been conducted using laboratory-based non-treponemal tests, which can identify active infection. Where capacity exists, confirmation with treponemal laboratory-based methodology is used, which is more specific for history of infection with *Treponema pallidum*, but cannot distinguish between current and past treated infection. Traditional laboratory-based non-treponemal and treponemal assays require a trained laboratory technician, equipment and electricity to store reagent and power equipment. However, over the past decade, new testing technologies have facilitated the expansion of syphilis testing in ANC settings, including rapid treponemal, combination rapid treponemal/non-treponemal, and dual rapid treponemal/HIV point-of-care tests (49).

To accurately interpret routine syphilis testing data, it is also important to account for the testing algorithm used. Two types of algorithms are typically used in ANC settings:

- *single test strategies*: either a single treponemal or non-treponemal assay; and
- *two test strategies*: a treponemal assay followed by a non-treponemal assay, or a non-treponemal assay followed by a treponemal assay.

In both algorithms, assays may be rapid or laboratory-based.

Finally, it is important to understand the type of test result recorded in routine records. Where two test strategies are used, routine data may capture the results of the full testing algorithm or just the screening result.

When analysing trends in syphilis seropositivity over time, it is critical to document the testing strategy and algorithms (assay format) used for each data point. If changes in testing practices have occurred, seropositivity rates must be adjusted to reflect changes in assay type in order to distinguish between true changes in burden of disease and artifactual changes due to changes in assay type. For example, a country has traditionally diagnosed syphilis seropositivity based on both treponemal and non-treponemal test results. If this country then switches to using only treponemal assays, the seropositivity rate based on only treponemal test results should be adjusted based on local epidemiology to reflect expected combined treponemal and non-treponemal seropositivity.

Analysis of syphilis data should follow the same approach as used for HIV data (Sections 5.8 and 6.8), including analysis by subnational unit, demographic subgroups, and trends over time. It is also important to include analysis of syphilis and HIV coinfection. Changes in syphilis prevalence, especially among pregnant women aged 15–24 years, may reflect changes in HIV risk behaviour before these changes are seen in HIV prevalence estimates.



9. DISSEMINATION OF SURVEILLANCE RESULTS

Public health surveillance is the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action (50). By definition, surveillance results should be widely and strategically disseminated to inform public health action. Following best practices in the dissemination of surveillance findings can optimize the impact of surveillance results, and ensure official and civil society support for surveillance systems.

Identify target audiences for dissemination

Results of HIV surveillance among pregnant women attending ANC should be disseminated to all persons and organizations implicated in the response to the HIV epidemic and STI control efforts, especially those responsible for decision-making. Potential target audiences for dissemination products include public health officials at the national, regional and district levels; government officials, policy-makers and planners; civil society organizations and advocacy groups; national and international partners; journalists and the media; the PLHIV community; and the general public. It is important that surveillance results are also disseminated to persons who contribute to the recording and collection of routine programme data used for surveillance (e.g. staff at surveillance sites) to maintain their engagement and commitment.

Disseminate messages to achieve objectives

Before dissemination, the surveillance programme should decide on the dissemination objectives for each audience. What does the surveillance programme want to achieve by communicating results? What actions does the surveillance or HIV/STI programme want the audience to take based on the communicated results? To achieve dissemination objectives, surveillance results should not be communicated as raw information or in language that is highly technical (unless the audience is highly technical). Rather, the dissemination of surveillance results should communicate specific messages that contextualize, summarize, interpret and prioritize the results, and that point towards action steps indicated by surveillance results. These functions can be summarized by three questions: What were the results? What do the results mean? What needs to be done? (51). Dissemination messages that address these questions will help to ensure that each target audience receives the intended message and takes the desired action.

Employ dissemination channels and products appropriate for the target audience

Dissemination products and channels (how the dissemination products are delivered) that are tailored to the target audience will be more successful in meeting dissemination objectives. A variety of products that can be used to disseminate surveillance results, including surveillance reports, fact sheets, briefings, slide sets, press releases and scientific

journals. Likewise, there are multiple options for dissemination channels, including workshops, press conferences and formal meetings.

Different audiences have different needs, expectations and capacity with regards to the consumption of surveillance results. Thus, the choice of dissemination products and channels should be tailored to the audience, based on that audience's data needs or expectations, their technical expertise, and their needs and expectations with regards to personal engagement.

A dissemination strategy should align the channel, product and target audience:

- Public health officials will have considerable technical expertise, require significant breadth and depth of data for use in programme planning and evaluation, and have substantial expectations with regards to engagement. A full surveillance report delivered in a workshop may be appropriate for this audience.
- Government officials, policy-makers and planners may have limited technical expertise, limited data needs and many demands on their time. However, government officials are important decision-makers and may require significant personal engagement. Briefings or slide sets focusing on major results, involving audience-friendly displays and delivered in formal meetings, may be appropriate for this audience.
- Journalists, media and the public have limited technical expertise and data needs. Press releases or fact sheets focusing on major results, involving audience-friendly displays and provided in press conferences, may be appropriate for this audience.
- Civil society organizations and the PLHIV community may have moderate technical expertise and data needs, but significant expectations with regards to personal engagement. Surveillance reports, fact sheets and slide sets provided in formal meetings may be appropriate for this audience.

Disseminate results in a timely fashion

For surveillance results to be used for public health action, they should provide a current picture of the HIV epidemic. This can only be achieved if surveillance results are disseminated in a timely fashion.



10. APPENDICES

A1. Sample-size for a single prevalence proportion in sentinel surveillance

Estimates of the number of pregnant women who should be sampled are determined by the presumed level of prevalence being measured, and the desired precision of the prevalence estimate.

Under simple random sampling with replacement, an approximate confidence interval around a prevalence proportion ρ is given by:

$$\hat{\rho} \pm Z_{1-\alpha/2} \sqrt{\frac{\hat{\rho}(1-\hat{\rho})}{\eta}}$$

where:

η = the sample-size

$Z_{1-\alpha/2}$ = the z score corresponding to the level of significance $1 - \alpha / 2$

$\hat{\rho}$ = the estimated proportion of HIV-positive women.

A 95% confidence interval is specified by choosing $\alpha = 0.05$. The 95% confidence intervals shown in Table 3 are based upon that approximation.

The margin of error d is half the width of the confidence interval, and is given by:

$$d = Z_{1-\alpha/2} \sqrt{\frac{\hat{\rho}(1-\hat{\rho})}{\eta}}$$

Therefore, an approximate estimator of sample-size for simple random sampling ($\hat{\eta}_{\text{SRS}}$) is given by:

$$\hat{\eta}_{\text{SRS}} = \frac{Z_{1-\alpha/2}^2 \hat{\rho}(1-\hat{\rho})}{d^2}$$

A2. Sample-size for detection of changes in HIV prevalence in sentinel surveillance

To calculate a sample-size that enables detection of statistically significant changes in HIV prevalence over time, surveillance programmes should decide on the magnitude of detectable change in prevalence that is desired (e.g. 10% change, 50% change) and the level at which a change in prevalence will be measured (e.g. national, regional).

The sample-sizes required to detect changes in prevalence are influenced by the choice of statistic used to express change and the precision required in the estimation of that statistic.

Two commonly used measures of change over time are the difference in prevalence between two points in time:

$$\delta_{t+\Delta t} = \rho_{t+\Delta t} - \rho_t$$

and the ratio of prevalence estimates at two points in time:

$$\theta_{t+\Delta t} = \rho_{t+\Delta t} / \rho_t$$

where:

t = time, in years, of an initial survey

Δt = the survey interval, in years.

For surveillance surveys conducted every 2 years, $\Delta t = 2$. The difference, $\delta_{t+\Delta t}$, represents changes in absolute prevalence proportions, and the quantity $\delta_{(t+\Delta t)/\Delta t}$ is the linear trend slope on the prevalence scale. The ratio $\theta_{t+\Delta t}$ is interpreted as the relative change in prevalence over time, and is analogous to the risk ratio where the "exposure" is time.

An approximate estimator of the sample-size required to obtain a $100(1-\alpha/2)\%$ confidence interval on the difference $\delta_{t+\Delta t} = \rho_{t+\Delta t} - \rho_t$ from SRS is given by:

$$\hat{n}_{\text{SRS}} = \frac{Z_{1-\alpha/2}^2 [\hat{\rho}_t(1-\hat{\rho}_t) + \hat{\rho}_{t+\Delta t}(1-\hat{\rho}_{t-\Delta t})]}{d^2}$$

where d is the desired margin of error on the difference between proportions. Typically $\hat{\rho}_t$ and $\hat{\rho}_{t-\Delta t}$ will be available from previous surveys.

In countries where HIV prevalence is decreasing, a moderately protective sample-size estimate may be obtained by assuming that prevalence is constant at $\hat{\rho}_t$, in which case the sample-size estimator is simply:

$$\hat{n}_{\text{SRS}} = \frac{2Z_{1-\alpha/2}^2 [\hat{\rho}_t(1-\hat{\rho}_t)]}{d^2}$$

In either case, the design-adjusted sample-size is again:

$$\hat{n}_{\text{SRS}} = \frac{\text{DEFF} \times \hat{n}_{\text{SRS}}}{r}$$

where $r \leq 1$ is the fraction of sampling units which are expected to provide complete data.

Sample-size estimation for the ratio $\theta_{t+\Delta t} = \rho_{t+\Delta t} / \rho_t$ requires numerical solution and is supported by commonly used sample-size estimation software.

A3 Sample-size considerations for young pregnant women in sentinel surveillance

Surveillance programmes may consider a sampling design and sample-size requirements that enable prevalence estimates of a specified precision among young (aged 15–24 years) pregnant women at the site level. There are two options to control the precision of prevalence estimates among young pregnant women at the site level: inflation and oversampling.

Inflation

The simpler option is to inflate the sample-size to ensure HIV prevalence estimates of a specified precision among young pregnant women. To do this, the original calculated sample-size is divided by the expected proportion of the sample that will be young pregnant women

$$\frac{\text{unadjusted sample-size}}{\text{expected proportion of sample that is young women}} = \text{inflated sample-size}$$

For example, suppose an unadjusted sample-size requirement of 300. Now suppose that pregnant women aged 15–24 years are expected to comprise 32% of eligible pregnant women. The original calculated sample-size is divided by 0.32 to produce an inflated sample-size of 938.

$$\frac{300}{0.32} = 938$$

Estimates of the proportion of eligible pregnant women who are aged 15–24 years should be available from previous surveillance rounds.

Oversampling

In an oversampling approach, an overall sample-size is calculated for all eligible pregnant women aged 15–49 years, and a separate sample-size is calculated for young pregnant women. All eligible pregnant women aged 15–49 years would be sampled until the overall

sample-size is met. At this point, sampling would continue only among young pregnant women until the sample-size for young pregnant women is met. For example, a site has an overall target sample-size of 300 pregnant women aged 15–49 years and a target sample-size of 300 for young pregnant women. When the overall target of 300 is reached, 200 of these are young pregnant women. Sampling would then continue only among young pregnant women until 100 additional young pregnant women are sampled.

This approach is more complicated than inflation. Oversampling requires the calculation and application of two different samples sizes to two different populations, and close monitoring of surveillance data to ensure proper sampling. Retrospective data collection (Section 6.6.3) would facilitate this more complicated approach. Further, this approach requires separate analyses. In our example above, the overall prevalence would be calculated for all eligible pregnant women aged 15–49 years using the first 300 women sampled. The 100 young pregnant women additionally sampled should be excluded from this analysis. Prevalence among young pregnant women would be separately calculated using the 200 young pregnant women from the initial 300 women, and the 100 additionally sampled young pregnant women.

A4 Modelling analysis of trends in HIV prevalence from sentinel surveillance

Statistical models are simplified mathematical representations of hypotheses about major factors that explain or predict some outcome which, in this case, is HIV prevalence among pregnant women who are served by sentinel sites surveillance sites. Statistical models are important because they filter out random variations to reveal the signals of effects on HIV prevalence. Therefore, statistical models provide the means to test our hypotheses against the evidence in the data, adjust estimates for effects of covariates, and make data-driven predictions of new observations.

By definition, prevalence proportion ρ should be confined to the closed interval $[0,1]$, and therefore the logit transformation $\text{logit}(\rho) = \log(\rho/(1-\rho))$ is commonly used to model factors that affect prevalence, and is useful for trend analysis. Linear trends on the logit scale correspond to curvilinear trends on the prevalence scale. Analyses of trends in prevalence can therefore be implemented as variations on logistic regression.

The simplest methods can be implemented using estimates of prevalence proportion from each surveillance round. Let $\hat{\rho}_{st}$ denote the survey-based estimates of prevalence from sentinel sites surveillance site s , $s = 1, \dots, S$, during surveillance round t , $t = 1, \dots, T$ and let $\text{Var}(\hat{\rho}_{st})$ denote the sampling variances where S and T are the numbers of sites and surveillance rounds, respectively. If the sampling variances are approximately equal, then the mixed-effects linear regression model

$$\text{logit}(\hat{\rho}_{st}) = \beta_0 + u_{0s} + (\beta_1 + u_{1s})t + \varepsilon_{st}$$

where β_0 and β_1 are the intercept and slope in time, respectively, on the logit scale, u_{0s} and u_{1s} are site-specific random effects on the intercept and slope, respectively, which are assumed to be normally distributed with mean 0, and the \mathcal{E}_{st} are observation errors, which are assumed to be independently and normally distributed with mean 0. The assumption of equal sampling variances can be relaxed by weighting by the reciprocals of the sampling variances. In this simple logit-linear regression model, $\hat{\rho}_{st}$ and $\hat{\rho}_{st-1}$ are assumed to be independent (serial independence). Therefore, this model provides a rather crude approximation to trends in prevalence because prevalence during surveillance round t should be some function of prevalence during surveillance round $t-1$ whenever surveillance is monitored every few years, because some of the same women remain in the population. However, a model assuming serial independence is the only viable option when there are few rounds of surveillance data.

In this model, the S site-specific logit-scale intercepts are given by the quantities $\beta_0 + u_{0s}$ and site-specific slopes are given by $\beta_1 + u_{1s}$. Similarly, the among-site median-profile logit-scale intercept and slope are β_0 and β_1 , respectively. Inferences about logit-linear trends are based on confidence intervals for estimates of those parameters. Most software can also produce pointwise confidence or prediction intervals on the model-based expected values prevalence. Let $\tilde{\rho}_{st}$ denote the model-based predictions of site-specific prevalences $\hat{\rho}_{st}$ given by

$$\tilde{\rho}_{st} = \text{expit}(\hat{\beta}_0 + \hat{u}_{0s} + (\hat{\beta}_1 + \hat{u}_{1s})t)$$

where $\text{expit}(\chi) = 1/(\exp(-\chi) + 1)$ is the inverse of the logit transformation of χ . Similarly, the among-site median-profile predicted prevalences are given by

$$\tilde{\rho}_t = \text{expit}(\hat{\beta}_0 + \hat{\beta}_1 t).$$

Note that these among-site predicted prevalences represent only those women who might have been served at the S sentinel sites surveillance sites, and do not represent national trends in prevalence. This mixed-effects logit-linear trend model can be fitted by commonly available statistical software including the SAS GLIMMIX procedure, the Stata melogit command and several R packages.

A5 Example of a sentinel surveillance data-collection form

Site:	Site code:	District:
Sociodemographic information		
1. Surveillance code	_____	
2. Date of first ANC visit	(dd/mm/yyyy) ____ / ____ / _____	
3. Age	____	
4. Marital status	<input type="checkbox"/> A) Single <input type="checkbox"/> B) Married <input type="checkbox"/> C) Divorced <input type="checkbox"/> D) Widowed	
5. Gravity and/or parity	____	
Information on HIV status at booking		
6. HIV status at booking	<input type="checkbox"/> A) Negative <input type="checkbox"/> B) Previously tested HIV-positive <input type="checkbox"/> C) Unknown	
7. ART status at booking	<input type="checkbox"/> A) On ART <input type="checkbox"/> B) Not on ART	
HIV-testing information		
8. HIV tested at first ANC visit	<input type="checkbox"/> A) Yes, tested <input type="checkbox"/> B) No, not tested <input type="checkbox"/> C) No, not tested because previously known HIV-positive <input type="checkbox"/> D) Testing not documented	
9. HIV assay #1	9.1 Assay #1 name:	
	9.2 Assay #1 result:	<input type="checkbox"/> A) Negative <input type="checkbox"/> B) Positive <input type="checkbox"/> C) Indeterminate <input type="checkbox"/> D) Result not documented
10. HIV assay #2	10.1 Assay #2 name:	
	10.2 Assay #2 result:	<input type="checkbox"/> A) Negative <input type="checkbox"/> B) Positive <input type="checkbox"/> C) Indeterminate <input type="checkbox"/> D) Result not documented

11. HIV assay #3	11.1 Assay #3 name:	
	11.2 Test #3 result:	<input type="checkbox"/> A) Negative <input type="checkbox"/> B) Positive <input type="checkbox"/> C) Indeterminate <input type="checkbox"/> D) Result not documented
12. Final HIV result	<input type="checkbox"/> A) Negative <input type="checkbox"/> B) Positive <input type="checkbox"/> C) Indeterminate <input type="checkbox"/> D) Result not documented	
Syphilis testing information		
13. Syphilis tested at first ANC visit	<input type="checkbox"/> A) Yes, tested <input type="checkbox"/> B) No, not tested <input type="checkbox"/> C) Testing not documented	
14. Syphilis assay #1	Assay #1 name:	
	Assay #1 result:	<input type="checkbox"/> A) Negative <input type="checkbox"/> B) Positive <input type="checkbox"/> C) Indeterminate <input type="checkbox"/> D) Result not documented
15. Syphilis assay #2	Assay #2 name:	
	Assay #2 result:	<input type="checkbox"/> A) Negative <input type="checkbox"/> B) Positive <input type="checkbox"/> C) Indeterminate <input type="checkbox"/> D) Result not documented
16. Final syphilis results	<input type="checkbox"/> A) Negative <input type="checkbox"/> B) Positive <input type="checkbox"/> C) Result not documented	
17. Syphilis treatment provided	<input type="checkbox"/> A) Yes → (specify) _____ <input type="checkbox"/> B) No <input type="checkbox"/> C) Treatment not documented	

11. REFERENCES

- 1 Guidelines for assessing the utility of data from prevention of mother-to-child transmission (PMTCT) programmes for HIV sentinel surveillance among pregnant women. Geneva, Joint United Nations Programme on HIV/AIDS/World Health Organization. 2013 (http://apps.who.int/iris/bitstream/10665/85512/1/9789241505611_eng.pdf, accessed 19 June 2015).
- 2 Diaz T, De Cock K, Brown T, Ghys PD, Boerma JT. New strategies for HIV surveillance in resource-constrained settings: An overview. *AIDS*. 2005;19 Suppl 2:S1-8 (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15930836, accessed 19 June 2015).
- 3 Unlinked anonymous screening for the public health surveillance of HIV infections: Proposed international guidelines Geneva, World Health Organization Global Programme on AIDS. 1989 (http://whqlibdoc.who.int/hq/1989/gpa_sfi_89.3.pdf, accessed 19 June 2015).
- 4 Guidelines for conducting HIV sentinel serosurveys among pregnant women and other groups. World Health Organization/Joint United Nations Programme on HIV/AIDS. 2003 (http://data.unaids.org/Publications/IRC-pub06/JC954-ANC-Serosurveys_Guidelines_en.pdf, accessed 19 June 2015).
- 5 Guidelines for second generation HIV surveillance. An update: Know your epidemic. Geneva, Joint United Nations Programme on HIV/AIDS/World Health Organization Working Group on Global HIV/AIDS and STI Surveillance. 2013 (http://apps.who.int/iris/bitstream/10665/85511/1/9789241505826_eng.pdf?ua=1, accessed 19 June 2015).
- 6 Spectrum/EPP 2014. Joint United Nations Programme on HIV/AIDS. 2015 (<http://www.unaids.org/en/dataanalysis/datatools/spectrumep2013/>, accessed 19 June 2015).
- 7 Consolidated guidelines on general HIV care and the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. Geneva, World Health Organization. 2013 (<http://www.who.int/hiv/pub/guidelines/arv2013/en/>, accessed 19 June 2015).
- 8 Global update on the health sector response to HIV, 2014. Geneva, World Health Organization. 2014 (<http://www.who.int/hiv/pub/progressreports/update2014/en/>, accessed 19 June 2015).
- 9 Guiding principles on ethical issues in HIV surveillance. Geneva, Joint United Nations Programme on HIV/AIDS (UNAIDS)/WHO Working Group on Global HIV/AIDS and STI Surveillance. 2013 (http://apps.who.int/iris/bitstream/10665/90448/1/9789241505598_eng.pdf, accessed 19 June 2015).

- 10 US Department of Health and Human Services. The Belmont report: Ethical principles and guidelines for research involving human subjects. Washington, D.C., The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1979 (<http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>, accessed 19 June 2015).
- 11 Zaba B, Boerma T, White R. Monitoring the AIDS epidemic using HIV prevalence data among young women attending antenatal clinics: Prospects and problems. *Aids*. 2000;14(11):1633-1645 (<Go to ISI>://WOS:000089021600019, accessed 19 June 2015).
- 12 Packages of interventions: Family planning, safe abortion care, maternal, newborn, and child health. Geneva, World Health Organization. 2010 (http://whqlibdoc.who.int/hq/2010/who_fch_10.06_eng.pdf, accessed 19 June 2015).
- 13 Elimination of new paediatric HIV infections and congenital syphilis in Asia-Pacific: Conceptual framework and M&E guide. Bangkok, United Nations Children's Fund East Asia and Pacific Regional Office. 2011 (http://www.unicef.org/eapro/PPTCT_CF_and_ME_guide_17Aug11.pdf, accessed 19 June 2015).
- 14 Pan American Health Organization. Regional initiative for elimination of MTCT of HIV and congenital syphilis in Latin America and the Caribbean: Concept document for the Caribbean. Washington, D.C., PAHO. 2010 (<http://www.paho.org/derechoalaSSR/wp-content/uploads/Documentos/Regional-initiative-to-eliminate-mother-to-child-HIV-and-syphilis-transmission.pdf>, accessed 19 June 2015).
- 15 Strategic framework for the elimination of new HIV infections among children in Africa by 2015. Brazzaville, World Health Organization Regional Office for Africa. 2013 (<http://www.afro.who.int/en/clusters-a-programmes/frh/making-pregnancy-safer/features/3898-strategic-framework-for-the-elimination-of-new-hiv-infections-among-children-in-africa-by-2015.html>, accessed 19 June 2015).
- 16 The global elimination of congenital syphilis: Rationale and strategy for action. Geneva, World Health Organization. 2007 (<http://www.who.int/reproductivehealth/publications/rtis/9789241595858/en/>, accessed 19 June 2015).
- 17 Methods for surveillance and monitoring of congenital syphilis within existing systems. Geneva, World Health Organization. 2011 (<http://www.who.int/reproductivehealth/publications/rtis/9789241503020/en/>, accessed 19 June 2015).
- 18 Data quality assurance tool for program-level indicators. Chapel Hill, MEASURE Evaluation. 2007 (<http://www.cpc.unc.edu/measure/tools/monitoring-evaluation-systems/data-quality-assurance-tools>, accessed 19 June 2015).
- 19 LFA guidelines for on-site data verification (OSDV) and rapid service quality assessment (RSQA) implementation. Geneva, The Global Fund to Fight AIDS, Tuberculosis and Malaria 2014 (<http://www.theglobalfund.org/en/me/documents/dataquality/>, accessed 19 June 2015).

- 20 Spohr M, Bailey C, Boucher P. Data exchange standards for electronic medical records. Roundtable consultation on Standards and Challenges in Health Information Systems, August 28, 2007, Nyon, Vaud, Switzerland, Geneva, 2007 (<http://www.who.int/healthmetrics/library/RoundTable001DataExchangeStandardsForEMRBoucher.pdf>, accessed 19 June 2015).
- 21 Framework and standards for country health information systems. Geneva, World Health Organization Health Metrics Network. 2008 (http://www.who.int/healthmetrics/documents/hmn_framework200803.pdf, accessed 19 June 2015).
- 22 Joint Inter-Ministerial Policy Dialogue on eHealth Standardization and Second WHO Forum on eHealth Standardization and Interoperability. Geneva, World Health Organization 2014 (http://www.who.int/ehealth/events/standardization_forum_2014/en/, accessed 19 June 2015).
- 23 Guidelines on protecting the confidentiality and security of HIV information. Proceedings from a workshop: 15-17 May 2006, Geneva, Joint United Nations Programme on HIV/AIDS. 2007 (http://data.unaids.org/pub/Manual/2007/confidentiality_security_interim_guidelines_15may2007_en.pdf, accessed 19 June 2015).
- 24 Three interlinked patient monitoring systems for HIV care/ART, MCH/PMTCT (including malaria prevention during pregnancy), and TB/HIV: Standardized minimum data set and illustrative tools. Geneva, World Health Organization. 2012 (http://www.who.int/hiv/pub/me/patient_monitoring_systems/en/, accessed 19 June 2015).
- 25 Evaluating a national surveillance system. Geneva, World Health Organization. 2013 (<http://www.who.int/hiv/pub/surveillance/2013package/module7/en/>, accessed 19 June 2015).
- 26 United States Agency for International Development. Data quality assurance tools Chapel Hill, MEASURE Evaluation. 2015 (<http://www.cpc.unc.edu/measure/tools/monitoring-evaluation-systems/data-quality-assurance-tools>, accessed 19 June 2015).
- 27 Data quality report card. World Health Organization. 2015 (http://www.who.int/healthinfo/topics_standards_tools_data_quality_analysis/en/, accessed 19 June 2015).
- 28 Data quality assurance standards and tools for PMTCT programmes. Geneva, World Health Organization Interagency Task Team on the Prevention and Treatment of HIV Infection in Pregnant Women. forthcoming.
- 29 Monitoring the building blocks of health systems: A handbook of indicators and their measurement strategies. Geneva, World Health Organization. 2008 (http://www.who.int/healthinfo/systems/WHO_MBHSS_2010_full_web.pdf, accessed 19 June 2015).
- 30 Parekh BS, Kalou MB, Alemnji G, Ou CY, Gershy-Damet GM, Nkengasong JN. Scaling up HIV rapid testing in developing countries: Comprehensive approach for implementing quality assurance. *Am J Clin Pathol*. 2010;134(4):573-584 (<http://www.ncbi.nlm.nih.gov/pubmed/20855638>, accessed 19 June 2015).

- 31 HIV assays operational characteristics. HIV rapid diagnostic tests (detection of HIV-1/2 antibodies): Report 17. Geneva, World Health Organization. 2013 (http://www.who.int/diagnostics_laboratory/evaluations/hiv/131107_hiv_assays17_final.pdf, accessed 19 June 2015).
- 32 Consolidated guidelines on HIV Testing Services. 2015 (http://apps.who.int/iris/bitstream/10665/179870/1/9789241508926_eng.pdf?ua=1&ua=1), accessed 19 June 2015).
- 33 US Department of Health and Human Services. Guidelines for assuring the accuracy and reliability of HIV rapid testing: Applying a quality system approach. Geneva, World Health Organization and US Centers for Disease Control and Prevention. 2005 (http://whqlibdoc.who.int/publications/2005/9241593563_eng.pdf, accessed 19 June 2015).
- 34 United States Agency for International Development. The DHS program: Demographic and health surveys. Rockville, ICF International. (<https://dhsprogram.com/>, accessed 19 June 2015).
- 35 Zaba B, Gregson S. Measuring the impact of HIV on fertility in Africa. *AIDS*. 1998;12 Suppl 1:S41-50 (<http://www.ncbi.nlm.nih.gov/pubmed/9677188>, accessed 19 June 2015).
- 36 Yeatman SE. The impact of HIV status and perceived status on fertility desires in rural Malawi. *AIDS Behav*. 2009;13 Suppl 1:12-19 (<http://www.ncbi.nlm.nih.gov/pubmed/19301116>, accessed 19 June 2015).
- 37 Burgos-Soto J, Balestre E, Minga A, Ajayi S, Sawadogo A, Zannou MD et al. Incidence of pregnancy after antiretroviral therapy initiation and associated factors in 8 West African countries. *J Acquir Immune Defic Syndr*. 2014;67(2):e45-54 (<http://www.ncbi.nlm.nih.gov/pubmed/25216079>, accessed 19 June 2015).
- 38 Global plan towards the elimination of new HIV infections among children by 2015 and keeping their mothers alive Geneva, United Nations Joint Programme on HIV/AIDS. 2011 (http://www.unaids.org/sites/default/files/en/media/unaids/contentassets/documents/unaidspublication/2011/20110609_JC2137_Global-Plan-Elimination-HIV-Children_en.pdf, accessed 19 June 2015).
- 39 Gopalappa C, Stover J, Pretorius C. HIV prevalence patterns by age and sex: Exploring differences among 19 countries. Analytical Studies No. 40, Calverton, Maryland,, IFC International. 2013 (<http://www.dhsprogram.com/publications/publication-AS40-Analytical-Studies.cfm>, accessed 19 June 2015).
- 40 UNAIDS report on the global AIDS epidemic 2013. Geneva, Joint United Nations Programme on HIV/AIDS (UNAIDS). 2013 (http://www.unaids.org/sites/default/files/media_asset/UNAIDS_Global_Report_2013_en_1.pdf, accessed 19 June 2015).

- 41 Todd J, Glynn JR, Marston M, Lutalo T, Biraro S, Mwita W et al. Time from HIV seroconversion to death: A collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy. *AIDS*. 2007;21 Suppl 6:S55-63 (<http://www.ncbi.nlm.nih.gov/pubmed/18032940>, accessed 19 June 2015).
- 42 Gouws E, Lyerla R, Garcia-Calleja JM, Barrerre B, Serrano D, Velasquez C et al. Trends in HIV prevalence and sexual behaviour among young people aged 15-24 years in countries most affected by HIV. *Sexually Transmitted Infections*. 2010;86:172-183 (<Go to ISI>://WOS:000284572700013, accessed 19 June 2015).
- 43 Mahy M, Garcia-Calleja JM, Marsh KA. Trends in HIV prevalence among young people in generalised epidemics: Implications for monitoring the HIV epidemic. *Sexually Transmitted Infections*. 2012;88:165-175 (<Go to ISI>://WOS:000311338000010, accessed 19 June 2015).
- 44 Ferrand RA, Corbett EL, Wood R, Hargrove J, Ndhlovu CE, Cowan FM et al. AIDS among older children and adolescents in Southern Africa: Projecting the time course and magnitude of the epidemic. *Aids*. 2009;23(15):2039-2046 (<Go to ISI>://WOS:000270475400013, accessed 19 June 2015).
- 45 Eaton JW, Garnett GP, Takavarasha FR, Mason PR, Robertson L, Schumacher CM et al. Increasing adolescent HIV prevalence in Eastern Zimbabwe: Evidence of long-term survivors of mother-to-child transmission? *PLoS One*. 2013;8(8):e70447 (<http://www.ncbi.nlm.nih.gov/pubmed/23950938>, accessed 19 June 2015).
- 46 Marston M, Becquet R, Zaba B, Moulton LH, Gray G, Coovadia H et al. Net survival of perinatally and postnatally HIV-infected children: A pooled analysis of individual data from sub-Saharan Africa. *Int J Epidemiol*. 2011;40(2):385-396 (<http://www.ncbi.nlm.nih.gov/pubmed/21247884>, accessed 19 June 2015).
- 47 HIV and adolescents: Guidance for HIV testing and counselling and care for adolescents living with HIV: Recommendations for a public health approach and considerations for policy-makers and managers. Geneva, World Health Organization. 2013 (http://www.youngpeopleandhiv.org/files/HIV_Testing_guideline.pdf, accessed 19 June 2015).
- 48 Guidelines for using HIV testing technologies in surveillance: Selection, evaluation and implementation—2009 update. Geneva, World Health Organization. 2009 (http://www.who.int/hiv/pub/surveillance/hiv_testing_technologies_surveillance.pdf, accessed 19 June 2015).
- 49 London School of Hygiene and Tropical Medicine. The rapid syphilis test toolkit: A guide to planning, management, and implementation. London, London School of Hygiene and Tropical Medicine. 2013 (https://www.lshtm.ac.uk/itd/crd/research/rapidsyphilistoolkit/rapid_syphilis_test_toolkit.pdf, accessed 19 June 2015).

- 50 Thacker SB, Berkelman RL. Public health surveillance in the United States. *Epidemiol Rev.* 1988;10:164-190 (<http://www.ncbi.nlm.nih.gov/pubmed/3066626>, accessed 19 June 2015).
- 51 Nelson DE, Hesse BW, Croyle RT. *Making data talk: Communicating public health data to the public, policy makers, and the press.* Oxford, New York, Oxford University Press. 2009 (Table of contents only <http://www.loc.gov/catdir/toc/ecip0824/2008033685.html>, accessed 19 June 2015).



For more information, contact:

World Health Organization
Department of HIV/AIDS
20, avenue Appia
1211 Geneva 27
Switzerland

E-mail: hiv-aids@who.int

www.who.int/hiv

ISBN 978 92 4 150972 5



9 789241 509725