HIV DRUG RESISTANCE

CONCEPT NOTE

SURVEILLANCE OF HIV DRUG RESISTANCE IN ADULTS INITIATING ANTIRETROVIRAL THERAPY (PRE-TREATMENT HIV DRUG RESISTANCE)

JULY 2014

World Health Organization

SURVEILLANCE OF HIV DRUG RESISTANCE IN ADULTS INITIATING ANTIRETROVIRAL THERAPY (PRETREATMENT HIV DRUG RESISTANCE)

JULY 2014

World Health Organization

WHO Library Cataloguing-in-Publication Data

Surveillance of HIV drug resistance in populations initiating antiretroviral therapy
(pre-treatment HIV drug resistance)

1.Anti-HIV agents – therapeutic use. 2.Drug resistance, Viral – drug effects. 3.HIV infections – epidemiology. 4.HIV infections – drug therapy 5.Population surveillance. I.World Health Organization.

ISBN 978 92 4 150719 6 (NLM classification: QV 268.5)

Layout by L'IV Com Sàrl, Villars-sous-Yens, Switzerland.

Printed by the WHO Document Production Services, Geneva, Switzerland.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

The work was led by Silvia Bertagnolio (World Health Organization), Jhoney Barcarolo (consultant) and Michael Jordan (Tufts University School of Medicine, USA). Statistical Methods were developed by Natalie Exner Dean and Marcello Pagano (Harvard University, USA). Natalie Exner Dean wrote the statistical annexes and analysis plan. The concept note received substantial technical inputs and revisions by the HIVResNet Core Group (Emiliano Bissio, Argentina; Avelin Aghokeng, Cameroun; Elliot Raizes, USA; and Andrea De Luca, Italy).

Jos Perriens, coordinator of the Technologies and Commodities Team at the HIV Department, World Health Organization, provided overall supervision. We thank the United States Centers for Disease Control and Prevention, Atlanta, USA and the President's Emergency Plan for AIDS Relief (PEPFAR) HIV Drug Resistance working group for their critical inputs to the document.

We are most grateful to the HIVResNet Steering Group (Chunfu Yang, USA; Chris Archibald, Canada; Achara Teeraratkul, Thailand; Avelin Aghokeng, Cameroun; AM.J. Wensing, the Netherlands; Christopher Duncombe, USA; Kerry Dierberg, USA; Do Thi Nhan, Vietnam; Emiliano Bissio, Argentina; Elliot Raizes, USA; Jonathan Shapiro, Israel; Natalia Nizova, Ukraine; Robert W. Shafer, USA; Wilford Kirungi, Uganda; Zhang Fujie, China; Tobias Rinke de Wit, the Netherlands; and Sergio Carmona, South Africa) for their strategic advice. We are also grateful to the contributions provided by WHO staff from regional and country offices.

Additionally, development of this guidance document benefitted greatly from inputs received from country ART programme managers and country representatives attending WHO regional workshops conducted in 2013 in Beijing, China; Brasilia, Brazil; Montpelier, France; Addis Ababa, Ethiopia; and Cape Town, South Africa. WHO is grateful to Amandine Cournil, Montpelier, France, for her review of the statistical annexes. Development and publication of this guidance was supported by the Bill and Melinda Gates Foundation (#38180); we are sincerely grateful for their support.

ACRONMYS

ADR	acquired HIV drug resistance
AIDS	acquired immunodeficiency syndrome
ART	antiretroviral therapy
ARV	antiretroviral drug
ATV/r	atazanavir/ritonavir
AZT	zidovudine (also ZDV)
DEFF	design effect
DBS	dried blood spot
DRV/r	darunavir/ritonavir
EFV	efavirenz
HIV	human immunodeficiency virus
ICC	intracluster correlation coefficient
HIVDR	HIV drug resistance
INI	integrase inhibitor
LPV/r	lopinavir/ritonavir
mL	milliliter
NNRTI	non-nucleoside reverse transcriptase inhibitor
NRTI	nucleoside reverse transcriptase inhibitor
N(t)RTI	nucleotide reverse transcriptase inhibitor
NVP	nevirapine
PDR	pre-treatment HIV drug resistance
PEP	post exposure prophylaxis
PEPFAR	President's Emergency Plan for AIDS Relief
PI	protease inhibitor
PPS	probability proportional to size
PPPS	probability proportional to proxy size
PR	protease region
PrEP	pre-exposure prophylaxis
RT	reverse transcriptase region
SI	sampling interval
SID	survey identification number
VL	viral load
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNGASS	United Nations General Assembly Special Session
WHO	World Health Organization

BACKGROUND

HIV Drug Resistance (HIVDR) emerges when HIV replicates in the presence of antiretroviral drugs. If HIV drug resistance becomes widespread, the drugs currently used to treat HIV infection may become ineffective. To date, levels of HIVDR in countries scaling up antiretroviral therapy (ART) remain manageable. However, resistance is slowly increasing. In East Africa, resistance rates of 10% to non-nucleoside drugs (such as nevirapine and efavirenz) have been recently described.

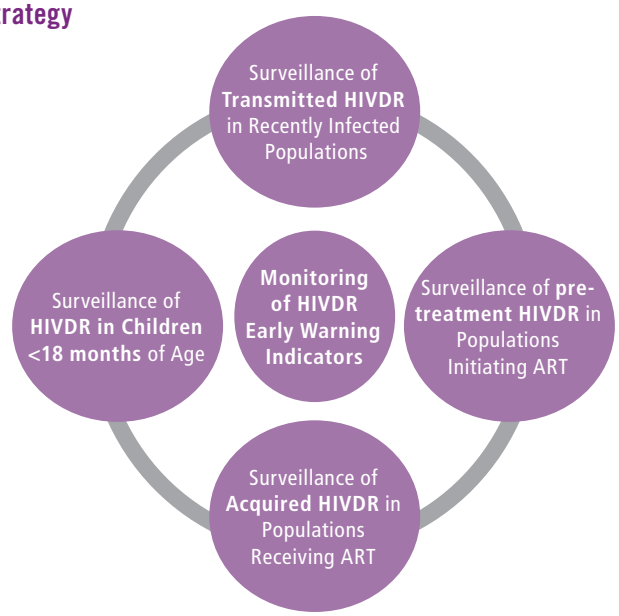
To maximize the long-term effectiveness of first-line ART and ensure the sustainability of ART programmes, it is essential to minimize the further spread of HIV drug resistance. Even in settings with optimal ART programme management, some degree of HIVDR is expected to emerge in populations on ART, and some HIVDR is expected to be transmitted to previously uninfected individuals. Therefore, WHO recommends that HIV treatment scale-up should always be accompanied by a robust assessment of drug resistance emergence and transmission. WHO's HIVDR monitoring and surveillance strategy is composed of five key elements:

i. Monitoring of Early Warning Indicators of HIV drug resistance
ii. Surveillance of HIVDR in recently-infected adult populations (transmitted HIVDR)
iii. Surveillance of pre-treatment HIVDR in adult populations initiating ART (pre-treatment HIVDR)
iv. Surveillance of acquired HIVDR in populations of adults and children receiving ART (acquired HIVDR)
v. Surveillance of HIV drug resistance in treatment-naive children less than 18 months of age

WHO's HIVDR Surveillance and Monitoring Strategy is a critical component of the public health approach to ART delivery. By obtaining population-level data on HIVDR in different populations, its various elements can inform programme-level decision making regarding, for example, optimal first and second lines, for both children and adults.

This document describes methods to assess HIVDR in adult populations about to initiate ART (surveillance of pre-treatment HIVDR).

Figure 1. HIV Drug Resistance Surveillance and Monitoring Strategy



1. INTRODUCTION

Nationally representative surveillance of HIVDR in populations initiating a standard triple-drug ART combination is critical to inform the selection of effective first-line ART combinations, as well as adequate pre-exposure prophylaxis (PrEP) and post-exposure prophylaxis (PEP) regimens. Detected HIVDR in populations initiating triple-drug ART regimens may have been transmitted at the time of initial infection or acquired due to previous exposure to antiretroviral (ARV) drugs (in the context of prevention of mother-to-child transmission (PMTCT) programmes, PrEP, PEP or previous disclosed or undisclosed ART).

2. SURVEY PURPOSE AND OVERVIEW

The purpose of this survey is to calculate (a) a nationally representative prevalence estimate of HIVDR among all ART initiators, and (b) a nationally representative prevalence estimate of HIVDR among initiators without prior exposure to ARV drugs (see Box 1). The size of the survey sample size (discussed in Section 4.4) has been calculated to provide sufficient statistical precision for the latter prevalence estimate. Operationally, in order to facilitate implementation, the survey will enrol all eligible individuals initiating ART during a predetermined period (see section 4.3 for patient eligibility criteria). At the same time, information on prior ARV exposure will be obtained at the time of specimen collection, and these data will then be used to stratify the sample and calculate the outcomes of interest (see Section 3 below).

Box 1: Why is it important to differentiate populations initiating ART with and without prior ARV exposure?

It is important to differentiate between populations initiating ART with and without prior ARV exposure as higher levels of HIVDR are anticipated among ART initiators with prior ARV exposure, and this population would be expected to contribute disproportionately to the level of observed resistance in a combined analysis. Critically, such a bias could lead to inaccurate conclusions and inappropriate public health policy-making. For example, high levels of resistance among women exposed to PMTCT could warrant changes in recommended first-line regimens (for example, switching from a non-nucleoside reverse transcriptase (NNRTI)-based to a protease inhibitor (PI)-based regimen) for this subgroup, but not necessarily for all patients initiating ART, especially those with no prior ARV exposures. Moreover, information on HIVDR among ART initiators without prior ARV exposure is key to inform the selection of optimal regimens for PrEP and PEP.

Prior ARV exposure may be ascertained through a variety of methods, including the application of a screening questionnaire and/or the review of patient medical records, where available and feasible. Initiators will be classified into one of the following three categories of prior ARVs exposure: yes, no, unknown. Countries should decide a *priori* which method(s) will be applied to identify patients with and without prior ARV exposure. If prior ARV exposure is identified, it should be characterized as to whether it refers to previous ART for one's own health, PrEP, PEP, PMTCT or a combination of exposures. This information may be used in a descriptive analysis at the national level and may be aggregated across surveys to generate regional/global estimates by type of ARV exposure.

3. SURVEY OUTCOMES

The survey has six main outcomes. Outcomes 1a, 1b and 1c are related to the prevalence of HIV drug resistance:

1a. Prevalence of HIVDR among all ART initiators, regardless of prior exposure to ARVs

Outcome 1a measures the proportion of any HIVDR[1,2] among all ART initiators, regardless of prior exposure to ARVs.

1b. Prevalence of HIVDR among ART initiators without prior exposure to ARVs

Outcome 1b measures the proportion of any HIVDR1 among ART initiators *without* prior ARVs drug exposure.

1c. Prevalence of HIVDR among individuals initiating ART with NNRTI-based regimens without prior exposure to ARVs

Outcome 1c measures the proportion of any HIVDR among individuals initiating NNRTI-based ART regimens *without* prior ARVs drug exposure.

In countries in which all patients initiate NNRTI-based regimens, Outcomes 1b and 1c will be identical. In some countries with more mature ART programmes, such as those in Latin America and the Caribbean, where ART has been available for over a decade, an important proportion of patients now initiate ART with non-NNRTI-based combinations, such as PI-based regimens. In such cases, Outcomes 1b and 1c should be calculated separately and the sample size must be further adjusted to account for the proportion of individuals receiving non-NNRTI-based regimens (see Box 2).

These three outcomes must be calculated taking into account observed clinic-level patient accrual rates and the actual number of ART initiators with sequences genotyped at each clinic. The analysis will account for these elements through adjustments of the survey weights (an example of a data analysis plan is provided in Annex 1.4 and additional technical background is available in the Statistical Appendix).

2a. Proportion of all ART initiators without prior exposure to ARVs

Outcome 2a measures the proportion of patients initiating ART without prior exposure to ARVs.

2b. Proportion of all ART initiators with prior exposure to ARVs

Outcome 2b measures the proportion of patients initiating ART with prior exposure to ARVs.

2c. Proportion of all ART initiators with unknown prior exposure to ARVs

Outcome 2c measures the proportion of patients initiating ART with unknown prior exposure to ARVs.

Outcomes 2a, 2b and 2c must be calculated taking into account observed clinic-level patient accrual rates. The analysis will account for these elements through adjustments of the survey weights.

1 Any HIV drug resistance is defined with respect to one or more of the following drugs or drug classes: NVP, EFV, any N(t)RTI, DRV/r, LPV/r or ATV/r. Sequences classified as low-level, intermediate or high-level resistance according to the Stanford HIVdb are aggregated as "HIV drug resistance".

2 In countries opting to genotype the integrase (IN) region of HIV, detected integrase inhibitor (INI) resistance is excluded when estimating the prevalence of any HIVDR.

4. OVERVIEW OF METHODS

4.1 Survey approach

This is a cross-sectional survey. For reasons of practicality, it is suggested that the survey be limited to a maximum of six months. For the purpose of this concept note the words "clinic" and "site" are used interchangeably. The survey is carried out in two stages[1]:

- In the first stage, a minimum of 15 clinics is selected (sampled) from a list of all clinics that initiate ART in the country. The number of clinics to be sampled is a country decision and should take into account the ability to enrol enough participants (and meet the required sample size) and the logistic issues associated with sampling a larger number of clinics.
- In the second stage, consecutive eligible patients initiating ART on or after a pre-defined survey start date are enrolled until the predetermined sample size for each clinic is achieved. Except in special situations (discussed later in this document), each clinic should contribute to the overall sample size with the same number of patients.

4.2 How to select clinics

To achieve a nationally representative estimate of resistance prevalence, the selection of clinics must be representative of all clinics which initiate ART in the country.

The first step in the selection of clinics is to create a *sampling frame*, also called a sampling table, which is a list of all ART clinics in the country where patients initiate treatment (as discussed below), alongside their respective sizes. The most accurate information to estimate clinic size is the number of treatment initiators during a recent time period in each clinic (such as the previous twelve months). Should this information not be available, countries can utilize the numbers of patients on ART in each clinic at the end of a determined previous time period (for example, at the end of the year prior to survey initiation). Once the *sampling frame* has been established, clinics must then be sampled.

If countries have information on the number of people who initiated ART at each clinic in the previous time period (for example, the previous six months), the recommended method to sample clinics is called *probability proportional to size (PPS) sampling*. In PPS sampling, the probability that a clinic is sampled is proportional to the number of treatment initiators observed at that clinic during a recent time period. In practice, this means that clinics with many treatment initiators are more likely to be sampled than clinics with fewer treatment initiators.

If the only information available is the total number of patients on ART at the end of a previous calendar year by clinic, another sampling method, known as *probability proportional to proxy size (PPPS) sampling*, must be used. In this case, the number of patients on ART at the end of a previous time period by clinic is used as a proxy for relative clinic size.

Whenever possible, PPS sampling is the preferred design because it is the most statistically efficient and thus requires fewer people to participate in the survey. Additional technical explanations are provided in the Statistical Appendix.

4.2.1 Sampling very small or difficult-to-access clinics

Some countries may have a number of clinics with extremely small populations of patients initiating ART or clinics that may be difficult to access for a variety of reasons, including political instability or geographical remoteness. Although not advisable, some countries may consider excluding some of these clinics from the systematic sampling table due to logistic and under-enrolment issues.

In general, if less than 10% of the population initiates at these clinics, countries may choose to exclude them from the systematic sampling table. This threshold seeks to limit the potential bias that such exclusion may introduce in the final results. In this case, exclusion of the clinic(s) from the systematic sampling table should be done a *priori* (and not after the clinic has been sampled). A list of all excluded clinics and reasons for their exclusion should be reported in any resulting technical report. On the other hand, if more than 10% of the population of interest initiates treatment at these clinics, it is not advisable to exclude them from the pool of clinics that can be sampled for the survey.

For example, suppose a country excluded difficult-to-access clinics that represented 10% of the population of treatment initiators. In this country, the prevalence of PDR at the remaining clinics in the sampling table was 10%. If the prevalence of PDR among the excluded, difficult-to-access to clinics was also 10%, then the true prevalence of PDR in the entire population was 10%; thus, excluding these clinics did not introduce bias into the national prevalence estimate. If the prevalence of PDR among the excluded, difficult-to-access clinics was much higher, for example 30%, then the true prevalence of PDR in the entire population was 12%. The survey among the clinics in the sampling frame would underestimate the prevalence of PDR (10% versus

1 This method of identifying study participants is known as a two-stage cluster design.

12%), though the magnitude of this bias is minor when the proportion of the population initiating therapy excluded is low (<10%) and the difference in prevalence between the included and excluded populations is small.

In general, if the excluded patients have a different prevalence of pre-treatment HIVDR than the observed patients, the national prevalence estimate will be biased.

4.2.2 Regional representation

Countries wishing to develop precise region-specific HIVDR estimates should implement a **separate survey** in each area of interest. However, if the goal is not to make region-specific inferences, but to balance administrative load and achieve regional representation, this can be achieved using a technique called *implicit stratification* by region. Operationally, this entails listing all clinics by administrative or geographical area prior to their selection (see Annex 1.1). Clinics will be sampled approximately proportionally to the size of the region. This method does not allow for the development of precise region-specific estimates nor does it always guarantee the geographical representation of all regions. If a country wishes to guarantee the representation of at least one clinic from each area of interest, this can be achieved using the method described in Annex 1.2.

If knowledge of resistance prevalence by other characteristics (for example, rural versus urban) is relevant for national decision-making, countries may consider *formal stratification* at the design phase (see Annex 1.3).

4.2.3 Countries with many ART clinics

Countries with a large number of clinics (for example, over 1000) are not required to sample more than 40 clinics to achieve a representative national prevalence estimate assuming that the level of heterogeneity across clinics is consistent with the available global data. Nonetheless, larger countries may have more heterogeneity across clinics than smaller countries. As a result, it is recommended that larger countries sample no fewer than 30 clinics, and they may prefer to sample a greater number for a more precise conclusion if resources are available. Alternatively, if larger countries would like to make region-specific statements with a particular precision, then they should consider conducting a separate survey in each region. These then typically lead to more precise national numbers when combined across regions.

4.3 Patient eligibility criteria

4.3.1 Inclusion criteria

- Adults[1] with HIV-1 infection[2] who can legally provide and do provide informed consent, and options B/B+)[3]
- All individuals initiating ART (including as first-line treatment of their own health or PMTCT
  - » for the first time
  - » reinitiating if they have stopped for more than three months[4].

4.3.2 Exclusion criteria

- Patients transferred-in already receiving ART
- In countries where routinely used antibody tests differentiate between HIV-1 and HIV-2, adults infected with HIV-2 or individuals with HIV-1/HIV-2 coinfection are excluded

4.4 Defining the survey sample size

4.4.1 Assumptions

Table 1 summarizes the critical assumptions which affect the calculation of the survey sample size, with their assumed values as used in the standard calculations presented in Section 4.4.2.

Table 1: Key model assumptions to calculate the sample size for Outcomes 1a, 1b and 1c

Assumptions	Proposed values
Expected prevalence of drug resistance among all ART initiators	10%
Expected genotyping failure rate	20%
Expected proportion of ART initiators without prior exposure to ARVs	75%
Expected proportion of individuals initiating ART with NNRTI-based regimens	100%
Desired confidence interval half-width	5%

1 Adults are generally defined as being 18-years-old and above, though the minimum age may be country specific.

2 HIV-2 infection is naturally resistant to NNRTIs, and thus requires different treatment approaches. WHO Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: Recommendations for a public health approach. Geneva, World Health Organization, 2013.

3 Option B+ is defined as: providing lifelong ART to all pregnant and breastfeeding women living with HIV regardless of CD4 count or clinical stage. Option B is defined as providing ART for pregnant and breastfeeding women with HIV during the mother-to-child transmission risk period and then continuing lifelong only for those women eligible for treatment for their own health.

4 Individuals who have stopped ART for less than three months are still deemed to be on ART and should not be enrolled in the survey.

With respect to the expected prevalence of drug resistance among treatment initiators, available evidence suggests that it is reasonable to conservatively assume an estimated prevalence of HIVDR among treatment initiators of 10%. This figure, which is greater than the 5% generally reported in the literature, including in the 2012 WHO HIV drug resistance report, is used as a conservative measure of expected HIVDR prevalence because higher levels of pre-treatment HIVDR – approximating 10% – have been documented in some regions.[1,2] If the observed prevalence of drug resistance is lower than 10%, the resulting estimate will have a narrower confidence interval.

The expected genotyping failure rate is assumed to be 20%.[3]

The expected proportion of ART initiators without prior ARV exposure is assumed to be 75%. It is not advisable to change this value except if comprehensive and recent national data are available from patient information systems or a previous PDR survey.

The expected proportion of individuals sampled initiating NNRTI-based regimens is assumed to be 100%. This assumption reflects the most common situation in countries using the public health approach. However, in some countries, a proportion of patients initiating ART are prescribed non-NNRTI-based combinations (for example, PI-based regimens). In such cases, the sample size must be further inflated to account for the proportion of individuals receiving non-NNRTI-based regimens. This assumption may be changed as countries adapt this concept note into national protocols.

A confidence interval of half-width of ±5% is suggested as an appropriate compromise between feasibility and precision.

4.4.2 Sample size calculations

In addition to the anticipated resistance prevalence, the desired precision of the estimate and the genotyping failure rate, the total sample size required for the survey is also affected by the number of clinics to be sampled. In general, as more clinics are sampled, better representation of the prevalence of resistance across clinics is achieved. For logistic or financial reasons, countries may wish to limit the number of clinics sampled. However, it is recommended to sample a minimum of 15 clinics to obtain a nationally representative estimate. Due to evidence that pre-treatment HIVDR outcomes cluster by clinic, it is better to sample more clinics than sampling additional patients within a clinic. As more clinics are sampled, better representation of the prevalence of pre-treatment HIVDR is achieved. While the total sample size required approximately stabilizes after the number of clinics sampled reaches 20, the number of ART initiators to be sampled per clinic continues to fall as more clinics are included in the survey. This may be particularly important to improve the survey's feasibility in countries where the average number of ART initiators per clinic is small. The survey design effect is most significant when fewer clinics are sampled (see Statistical Appendix)[5].

Table 2 provides sample sizes for countries intending to sample a subset of their ART clinics using PPS or PPPS sampling, based on the assumptions outlined in Table 1.

Box 2: Adjusting the sample size in countries where patients may initiate ART with non-NNRTI-based combinations

The calculations presented in section 4.4.2 assume that the vast majority of individuals will initiate ART with a standard NNRTI-based regimen, as recommended by WHO in the context of the public health approach to ART[4]. However, in countries where individuals may initiate ART with PI-based regimens, the sample size must be further adjusted to account for the proportion of individuals initiating ART with non-NNRTI-based regimens. All individuals initiating ART are eligible to be enrolled, irrespective of their prior ART history. To achieve a certain precision for the confidence interval around the estimate of resistance among ART initiators without prior ARV exposure, the sample size must be adjusted to maximize the likelihood of enrolling a sufficient number of patients in this category. For example, if in Country X only 75% of individuals initiate ART with a NNRTI-based combination, and the unadjusted sample size is 300, this should be further increased appropriately by dividing it by 0.75 to arrive at a sample size of approximately 400. The use of the Excel-based tool discussed in Section 5.5 can perform this calculation with the appropriate rounding. Countries in this situation should report, in addition to Outcomes 1a and 1b, HIVDR prevalence among individuals starting NNRTI-based ART without prior ARV exposure (Outcome 1c).

1 WHO HIV drug resistance report 2012. Geneva, World Health Organization, 2012.

2 Hamers et al. (2011). HIV-1 drug resistance in antiretroviral-naive individuals in sub-Saharan Africa after rollout of antiretroviral therapy: a multicentre observational study. Lancet Infect Dis. 11(10):750-9. doi: 10.1016/S1473-3099(11)70149-9. Epub 2011 Jul 27.

3 This is based on data from the 2012 WHO HIV drug resistance report and includes an estimated 10% of individuals initiating ART with VL ≤1000 copies/ml (Italian Cohort of Antiretroviral Naïve Patients (ICONA), A. de Luca, personal communication).

4 Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendation for a public health approach. Geneva, World Health Organization, 2013.

5 The sample sizes are also impacted by the fact that the per-clinic sample size must be a whole number, and this is then multiplied by the number of clinics to yield the total sample size. Variability in how the per-clinic sample sizes are rounded can result in sample sizes that stay constant or slightly increase with additional clinics.

The standard sample sizes described in Table 2 were calculated assuming an extremely large number of sites and patients initiating ART. However, as countries develop their national protocols, sample sizes can be tailored to their local circumstances by using national data on (i) the number of people who initiated ART in the year prior to the survey and (ii) the number of ART sites in the sampling table. This adaptation process will result in a decrease in the estimated total sample size required to reach the same confidence interval. This is particularly relevant for countries with few patients initiating ART, such as settings characterized by low HIV prevalence or concentrated HIV epidemics. Nevertheless, the figures presented in Table 2 provide standard sample sizes (based on the assumptions discussed in Section 4.4.1) that can be used for budgetary and planning purposes.

Countries unable to achieve this sample size can consider increasing the confidence interval width to ±6% using the Excel-based calculator discussed in Section 5.5.

Table 2: Standard sample size calculations, survey of pretreatment HIVDR in populations initiating ART

	PPS sampling (information required: number of ART initiators by clinic during a six-month period[1])		PPPS sampling (information required: number of people on ART by clinic at the end of previous calendar year (for example, end of 2013)	
Number of clinics to be sampled	Number of ART initiators to be sampled per clinic	Total number of ART initiators to be sampled	Number of ART initiators to be sampled per clinic	Total number of ART initiators to be sampled
15	23	345	34	510
16	22	352	31	496
17	20	340	28	476
18	19	342	26	468
19	17	323	24	456
20	16	320	23	460
21	15	315	21	441
22	14	308	20	440
23	14	322	19	437
24	13	312	18	432
25	13	325	17	425
26	12	312	17	442
27	11	297	16	432
28	11	308	15	420
29	11	319	15	435
30	10	300	14	420
31	10	310	14	434
32	9	288	13	416
33	9	297	13	429
34	9	306	12	408
35	9	315	12	420
36	8	288	11	396
37	8	296	11	407
38	8	304	11	418
39	8	312	10	390
40	8	320	10	400

Notes: assumed pretreatment HIVDR prevalence = 10%, confidence interval width = ±5%, laboratory failure rate = 20%, assumed proportion of ART initiators with (or unknown) prior ARV exposure = 25%, assumed proportion of individuals initiating ART receiving NNRTI-based regimens = 100%.

1 This refers to the number of ART initiators in a 12-month period divided by 2.

4.4.3 Countries sampling all clinics where ART is initiated

It is recommended that countries with few ART clinics sample all of them. For countries sampling all ART clinics, the standard total sample size, using the same assumptions discussed in Table 1, and based on PPS sampling, is 254. If information on ART initiators per clinic is not available and PPPS sampling is used, then the standard total sample size is 346. The number of sampled patients at each clinic should be proportional to the size of the clinic (that is, the number of eligible patients at that clinic). For example, if 25% of patients initiate treatment at one clinic in the country, 25% of the actual sample size should be collected at this clinic.

This standard total sample size was calculated assuming an infinitely large population size. Countries can tailor their sample size calculations to their local circumstances by using national data on the number of people who initiated ART in the year prior to the survey. This adaptation process will result in a decrease in the estimated total sample size required to reach the same confidence interval.

4.5 Laboratory methods

4.5.1 Specimen collection, handling, processing and tracking

Dried blood spot (DBS) or plasma can be used as the specimen type for this survey. DBS has been shown to be a reliable specimen type for HIVDR genotyping.[1] DBS specimens should be collected and handled according to the WHO Guidance for DBS specimen collection and handling for HIV drug resistance testing.[2] Countries using plasma specimens for this survey should refer to the WHO recommendations on plasma collection, processing and storage for HIVDR testing.[3]

4.5.2 HIVDR genotyping and quality assurance of sequences

Specimens collected for PDR surveillance should be tested in WHO-designated HIVDR genotyping laboratories. These laboratories are members of the WHO HIVResNet Laboratory Network, undergo a rigorous inspection process and participate in annual proficiency panel testing. Use of WHO-designated laboratories promotes quality assured results for the purpose of public health surveillance. If a country does not have a WHO-designated laboratory for HIVDR testing, it is encouraged to send specimens to a WHO-designated regional or specialized laboratory. A list of WHO-designated laboratories may be found on the WHO HIV drug resistance webpage.[4]

Designated laboratories perform extensive quality assurance of sequences and follow the WHO Laboratory Standard Operating Procedures (SOP) for Post-Testing Quality Assurance of HIV Drug Resistance Genotyping. This SOP outlines steps for standardized and automated chromatogram interpretation using Web Recall, quality assurance using MEGA and additional quality assurance and HIVDR interpretation using Stanford HIVdb. This document will be available on the WHO website.

1 Bertagnolio S et al., Dried blood spots for HIV-1 drug resistance and viral load testing: A review of current knowledge and WHO efforts for global HIV drug resistance surveillance. AIDS Rev. 2010. 12(4):195-208.

2 http://www.who.int/hiv/topics/drugresistance/dbs_protocol.pdf

3 http://www.who.int/hiv/pub/drugresistance/hiv_reslab_strategy.pdf

4 http://www.who.int/hiv/topics/drugresistance/en/

5. IMPLEMENTATION CONSIDERATIONS

5.1 Duration of the survey, patient screening and sampling

To ensure results are available to decision-makers in a timely fashion, it is preferable to limit the duration of patient sampling to a maximum of six months. Operationally, once ART clinics to be included in the survey have been selected, a convenient starting date is chosen. Staff at the selected clinics should then screen all patients attending the clinics for the eligibility criteria discussed in Section 4.3.

For eligible patients, the survey should proceed in two steps:

1. Step 1 (minimum information step also referred to as "screening period"): obtain verbal consent and collect minimum information from all ART initiators though a screening questionnaire- clinic ID, patient ID, prior ARV exposure (Yes/no/unknown) and type of ARV exposure (mark all that apply: PMTCT, prior ART for one's health, PrEP, PEP, unspecified).
2. Step 2 (specimen collection also referred to as "enrolment period"): obtain full consent, collect the remaining necessary information (see Section 5.2 below) and obtain blood specimen.
3. In clinics where the required number of specimens is obtained in less than three months, Step 1 should continue to be applied until the end of the third month from survey initiation.

Information on prior exposure to ARVs can be obtained using a screening questionnaire if it is not captured routinely in patient medical records. After an eligible patient is identified and consent is obtained, a unique de-identified HIVDR survey identifier (PDR-SID) should be assigned to the patient (Box 4). Data abstracted by clinic staff should never include identifying information.

All eligible patients should be consecutively enrolled until the required sample size per clinic is reached or until the maximum enrolment period of six months has passed. If the desired number of specimens in a particular clinic is obtained before six months, specimen collection ("enrolment period") can stop at that clinic. However, if the desired number of specimens is achieved at a particular clinic before three months have elapsed since survey initiation, this clinic must continue to apply the screening questionnaire ("screening period") to all ART initiators for a minimum of three months. Information obtained in the screening period is essential to establish reliable estimates of relative clinic sizes, patient accrual rates and of the prevalence of prior ARV exposure, data which will be used at the analysis stage and for planning the next round of HIVDR surveillance.

Box 3: Important consideration prior to specimen collection

On the date of enrolment, a blood specimen should be collected from the patient prior to ART initiation (or reinitiation) at the clinic.

Box 4: Convention for assigning patient identification numbers

Individuals enrolled in the survey will be assigned a survey identification (SID) number, or unique survey ID. This number will be used to identify the patient as well as the sequence generated by the genotyping assay and is composed of the following five elements delimited by a dash character ("-"):

- country abbreviation: the ISO standard 3-letter abbreviation[1]
- survey type: PDR
- year survey started
- site abbreviation (a 3-letter abbreviation for the site, unique within the country; by default, the first three letters of the site name unless this is not unique)
- 4-digit unique patient number, that is, a consecutive unique patient number assigned to a participant at that site.

For example, if the "University HIV Clinic" was a site that participated in a national survey of PDR in South Africa in 2014, a participant's PDR-SID would be: ZAF-PDR-2014-UHC-0001

1 http://www.worldatlas.com/aatlas/ctycodes.htm

5.2 List of variables to be collected

5.2.1 Patient-level information

5.2.1.1 Patient-level epidemiological and laboratory variables to be captured for all patients who will have blood drawn for genotyping

This section describes the minimal set of patient information that must be captured in the survey database. Some will be obtained using a questionnaire applied to patients at the time of enrolment, while others will come from laboratory records. Once eligible patients have been identified, the following information must be captured for all patients *who will have blood drawn* for genotyping:

- Clinical/demographic information
  - i. Clinic ID
  - ii. Patient ID (see Box 4 for identification convention)
  - iii. Prior ARV exposure (yes/no/unknown)
  - iv. Type(s) of prior ARV exposure (mark all that apply: PMTCT, prior ART for one's health, PrEP, PEP, unspecified, more than one type of previous exposure)
  - v. If prior exposure to PMTCT, specify type of prophylaxis (A/B/B+/other/unknown)
  - vi. Age
  - vii. Gender (female/male/other)
  - viii. CD4 cell count before ART initiation (if results are already available and are from a test performed less than six months prior to survey enrolment)
  - ix. Regimen prescribed: list drugs prescribed

- Laboratory information
  - x. Specimen ID (see Box 4 for identification convention)
  - xi. Reverse transcriptase (RT) region of pol gene successfully sequenced[1]? (yes/no)
  - xii. Protease (PR) region of pol gene successfully sequenced (yes/no/not applicable[2])
  - xiii. Integrase (IN) region of pol gene successfully sequenced (yes/no/not applicable)[3]
  - xiv. Drug resistance (see also Section 6). For all drugs, choose the appropriate level according to the Stanford database algorithm interpretation: susceptible, potential low-level , low-level, intermediate or high-level resistance

5.2.1.2 Minimum information to be captured for all eligible patients for at least the first three months from survey initiation

As discussed in Section 5.1, if the recruitment objective in a particular clinic is reached before the six-month limit, specimen collection can stop at that clinic. However, if this happens within the first three months of survey initiation, these clinics should continue to count and apply the screening questionnaire regarding prior ARV exposure (minimum information step) to all ART initiators for a minimum of three months from the survey start date. In this case, the following minimum information should be collected from ART initiators until the end of the third month:

- i. Clinic ID
- ii. Patient ID
- iii. Prior ARV exposure (yes/no/unknown)
- iv. Type(s) of prior ARV exposure (mark all that apply: PMTCT, prior ART for one's health, PrEP, PEP, unspecified)
- v. If prior exposure to PMTCT, specify type of prophylaxis (A/B/B+/other/unknown)

5.2.2 Clinic-level information

In addition to individual patient-level information, the following information should be collected for each clinic included in the survey:

- i. Clinic name
- ii. Clinic ID
- iii. Date of survey initiation (date when patient screening starts) (DD/MM/YYYY)
- iv. Date when specimen collection ended (DD/MM/YYYY)
- v. Date when patient screening ends (DD/MM/YYYY). If the specimen collection takes less than three months, then screening must proceed for AT LEAST three months. If specimen collection takes more than three months, then the specimen collection end date and patient screening end date will be the same date. If the country is unable to collect enough specimens in the six-month survey period, the patient screening end date will be six months after the start date. (See section 5.1)
- vi. Number of ART initiators between date when of survey initiation and patient screening ended (if specimen collection ends earlier than three months)

1 A specimen is considered to be successfully sequenced only when it passes the appropriate quality assurance as recommended by WHO.

2 In general, most surveys will include genotyping of the protease region; only in highly resource-constrained environments where a very small proportion of the population is treated with PI-based regimens should a country opt to not genotype the protease region.

3 Integrase inhibitors (INI) are infrequently used in resource limited settings. Genotyping of the IN region should only be considered if INI containing regimens are used as part of national treatment guidelines. At present, it is anticipated that most countries will not opt to genotype this region.

vii. Estimated number of ART initiators during the six-month period[1]
viii. Clinic size as contained in the table used for systematic sampling (an example of a systematic sampling table is contained in Annex 1)
ix. If stratification is used, specify stratum name (for example old/new clinics) to which each clinic belongs[2]
x. Type of clinic: urban/rural

5.2.3 Survey-level information

i. Total number of clinics sampled
ii. Total number of clinics in the sampling table (an example of a systematic sampling table is contained in Annex 1)
iii. If stratification was used, number of strata and total number of clinics in each sampling table
iv. Sampling interval from systematic sampling table
v. If stratification was used, sampling interval for each stratum

5.3 Patient under-enrolment

If the required sample size per clinic is not achieved during the recommended survey enrolment period of six months, the survey will not achieve the predetermined sample size. If the amount of under-enrolment is minimal, under-enrolment will not greatly affect the precision of the survey. However, if under-enrolment is large, the resulting prevalence estimate of HIVDR will have a wider confidence interval than originally planned. If a country expects to encounter significant under-enrolment (for example, because at random relatively small clinics were sampled), the expected difference should ideally be *equally* distributed to the largest clinics among those sampled which could absorb them. For example, if the overall expected under-enrolment is 40 patients, and among the clinics sampled there are five large clinics that could enrol more patient than the attributed quota, an additional eight patients should be sampled from each large clinic on top of the required quota for that clinic.

5.4 Repeating the survey

This survey is designed to allow for the assessment of trends of HIVDR prevalence in populations of ART initiators over time. Thus, it should be repeated periodically, generally every three years or earlier.

When the survey is repeated, countries are advised to update the sampling table and perform a new random sample of clinics to ensure the new survey is adequately representative of changes in the ART programme.

5.5 Tools for country adaptation

WHO has developed a user-friendly Excel-based calculator to assist countries to determine their locally appropriate sample sizes based on local information (for example, the number of clinics in the country and the number of people who initiated ART in the previous six months). It will be available for download on the WHO HIV drug resistance website at http://www.who.int/hiv/topics/drugresistance/en/index.html.

5.6 Combining pre-treatment and acquired drug resistance surveys

WHO has developed a method to survey acquired HIVDR (ADR) in populations receiving ART[3]. Some countries may wish to concomitantly perform PDR and ADR surveys by sampling the same clinics. If, generally, patients initiate treatment at the same clinics as they receive their follow-up care, then the same sampling table can be constructed and used for both surveys using the proxy measure of *total number of patients on ART* at each clinic. The pre-treatment survey design will be PPPS, and this will result in an increase in the required sample size for the pre-treatment survey. Nonetheless, this increase in sample size may be outweighed by the decrease in logistic complexity and costs associated with sampling fewer clinics overall (between the two surveys). Additional details can be found in the concept note describing Surveillance of Acquired HIV Drug Resistance.

WHO has developed an Excel-based calculator to guide countries in the sampling of ART clinics when the PDR and ADR surveys are conducted simultaneously. It will be made available on the WHO HIV drug resistance website at http://www.who.int/hiv/topics/drugresistance/en/index.html.

1 This estimate is achieved by multiplying two numbers. The first number is equal to the number of ART initiators observed at that clinic between the date of survey initiation and the date when patient screening ends. The second number is equal to 180 divided by the number of days between the date of survey initiation and the date when patient screening ends. Their product is an estimate of the number of ART initiators observed at that clinic during a 6 month period.

2 Stratification is discussed in Section 4.2 (Regional representation) and in Annex 1.3.

3 Available at http://www.who.int/hiv/topics/drugresistance/en/index.html

6. DATA ANALYSIS

For this survey, the Stanford HIVdb algorithm[1] is used to classify HIVDR. The Stanford algorithm classifies HIVDR in five levels: susceptible, potential low-level, low-level, intermediate or high-level drug resistance.

Outcomes 1a, 1b and1c measure the prevalence of ANY HIVDR, defined as low-, intermediate- or high-level resistance according to the Stanford HIVdb to one or more of the following drugs or drug classes: NVP, EFV, any N(t)RTI, DRV/r, LPV/r or ATV/r. Sequences classified as susceptible and potential low-level resistance are considered as having no HIVDR.

Resistance to NNRTI is defined as low-, intermediate- or high-level resistance (according to the Stanford HIVdb) to NVP, EFV or both. Sequences classified as susceptible and potential low-level resistance are considered as having no HIVDR.
Once all data collection has been completed, estimates of (i) overall pre-treatment HIVDR, (ii) HIVDR among treatment initiators without prior ARV exposure, and, if applicable (iii) HIVDR among individuals initiating NNRTI-based ART without prior ARV exposure, along with their respective confidence intervals, will be calculated. Data will be weighted taking into account observed clinic-level patient accrual, number of patients screened and the number of individuals with sequences genotyped. Guidance on data analysis is provided in Annex 1.4. Additional technical background can be found in the Statistical Appendix.

The survey is not powered to generate precise estimates of resistance among initiators with prior ARV exposure or with unknown prior ARV exposure. However, data from initiators with prior ARV exposure can be aggregated across surveys to obtain regional and global estimates with an acceptable confidence interval; these aggregated analyses can critically inform global-level treatment guidelines.

1 Available at: http://sierra2.stanford.edu/sierra/servlet/JSierra

ANNEXES

Annex 1.1: Selecting the clinics to survey

This section describes how to sample clinics from the list of all clinics initiating ART in the country. Sampling of clinics is performed using systematic sampling for generating probability proportional to size (PPS) samples[1] and can be readily adapted for the more general setting of probability proportional to proxy size (PPPS) sampling when recent data on numbers of initiators by clinic is unavailable.

To execute systematic sampling, all clinics initiating patients on ART in the country are listed (Table A1). To enhance the geographical representativeness of the sample, clinics can be listed by administrative or geographical area. Within geographical regions, clinics should be ordered by size.

Operationally, a) List all eligible clinics initiating patients on ART in order of size by region along with the number of ART initiators observed in the previous calendar year (to reflect the relative sizes of the patient populations), b) Calculate the cumulative population size for each clinic listed (described below), c) Determine the sampling interval, d) Pick a random starting-point, f) Select clinics based off the random starting-point, sampling interval and cumulative population size. Table A1 below illustrates these steps in greater detail.

Detailed instructions

1. List geographical or administrative regions within the country in alphabetical order.
2. Within each region, list all clinics initiating patients on ART in that region in order of size.
3. Record the number of eligible patients who initiated ART in the previous calendar year by clinic.
4. Estimate the eligible population size for that clinic during the survey period. For example, if the recorded information is for the previous calendar year, and the survey duration is six months, then create a new column that divides the whole year's numbers by two (= 12 months/2).
5. Starting at the top of the table, calculate the cumulative eligible population size for each clinic in another column. The cumulative eligible population size is the size of the clinic plus the size of all clinics previously listed in the table.
6. Determine the sampling interval by dividing the cumulative population size over all listed clinics by the number of clinics to be sampled. In the case of our example, the cumulative population size is 13,666 and the number of clinics to be sampled is 20. Therefore the sampling interval is 13666/20 = 683.3, rounded to 683.
7. Pick a random starting-point. To select the first clinic, obtain a random number between 1 and the sampling interval 683. A random number generator can be found at http://www.random.org/. For example, the random number obtained in this example was 500.
8. Select clinics based off of the random starting-point, sampling interval and cumulative population size.
   a. Select the first clinic in which the cumulative size is greater than or equal to the random number. Clinic E has a cumulative population size of 500. Because Clinic E is the first clinic such that the cumulative size is greater than or equal to the random start, Clinic E is selected.
   b. Add the initial random number and the sampling interval (500 + 683 = 1183), and then select the first clinic listed in which the cumulative total is greater than or equal to this number (1183). The cumulative size for Clinic F is 856, which is less than 1183. The cumulative size for Clinic G is 1209, making clinic G the first clinic with cumulative size greater than or equal to 1183. Thus, Clinic G is selected. Continue adding the sampling interval to the result obtained until all 20 clinics have been selected.

It is possible for a clinic to be selected more than once if its eligible population size is larger than the sampling interval. In our example, Clinic S is selected twice. If a clinic is picked twice, for example, then twice the sample size must be taken from this clinic. For example, if the sample size is 14 per clinic, then sample size for that clinic is 28. If a clinic is picked k times, then k times the sample size must be taken. The result is that fewer than 20 unique clinics are sampled. In our example, 19 clinics are sampled.

If PPPS sampling is used, clinic size is measured using the number of people on ART at each clinic during a previous time period.

Annex 1.2: Guaranteeing representation from all regions

To guarantee that at least one clinic is sampled from each region, a country must determine prior to sampling the minimum number of clinics that must be sampled. The method is as follows:

- Prepare the table that will be used for systematic sampling (one row per clinic) with clinics sorted by region
  - » 1st column: region
  - » 2nd column: clinic name
  - » 3rd column: clinic size

1 Lohr, Sampling: Design and analysis, 2nd edition, Section 6.2

Table A1: Systematic sampling table for clinic selection (PPS sampling)

A	B	C	D (C/2)	E	F	G
Region	Clinic name	Number of patients initiating ART in previous 12 months	Number of patients estimated to initiate ART during 6-month survey period	Cumulative total of eligible patients	Selection	Sample clinic
A	Clinic A	600	300	300		
A	Clinic B	222	111	411		
A	Clinic C	106	53	464		
A	Clinic D	40	20	484		
A	Clinic E	32	16	500	500 (Random Start)	Clinic 1
B	Clinic F	712	356	856		
B	Clinic G	706	353	1209	500 + 683 = 1183	Clinic 2
B	Clinic H	250	125	1334		
B	Clinic I	90	45	1379		
C	Clinic J	1208	604	1983	1183 + 683 = 1866	Clinic 3
C	Clinic K	1200	600	2583	1866 + 683 = 2549	Clinic 4
C	Clinic L	800	400	2983		
C	Clinic M	766	383	3366	2549 + 683 = 3232	Clinic 5
C	Clinic N	402	201	3567		
C	Clinic O	230	115	3682		
C	Clinic P	210	105	3787		
C	Clinic Q	198	99	3886		
C	Clinic R	50	25	3911		
D	Clinic S	1374	687	4598	3232 + 683 = 3915 3915 + 683 = 4598	Clinic 6 (selected twice)
D	Clinic T	1266	633	5231		
D	Clinic U	1170	585	5816	4598 + 683 = 5281	Clinic 7
E	Clinic V	1302	651	6467	5281 + 683 = 5964	Clinic 8
E	Clinic W	1034	517	6984	5964 + 683 = 6647	Clinic 9
E	Clinic X	706	353	7337	6647 + 683 = 7330	Clinic 10
E	Clinic Y	660	330	7667		
E	Clinic Z	558	279	7946		
E	Clinic AA	334	167	8113	7330 + 683 = 8013	Clinic 11
F	Clinic BB	1260	630	8743	8013 + 683 = 8696	Clinic 12
F	Clinic CC	928	464	9207		
F	Clinic DD	316	158	9365		
F	Clinic EE	66	33	9398	8696 + 683 = 9379	Clinic 13
G	Clinic FF	1376	688	10086		
G	Clinic GG	1196	598	10684	9379 + 683 = 10062	Clinic 14
G	Clinic HH	1112	556	11240	10062 + 683 = 10745	Clinic 15
G	Clinic II	930	465	11705	10745 + 683 = 11428	Clinic 16
G	Clinic JJ	798	399	12104		
G	Clinic KK	570	285	12389	11428 + 683 = 12111	Clinic 17
G	Clinic LL	362	181	12570		

A	B	C	D (C/2)	E	F	G
Region	Clinic name	Number of patients initiating ART in previous 12 months	Number of patients estimated to initiate ART during 6-month survey period	Cumulative total of eligible patients	Selection	Sample clinic
G	Clinic MM	286	143	12713		
H	Clinic NN	1336	668	13381	12111 + 683 = 12794	Clinic 18
H	Clinic OO	570	285	13666	12794 + 683 = 13477	Clinic 19
		Sampling Interval	683			
		Random Start	500	*generated by www.random.org		

- For each region, add up the sizes of all clinics in the region to determine the region size
- Identify the size of the smallest region
- Determine the preferred number of sampled clinics, n (for example 20), based on feasibility
- Calculate the sampling interval, SI = (sum of all clinic sizes over all regions)/*n*
- Test whether the sampling interval is SMALLER than the SMALLEST region size
  - » If the sampling interval is smaller than the smallest region size, the preferred n is appropriate and **at least one clinic per region will be sampled with certainty.**
  - » If the sampling interval is greater than the smallest region size, increase the preferred n to sample one additional clinic (for example *n* was 20, now *n* is 21), and repeat the last two steps. Continue to increase number of clinics until an appropriate number is identified.

For example, using the data presented in **Table A1 Annex 1.1**, if a country wants to guarantee that at least one clinic is sampled from each of the eight regions (A through H), it is necessary to follow the procedure described above to determine the minimum number of clinics to sample.

- For each region, add up the sizes of all clinics in the region to determine the region size.

Region	A	B	C	D	E	F	G	H
Size (6 months)	500	879	2532	1905	2297	1285	3315	953

- Identify the smallest region.
  - » A is the smallest region with 500 patients observed in a six-month period.
- Determine the preferred number of sampled clinics, *n*, based on feasibility.
  - » For example, the country would like to sample *n* = 20 clinics.
- Calculate the sampling interval, SI.
  - » The sampling interval is the sum of all clinic sizes in a six-month period, 13,666, divided by the number of clinics, 20. SI = 13,666/20 = 683.
- Test whether the sampling interval is SMALLER than the SMALLEST region size.
  - » The sampling interval is larger than the size of region A. Thus, it is not guaranteed that at least one clinic from region A will be sampled.
- Since the sampling interval is too large, increase the number of clinics, *n*, until the sampling interval is SMALLER than the SMALLEST region size (that is, less than 500).
  - » For example, for *n* = 27, SI = 13,666/27 = 506.1481.
  - » For example, for *n* = 28, SI = 13,666/28 = 488.0714.
  - » Thus, *n* = 28 is the fewest number of clinics that must be sampled to guarantee at least one clinic from each region is sampled using systematic sampling.
  - » If 28 clinics are too many, the country should consider combining small regions into larger regions.

Annex 1.3: Stratification

Countries may choose to stratify clinics based on certain characteristics, including clinic location (for example, urban/rural) and clinic type (for example, primary/secondary/tertiary, etc.). The benefit of stratification is that countries can predetermine how many clinics are sampled from each stratum. If done appropriately, stratification can increase the precision of the survey. If done inappropriately, stratification can decrease the precision of the survey.

Stratification should only be employed if the country (1) has a strong desire to fix the number of clinics sampled per stratum (for example, to guarantee the same number of clinics per region), or (2) wishes to ensure sufficient sample size to make

precise stratum-specific statements, or (3) wishes to adjust the design so that different sample sizes are requested of different strata (for example, define a small clinic stratum and require only a small number of patients per small clinic), or if (4) the country has knowledge of a clinic-level factor that is associated with pre-treatment resistance (for example, urban clinics tend to have more pre-treatment resistance than rural clinics).

The following guidelines should be observed for the appropriate implementation of stratification:

Defining strata

- The number of stratifying variables should be limited to only those that are most associated with the outcome or those that are most relevant to the survey designers. Extraneous stratifying variables should be avoided.
- The number of levels of the stratifying variables should be limited, where possible. For example, urban/rural is two levels; primary/secondary/tertiary is three levels. If more than one stratifying variable is used, the number of strata is equal to the product of the number of levels in each stratifying variable. For example, if both urban/rural and primary/secondary/tertiary are used, there are six (= 2 x 3) levels. Again, the number of combinations should be limited.
- If a regional variable is used, larger regions (such as north, central and south) are preferable over smaller regions (such as districts, of which there may be many). Small regions (as defined by having few eligible patients) may be combined with other similar regions.
- All strata must be non-empty. If there are no urban tertiary clinics, this stratum should be eliminated. If there are very few patients in urban tertiary clinics, this stratum should be combined with a similar stratum, such as urban secondary clinics.
- In general, no single stratum should be too small, that is, contain too few eligible patients.

Designing the survey

Step 1: calculate the effective sample size

The first step in designing the survey, after determining the number of stratifying variables, is to calculate the effective sample size. The method for calculating the effective sample size given an assumed prevalence, a desired precision, a desired number of clinics to be sampled and the predetermined number of strata, is described in the Statistical Appendix.

» EXAMPLE: prevalence of HIVDR is 10%, confidence interval half-width is 5%, the desired number of clinics sampled is 30, and there are two strata: urban and rural. Then, the total effective sample size is 152 (without adjustment for laboratory failure and design effect).

Step 2: allocate the effective sample size to different strata

Next, the effective sample is allocated to the strata. In general, the effective sample size should be allocated proportionally to each stratum.

» EXAMPLE: If 60% of eligible patients reside in urban areas and 40% in rural areas, $152 \times 0.60 \approx 91$ is the effective sample size of the urban stratrum, and $152 \times 0.40 \approx 61$ is the effective sample size of the rural stratum.

Any design that does not allocate the sample size proportionally to the size of the strata will be less efficient (such as taking equally sized samples from regions of vastly different size) and therefore not recommended.

Step 3: calculate the actual sample size

Now, in each stratum, the effective sample is used to determine an appropriate actual sample size. The effective sample size must be inflated by the design effect, estimated laboratory failure, and other factors described for sample size calculations.

» EXAMPLE: For a PPS design, assuming a desired total number of clinics of 30, the country could choose to sample 18 urban clinics with nine patients per clinic and 12 rural clinics with nine patients per clinic (adjusting for laboratory failure, expected prevalence of prior ARV exposure, and expected proportion of patients initiating NNRTI-based regimens).

Countries do not need to sample the same number per clinic across all strata.

» EXAMPLE: Because urban clinics are much larger than rural clinics, the country could choose a different type of PPS design. The country could choose to sample 10 urban clinics with 17 patients per clinic and 20 rural clinics with six patients per clinic.

COUNTRIES MUST SAMPLE AT LEAST TWO CLINICS FROM EACH STRATUM, even if the stratum is small. This is very important for the analysis stage of the survey.

Executing the survey

To perform the survey, countries create a sampling table for each stratum. Within each stratum, countries can use systematic sampling to sample the desired number of clinics. The general method for systematic sampling is the same as described previously in Annex 1.1.

» EXAMPLE: List all urban clinics. Use systematic sampling to sample 18 urban clinics (if first design described in Step 3 is selected). List all rural clinics. Use systematic sampling to sample 12 rural clinics.

Annex 1.4: Data analysis plan

It is recommended to conduct data entry in Excel, and data analysis in Stata. Instructions are provided below for data analysis in Stata.

Alternative statistical packages can be used to perform data analysis as long as they properly adjust for survey weights, clustering and stratification (if necessary). All statistical packages are expected to yield identical point estimates, but not all statistical packages are expected to yield identical standard error estimates and confidence intervals. Statistical packages that do not allow users to specify the finite population correction at each stage of sampling will overestimate the standard error, especially in countries with small eligible populations. Stata, SUDAAN, and R's survey package allow users to specify finite population corrections at each stage of sampling. SAS and WesVar do not allow users to specify finite population corrections beyond the first stage of sampling. Epi Info does not allow users to specify any finite population corrections.

Deviations in Stata code for survey designs using **Stratification** or sampling **All Sites** are indicated. A survey design without stratification and including only a subset of sites is referred to as a **Standard** design.

An example is provided of a **Standard** survey design. In this example, 19 unique clinics are selected via PPS systematic sampling from the sample population described in Annex 1.1. From each clinic, 14 patients are enrolled into the survey. Because Clinic S was sampled twice during systematic sampling, 28 patients are enrolled from this clinic. Clinic-specific sample sizes also vary because of differential laboratory failure or under-enrolment.

Step 1: Create a table summarizing the necessary information from each clinic

In Excel, create a spreadsheet summarizing the necessary information from each site.

1. List the unique site IDs of the sampled sites in a column labelled "SITE_ID"[1,2].
2. Stratification: List the stratum IDs of each of the sampled sites in a column labelled "STRATUM_ID".
3. Calculate the site sampling weight for each site and record it in a column labelled "SITE_SAMPLING_WT".
   a. **Standard:** The site sampling weight is equal to the sampling interval from the systematic sampling table (for example, 683) divided by the estimated site size from Column D of the systematic sampling table (for example, for site E, the weight is 683/16 = 42.6875, rounded to 42.688). *Note:* smaller sites will have larger site sampling weights.
   b. **Stratification:** For a site from a particular stratum, the site sampling weight is equal to the sampling interval from the stratum-specific systematic sampling table divided by the estimated site size from the same stratum-specific systematic sampling table. Note: each stratum will have a different sampling interval.
   c. **All Sites:** The site sampling weight is equal to 1 for all sites.
4. List site-specific data.
   a. List the estimated 6-month eligible population sizes from each site in a column labelled "N_INITIATORS_6MOS" (this value may be extrapolated using the procedure described in Section 5.2.2 if the eligible population size is observed for less than 6 months).
   b. List the number of initiators screened for prior ARV exposure (for a minimum of three months – see section 5.1) from each site in a column labelled "N_INITIATORS_SCREENED_ PRIOR_ARV". The number of patients screened is equal to the number of patients with recorded ARV exposure information (yes/no/unknown).
   c. List the number of patients with genotyped sequences from each site in a column labelled "N_INITIATORS_GENOTYPED".

Table A2: Example of site-level data

SITE_ID	SITE_SAMPLING_WT	N_INITIATORS_6MOS	N_INITIATORS_SCREENED_PRIOR ARV	N_INITIATORS_GENOTYPED
E	42.688	18	14	14
G	1.935	402	297	11
J	1.131	580	367	12
K	1.138	633	392	11
M	1.783	420	277	12
S	0.994	788	662	22
…				

Step 2: Create a table summarizing the necessary information for each patient sampled

In Excel, create a spreadsheet summarizing the necessary information for each patient screened. (See Table A3 for an example.)

1 At the analysis stage, all variable names should be indicated with capital letters without quotation marks in the column headers.

2 It is recommended that the site ID correspond exactly to the three-letter site code described in Box 4. In this highly simplified example, a one- to two-letter site code is used.

1. List the unique patient ID in a column labelled "ID".
2. List the site ID in a column labelled "SITE_ID" (must be identical to the site ID in the column labelled "SITE_ID" in Table A2).[1]
3. List a number indicating the patient's exposure to previous ARVs (0 if no prior exposure, 1 if prior exposure, 2 if unknown) in a column labelled "ARV_PREVIOUS_EXPOSURE_RF".
4. List a binary variable indicating whether a patient is initiating an NNRTI-based regimen in a column labelled "NNRTI_REGIMEN_BN" (1 if NNRTI-based; 0 if not NNRTI-based; missing if unknown).
5. List a binary variable indicating whether a patient had detected HIVDR in a column labelled "ANY_HIVDR_BN" (1 if HIVDR[2]; 0 if no HIVDR; missing if no specimen was collected or specimen was collected but no data are available).
6. Save data in a spreadsheet, such as "PDR_PT_DATA.xlsx".

Table A3: Example of patient-level data

ID	SITE_ID	ARV_PREVIOUS_EXPOSURE_RF	NNRTI_REGIMEN_BN	ANY_HIVDR_BN
PDR-2014-E-0001*	E	0	1	0
PDR-2014-E-0002*	E	0	0	1
PDR-2014-E-0003*	E	0	1	
PDR-2014-E-0004*	E	1	1	1
….				
PDR-2014-G-0001*	G	1	0	0
PDR-2014-G-0002*	G	2		1
PDR-2014-G-0003*	G	2	1	

*simplified patient IDs; but true patient IDs should follow naming convention in Box 4

In the above sample table:

- Patient PDR-2014-E-0001 (from Site E) did not report prior exposure to ARV and did not have evidence of HIVDR. This patient is initiating a NNRTI-based regimen.
- Patient PDR-2014-E-0002 (from Site E) did not report prior exposure to ARV and did have evidence of HIVDR. This patient is initiating a non-NNRTI-based regimen.
- Patient PDR-2014-E-0003 (from Site E) did not report prior exposure to ARV and did not have a genotype available because the specimen was of poor quality. This patient is initiating a NNRTI-based regimen.
- Patient PDR-2014-E-0004 (from Site E) did report prior exposure to ARV and did have evidence of HIVDR. This patient is initiating a NNRTI-based regimen.
- Patient PDR-2014-G-0001 (from Site G) did report prior exposure to ARV and did not have evidence of HIVDR. This patient is initiating a non-NNRTI-based regimen.
- Patient PDR-2014-G-0002 (from Site G) had unknown prior exposure to ARV and did have evidence of HIVDR. This patient is initiating an unknown regimen.
- Patient PDR-2014-G-0003 (from Site G) had unknown prior exposure to ARV and did not have a genotype available because a specimen was not collected. This patient is initiating a NNRTI-based regimen.

Step 3: Import data into Stata

1. Import site data using the import data option (*File/Import/Excel Spreadsheet*)[3]. Use the Browse button to identify the spreadsheet. Select the option to import the first row as variable names. Change the variable case to upper to preserve variable names.
2. Save site data as a .dta file using the save option (*File/Save*). In this example, we save the data as "PDR_SITE_DATA.dta".
3. Import patient data using the import data option (*File/Import/Excel Spreadsheet*). Use the Browse button to identify the spreadsheet. Select the option to import the first row as variable names. Change the variable case to upper to preserve variable names.
4. Save patient data as a .dta file using the save option (*File/Save*). In this example, we save the data as "PDR_PT_DATA.dta". Press "Yes" to overwrite data currently in memory.
5. Merge the two datasets using the merge option (*Data/Combine datasets/Merge two datasets*). Select the "Many-to-one" option. Select or type in "SITE_ID" as the Key Variable. Use the Browse button to select "PDR_SITE_DATA.dta". Press OK.

Step 4: Create survey weights and other necessary variables

1. Create survey weight for Outcomes 1a, 1b, and 1c. In the command line, type in "generate OUTCOME1_WT = SITE_SAMPLING_WT*(N_INITIATORS_6MOS/N_SPECIMENS_GENOTYPED)"[4].
2. Create survey weight for Outcomes 2a, 2b and 2c. In the command line, type in "generate OUTCOME2_WT = SITE_SAMPLING_WT*(N_INITIATORS_6MOS/N_INITIATORS_SCREENED)".
3. Create a variable indicating the total number of sites in the sampling frame (prior to systematic sampling).

1 It is recommended that the site ID correspond exactly to the three-letter site code described in Box 4. In this highly simplified example, a one- to two-letter site code is used.

2 Refer to Section 6

3 (*Menu/Option/Sub-option*) indicates using the drop-down menus to select an option.

4 For all Stata commands in the command line, do not include the quotation marks. Type only the text between the quotations marks.

a. **Standard or All Sites:** In our example, there were 41 total sites. In the command line, type in "generate N_TOTAL_SITES = 41". In practice, replace 41 with your country-specific number.
b. **Stratification:** The variable should refer to the number of total sites in each stratum-specific sampling frame. In the command line, type in "generate N_TOTAL_SITES = .". Then, for each stratum, use the replace command to identify the stratum-specific number. For example, if there are 50 sites in stratum 1 and 40 sites in stratum 2, type in "replace N_TOTAL_SITES = 50 if STRATUM_ID == 1" and "replace N_TOTAL_SITES = 40 if STRATUM_ID == 2".

4. Create a variable to be used for reporting results for global aggregation. In the command line, type in "generate POP_SIZE = 1".

Step 5: Declare survey design and analyse data

1. Declare survey design for Outcomes 1a, 1b, and 1c (*Statistics/Survey data analysis/Setup & utilities/Declare survey design for dataset*)..
   a. In the Main tab, change "Number of stages" to 2.
   b. Select "SITE_ID" as the "Stage 1: Sampling units".
   c. **Standard or All Sites:** Leave "Stage 1: Strata" blank.
   d. **Stratification:** Select "STRATUM_ID" as the "Stage 1: Strata".
   e. Select "N_TOTAL_SITES" as the "Stage 1: Finite pop. correction".
   f. Select "ID" as the "Stage 2: Sampling units".
   g. Leave "Stage 2: Strata" blank.
   h. Select "N_INITIATORS_6MOS" as the "Stage 2: Finite pop. correction".
   i. In the Weights tab, select "OUTCOME1_WT" as the "Sampling weight variable".
   j. In the SE tab, select "Center at the grand mean" for Strata with a single sampling unit. Press OK.

2. Analyse Outcome 1a. (*Statistics/Survey data analysis/Means, proportions, ratios, totals/Proportions*).
   a. In the Model tab, select "ANY_HIVDR_BN" as the "Variable". Press OK.
   b. The point estimate, standard error and 95% confidence interval for the prevalence of HIVDR are located in the row labelled "1".
   c. In the output, the number of PSUs (Primary Sampling Units) is equal to the number of unique clinics sampled in the survey. The design degrees of freedom are equal to the number of PSUs minus the number of strata, where the number of strata is equal to 1 if no stratification is used.

. svy linearized : proportion ANY_HIVDR_BN

Survey: Proportion estimation			
Number of strata =	1	Number of obs =	242
Number of PSUs =	19	Population size =	13408.1
		Design df =	18

	Proportion	Linearized Std. Err.	[95% Conf.	Interval]
ANY_HIVDR_BN				
0	.9537658	.0119793	.9209964	.9733365
1	.0462342	.0119793	.0266635	.0790036

3. Analyse Outcome 1b. (*Statistics/Survey data analysis/Means, proportions, ratios, totals/Proportions*).
   a. In the Model tab, select "ANY_HIVDR_BN" as the "Variable(s)".
   b. In the if/in/over tab, type in "ARV_PREVIOUS_EXPOSURE_RF==0" (selects only patients without prior exposure) into the "If: (expression)" box. Press OK.
   c. The point estimate, standard error and 95% confidence interval for the prevalence of HIVDR among patients without prior exposure are located in the row labelled "1".
   d. The number of observations used to calculate this outcome is labelled "Subpop. no. obs". Disregard the value labelled "Number of obs" as it is not a meaningful quantity.

. svy linearized, subpop(if ARV_PREVIOUS_EXPOSURE_RF == 0) : proportion ANY_HIVDR_BN

Survey: Proportion estimation			
Number of strata =	1	Number of obs =	1849
Number of PSUs =	19	Population size =	103577
		Subpop. no. obs =	186
		Subpop. size =	10427.1
		Design df =	18

	Proportion	Linearized Std. Err.	[95% Conf.	Interval]
ANY_HIVDR_BN				
0	.9672758	.0117146	.9314282	.9846912
1	.0327242	.0117146	.0153088	.0685718

4. Analyse Outcome 1c in the same manner as Outcome 1b, selecting patients without prior exposure to ARVs initiating first-line NNRTI-based regimens as the appropriate subpopulation ("ARV_PREVIOUS_EXPOSURE_RF==0 & NNRTI_REGIMEN_BN==1").
5. Declare survey design for Outcomes 2a, 2b and 2c (*Statistics/Survey data analysis/Setup & utilities/Declare survey design for dataset*).
   a. In the Main and SE tabs, select the same options described in Step 1.
   b. In the Weights tab, select "OUTCOME2_WT" as the "Sampling weight variable". Press OK.
6. Analyse Outcome 2a, 2b and 2c (*Statistics/Survey data analysis/Tables/One-way tables*).
   a. In the Model tab, select "PREVIOUS_ARV_EXPOSURE_RF" as the "Categorical variable".
   b. In the Table items table, check the boxes for "Standard errors" and "Confidence intervals". Press OK.
   c. The point estimate, standard error and 95% confidence interval for the prevalence of no previous exposure to ARVs are located in the row labelled "0".
   d. The point estimate, standard error and 95% confidence interval for the prevalence of previous exposure to ARVs are located in the row labelled "1".
   e. The point estimate, standard error and 95% confidence interval for the prevalence of unknown previous exposure to ARVs are located in the row labelled "2".

. svy linearized : tabulate ARV_PREVIOUS_EXPOSURE_RF, se ci

Survey: Proportion estimation			
Number of strata =	1	Number of obs =	6387
Number of PSUs =	19	Population size =	13408.127
		Design df =	18

ARV_PREVIOUS_EXPOSURE_RF	proportions	se	lb	ub
0	.7443	.0052	.7333	.7551
1	.2003	.0055	.1889	.2122
2	.0554	.0036	.0482	.0634

Key:	
	proportions = cell proportions
	se = linearized standard errors of cell proportions
	lb = lower 95% confidence bounds for cell proportions
	ub = upper 95% confidence bounds for cell proportions

7. Analyse data for aggregate reporting to WHO. Example shown is the prevalence of HIVDR among initiators with prior ARV exposure.
   a. Repeat Step 1 to declare survey design used for Outcomes 1a and 1b.
   b. Repeat Step 3, but in the if/in/over tab, type in "ARV_PREVIOUS_EXPOSURE_RF==1" (selects only patients with prior exposure) into the "If: (expression)" box. Press OK.
   c. The point estimate and standard error for the prevalence of HIVDR among patients with prior exposure are located in the row labelled "1". The number of observations used to construct this estimate is labelled "Subpop. no. obs"; this is equal to the number of patients with prior ARV exposure and successfully genotyped). The design degrees of freedom are labelled "Design df".

. svy linearized, subpop(if ARV_PREVIOUS_EXPOSURE_RF == 1) : proportion ANY_HIVDR_BN

Survey: Proportion estimation			
Number of strata =	1	Number of obs =	5104
Number of PSUs =	19	Population size =	286646
		Subpop. no. obs =	40
		Subpop. size =	2127.8
		Design df =	18

	Proportion	Linearized Std. Err.	[95% Conf.	Interval]
ANY_HIVDR_BN				
0	.8960278	.038704	.7826241	.9537646
1	.1039722	.038704	.0462354	.2173759

   d. To aggregate the data at a global level, it is also necessary to report the numerator of the prevalence estimate (and its associated standard error) and the denominator of the prevalence estimate (and its associated standard error). In this scenario, the numerator is an estimate of the total number of ART initiators in the country during the 6-month survey period with prior exposure to ARVs and detected HIVDR. The denominator is an estimate of the total number of ART initiators in the country during the 6-month survey period with prior exposure to ARVs. The prevalence is equal to the numerator divided by the denominator. Select (*Statistics/Survey data analysis/Means, proportions, ratios, totals/Totals*).

In the Variables box in the Model tab, type or select "ANY_HIVDR_BN POP_SIZE". In the if/in/over tab, type in "ARV_PREVIOUS_EXPOSURE_RF==1" (selects only patients with prior exposure) into the "If: (expression)" box. Press OK.

e. The numerator estimate and its standard error are located in the row labelled "ANY_HIVDR_BN". The denominator estimate and its standard error are located in the row labelled "POP_SIZE".

. svy linearized, subpop(if ARV_PREVIOUS_EXPOSURE_RF == 1) : total ANY_HIVDR_BN POP_SIZE

Survey: Total estimation			
Number of strata =	1	Number of obs =	5104
Number of PSUs =	19	Population size =	286646
		Subpop. no. obs =	40
		Subpop. size =	2127.8
		Design df =	18

	Proportion	Linearized Std. Err.	[95% Conf.	Interval]
ANY_HIVDR_BN	221.232	90.36867	31.37443	411.0895
POP_SIZE	2127.799	286.2371	1526.437	2729.161

Annex 1.5: Reporting of HIVDR data

All countries are expected to report to WHO a dataset including (1) individual patient information (demographic and matching laboratory data), (2) clinic data and (3) survey variables discussed in Section 5.2, in addition to the patient sequences in FASTA file format. It is recommended that sequence IDs, patient IDs and specimen ID numbers be identical.

In countries where individual patient information and sequences cannot be reported, survey outcomes and additional data on the prevalence of HIVDR in different subpopulations should be reported in an aggregated fashion. **An Excel data collection and reporting tool will be available on the WHO website.** Prevalence data should be accompanied by numerator, denominator, standard error of prevalence, standard error of numerator, standard error of denominator and design degrees of freedom, to allow pooling of regional and global data.

For this survey, the Stanford HIVdb algorithm[1] is used to classify HIVDR. The Stanford HIVdb algorithm classifies HIVDR in five levels: susceptible, potential low-level, low-level, intermediate or high-level drug resistance. Sequences classified as susceptible and potential low-level resistance are considered as having "no HIVDR".

The utilization of these different categories is summarized below:

1. HIVDR by individual drug

When reporting HIVDR by individual drug, sequences classified as low-, intermediate- or high-level resistance according to the Stanford HIVdb should be classified as "HIV drug resistance". This classification applies to all drugs.

2. HIVDR by drug class

When reporting HIVDR by drug class, the following operational definitions for drug class should be used:

- NNRTI class refers to any NNRTI
- NRTI class refers to any NRTI
- boosted PI class refers only to DRV/r, or LPV/r or ATV/r
- Integrase inhibitor class refers to any integrase inhibitor

Sequences classified as low-, intermediate- or high-level resistance according to the Stanford HIVdb should be aggregated as "HIV drug resistance".

3. Any HIVDR

"Any HIV drug resistance" is defined in sequences classified as low-, intermediate- or high-level resistance according to the Stanford HIVdb with respect to one or more of the following drugs: NVP, EFV, any N(t)RTI, DRV/r, LPV/r or ATV/r[2].

In countries where these individual variables and sequences cannot be reported, aggregate data on HIVDR among individuals without, with and unknown prior ARV exposure should be recorded and reported by drug. A standardized Excel-based reporting form will be available for download on the WHO HIV drug resistance website at http://www.who.int/hiv/topics/drugresistance/en/index.html.

1 Available at: http://sierra2.stanford.edu/sierra/servlet/JSierra

2 INI should not be included.

STATISTICAL APPENDIX

1. Sample size calculations

To determine the necessary sample size for the survey, we start by determining the *effective* sample size. The effective sample size, k_{eff}, refers to the number of patients we would need to sample to achieve a desired confidence interval half-width if we were conducting a simple random sample with replacement. The effective sample size is determined by the prevalence of the outcome and the desired width of the confidence interval. The effective sample size is then multiplied by the estimated design effect to yield the actual sample size of the survey.

To determine the effective sample size, consider the following formula for a Wald-type confidence interval. Here, $\tilde{p}_{DR}$ refers to the assumed prevalence of drug resistance among all ART initiators, n refers to the number of clinics sampled, and df are the design degrees of freedom:

$$95\%\ CI = \left(\tilde{p}_{DR} - t_{df,0.975}\sqrt{\frac{\tilde{p}_{DR}(1-\tilde{p}_{DR})}{k_{eff}}}, \tilde{p}_{DR} + t_{df,0.975}\sqrt{\frac{\tilde{p}_{DR}(1-\tilde{p}_{DR})}{k_{eff}}} \right)$$

The design degrees of freedom are defined as df = (number of clinics sampled) – (number of strata). If stratification is not used, $df = n - 1$.

The half-width of this confidence interval (referring to the distance from the midpoint to either end of the confidence interval) is:

$$L = t_{df,0.975}\sqrt{\frac{\tilde{p}_{DR}(1-\tilde{p}_{DR})}{k_{eff}}}$$

So, the effective sample size k_{eff} is the smallest sample size such that $t_{df,0.975}\sqrt{\frac{\tilde{p}_{DR}(1-\tilde{p}_{DR})}{k_{eff}}}$ is less than L. Once the effective sample size is calculated, it has to be adjusted to reflect the design effect. This is addressed in Section 2 below.

The effective sample size can be calculated using the following formula[1]:

$$k_{eff} = \frac{(t_{df,0.975})^2 \tilde{p}_{DR}(1-\tilde{p}_{DR})}{L^2}$$

k_{eff} should be rounded up to the nearest integer. An example of how the effective sample size is calculated is provided in Box A1.

Note: Because the method for calculating a confidence interval in the setting of clustered surveys uses a t distribution with degrees of freedom equal to the design degrees of freedom, our effective sample size is also a function of the number of clinics sampled. When the design degrees of freedom are large (around 40 or greater), it is standard to assume that $z_{0.975} \approx t_{df,\,0.975}$ as this simplifies calculations. This is only appropriate when the design degrees of freedom are large. Since this design requires sampling of around 15–40 clinics, the design degrees of freedom will be small, and it is thus inadvisable to make this simplification. The consequence of using this simplification would be an underestimation of the total sample size required to achieve a given confidence interval half-width.

1 Lohr (2010) Sampling: Design and analysis, 2nd edition, Boston, MA, Cengage Learning. Section 7.5.2

Box A1: Calculating the effective sample size

A country plans to sample $n = 20$ clinics without stratification ($df = 20–1 = 19$, and $t_{19,0.975} = 2.093$), assumes that the prevalence of drug resistance among all ART initiators is 10%, ($\tilde{p} = 0.10$), and desires a confidence interval width of ±5% ($L = 0.05$), then the following effective sample size is required:

$$k_{eff} = \frac{2.093^2 * 0.10 * (1 - 0.10)}{0.05^2} = 157.70$$

Thus, the required effective sample size in this example is approximately 158 individuals. (Note: always round up when performing sample size calculations.) Thus, if a simple random sample of treatment initiators in the country was being performed, a sample size of 158 individuals would be required to achieve the desired precision.

The effective sample size must be inflated to determine the actual sample size required for the survey. The amount by which the sample size is increased is called the design effect. The elements of the study design that contribute to the design effect are (1) clustering of the outcome by clinic ($DEFF_{clust}$), and (2) imperfect information from using data from a previous year or from a slightly different population ($DEFF_{info}$). These elements are described in greater detail below.

2. Calculating the contribution to the design effect due to clustering of the outcome by clinic

It is first necessary to calculate the design effect due to clustering of the outcome. Clustering of the outcome occurs because the amount of drug resistance among initiators varies by clinic. Initiators from the same clinic may have more similar drug resistance outcomes than initiators from different clinics in the same country. The similarity of initiators within clinics is measured via the intracluster correlation coefficient, or ICC.

If m is the number of initiators without prior exposure to ARVs sampled per clinic and ICC_{DR} is the estimated intracluster correlation for the drug resistance outcome, the design effect due to clustering can be estimated using the following formula:

$$DEFF_{clust} = 1 + (m - 1)ICC_{DR}$$

The design effect due to clustering increases as more initiators are sampled from the same clinics (m increases).
Box A2 summarizes the process for estimating intracluster correlations.

Box A2: Estimating intracluster correlations

In order to estimate the intracluster correlation, global data from WHO's HIV Drug Resistance Report 2012 were used. For each clinic in each country, the estimated probability of drug resistance for treatment initiators was used to calculate the intracluster correlation using an analysis of variance estimator[2]. Although intracluster correlation is defined as capturing the clustering of outcomes by clinics within the same country, clinics in the dataset were not separated by country.

For the outcome of pre-treatment HIVDR, the estimated ICC using data from the 2012 WHO drug resistance report is $ICC_{DR,raw} = 0.005.$. The observed prevalence of pre-treatment HIVDR in the global data is 4.5%. As the assumed prevalence of HIVDR among initiators is 10%, and since the ICC and prevalence are generally correlated, the ICC was adjusted to reflect the assumed prevalence. To perform this adjustment, a linear model predicting log(ICC) by log(prevalence) was applied[3]. The equation is $ICC_{DR} = ICC_{DR,raw} \times \exp(0.91 \times 1n[\tilde{p}_{DR}/0.045])$. For the assumed prevalence of 10%, the multiplicative factor was 2.06, resulting in an estimated ICC of $ICC_{DR} = 2.06 \times ICC_{DR,raw} = 0.010.$

It is important to note that there are limitations to this estimate. In particular, the ICC is only based on the data available in the WHO's global HIV drug resistance report 2012. As this survey is repeated over time, results of the surveys can be used to better inform the estimate of ICC and future iterations of the survey.

3. Calculating the contribution to the design effect due to imperfect weighting information

Ideally, we sample clinics proportional to the number of individuals who initiate ART. However, these numbers are generally not available in most countries. A reasonable alternative is the number, by clinic, of individuals who initiate ART during a previous time period. If sampling is conducted using this information, we refer to this as *probability proportional to size (PPS) sampling.*

If the number of initiators by clinic is not available, countries may employ probability *proportional to proxy size (PPPS) sampling. In PPPS sampling*, clinics are sampled with probabilities proportional to a proxy measure, such as the total number of individuals on ARVs by clinic. Ideally, the number of treatment initiators at each clinic will be highly correlated with the proxy measure; thus, the overall design will be close to proportional. If the number of treatment initiators is not highly correlated with the proxy measure, then the design will be further from proportional and, thus, less efficient.

Because information from a previous time period or from a slightly different population is used to conduct sampling, the weights will not be perfectly proportional. It is expected that the sizes of the eligible populations within the clinics will change over time, although it is assumed that the changes in the relative clinic sizes will not be dramatic. If clinic populations change dramatically over time, for example because of decentralization, this information should be incorporated into the estimated population sizes used for sampling. The goal is to use estimated population sizes that will be most predictive of the population sizes to be observed during the survey period. The survey period is defined as the 6 month period starting on the date of survey initiation; it is the maximum duration of the survey and the period for which the survey results will be generalized.

To estimate the effect of imperfect information on the design effect, we can use a formula estimating the variance contribution for disproportionate weights[3]. The design effect can be approximated by $DEFF_{info} = 1+cv^2\,(weights)$, where $cv()$ refers to the

1 Ridout et al. (1999) Estimating intraclass correlation for binary data. Biometrics 55, 137-148.

2 Guillford et al. (2005) Intraclass correlation coefficient and outcome prevalence are associated in clustered binary data. Journal of Clinical Epidemiology 58, 246-251. Note: log prevalence coefficient used is 0.91 (from HTA data)

3 Kalton, et al. (2005), Section B, Chapter VI: Estimating components of design effects for use in sample design. In: Household sample surveys in developing and transition countries. New York, United Nations. Retrieved at http://unstats.un.org/unsd/hhsurveys/. See p. 110, eqn. 23.

coefficient of variation. For PPS sampling, it is estimated that $DEFF_{info} = 1.10$ is reasonable for this setting; this corresponds to inflating the sample size by 10%. For PPPS sampling, it is estimated that $DEFF_{info} = 1.50$ is reasonable for this setting; this corresponds to inflating the sample size by 50%. PPPS sampling requires greater inflation than PPS sampling because the weights are expected to be less proportional, yielding a less efficient design. These numbers were calculated from observing the differences in population sizes between treatment initiators and patients on ART at clinics in an African country over a two-year period.

4. Other sources of design effect

The design effect is also influenced by other sources of variability. For example, different clinics will have different levels of genotyping failure. This will induce additional variability in the weights. It is estimated that this source of design effect will be small, so it is ignored in the calculations to increase the simplicity of the design.

5. Calculating the sample size

The total design effect is estimated using the following equation[1]:

$$DEFF = DEFF_{clust} \times DEFF_{info}$$

The following procedure can be used to identify an appropriate sample size for the survey:

- Calculate necessary effective sample size k_{eff} for a given number of clinics n
- Determine the appropriate value of $DEFF_{info}$ based on the sampling design
- Determine the appropriate value of the ICC, for example $ICC_{DR} = 0.010$
- Solve the following equation for m :

$$m = \frac{1 - \mathrm{ICC_{DR}}}{\left[\frac{n}{\mathrm{DEFF_{info}}\, k_{eff}} - \mathrm{ICC_{DR}}\right]}$$

- If such an m does not exist, or if the calculated value of m is too large to be practical in a particular setting, consider increasing the number of clinics sampled, n. Because of the design effect, sampling a larger number of clinics may require fewer samples per clinic, and it will also require a smaller overall sample size.

In order to retain statistical power at the analysis stage when considering only patients without prior ARV exposure initiating NNRTI-based regimens, the sample size needs to be adjusted for three additional parameters: (i) laboratory failure when genotyping; for example, if we expect a 10% genotyping failure rate, we need to divide the required sample size by 0.90, (ii) expected proportion of initiators without prior exposure to ARVs; for example, if we assume that 75% of initiators will not have had prior exposure to ARVs, we need to divide the required sample by size by 0.75, and (iii) the proportion of patients initiating NNRTI-based regimens; for example, for a country in which all patients (100%) initiate NNRTI-based regimens, we must divide by 1.00, which will have no impact on the calculations.

$$m_{samp} = m/(0.90 \times 0.75 \times 1.00)$$

Because the prevalence of HIVDR among initiators without prior exposure is expected to be less than the prevalence of HIVDR among all initiators, the sample size calculations to achieve a particular precision for Outcome 1c are likely conservative. Because the sample sizes for Outcome 1c are then inflated to generate the overall sample sizes, it is likely that the sample size calculations to achieve a particular precision for Outcome 1a are conservative as well.

1 Park and Lee (2004) Design effects for the weighted mean and total estimators under complex surveys sampling, Survey Methodology 30: 183-193.

6. Incorporating the finite population correction

The formula for the design effect due to clustering can be revised to incorporate the predicted effect of the finite population corrections which will be applied at the analysis stage. The design effect due to clustering in the absence of finite population corrections is $DEFF_{clust} = 1 + (m - 1)ICC_{DR}$ where m is the number of patients with available genotypes per clinic, and ICC is the intracluster correlation. For a country with N total clinics in the sampling frame and an average of $\bar{M}$ eligible patients per clinic, it can be shown that the design effect due to clustering can be approximated by:

$$DEFF_{clust} = \left(1 - \frac{m}{\bar{M}}\right) + \left[\left(1 - \frac{n}{N}\right)m - \left(1 - \frac{m}{\bar{M}}\right)\right] ICC_{DR}$$

The average number of eligible patients per clinic can be estimated as the total number of eligible patients (estimated from available data) divided by the total number of clinics in the sampling frame ($\bar{M} = M/N$). It can be shown that the necessary number of patients per clinic to be sampled per clinic to achieve a desired precision is:

$$m = \frac{1 - ICC_{DR}}{\left[\frac{n}{k_{eff} DEFF_{info}} - ICC_{DR}\left(1 - \frac{n}{N}\right) + \frac{N}{M}(1 - ICC_{DR})\right]}$$

The sample size must then be adjusted for expected genotyping failure and the expected proportion of initiators without prior exposure to ARVs. This calculation assumes that at least m patients can be sampled from all clinics; if fewer than m patients are sampled from a clinic, the confidence interval will be slightly wider than planned for.

7. Sample size calculations when all clinics are sampled

If all clinics in the sampling frame will be included in the survey, the following modifications can be made to the sample size calculations. Briefly, the survey effective sample size is calculated, and this effective sample size is multiplied by a design effect due to imperfect information, the expected laboratory failure, and the expected proportion of patients without prior ARV exposure. It is not necessary to multiply the calculations by a design effect due to clustering because all clinics in the sampling frame are included in the survey. The effective sample size necessary to achieve a confidence interval of half-width L is:

$$k_{eff} = \frac{3.84\, \tilde{p}_{VLS}(1 - \tilde{p}_{VLS})}{L^2}$$

If the finite population correction is incorporated into the calculations (where M is the total eligible population size in the country), then the effective sample size can be calculated using the following equation:

$$k_{eff} = \frac{M \times 3.84\, \tilde{p}_{VLS}(1 - \tilde{p}_{VLS})}{L^2 \times M + 3.84\, \tilde{p}_{VLS}(1 - \tilde{p}_{VLS})}$$

Because information on patient enrolment from a prior time period will be used to allocate the sample, it is recommended that the sample size be inflated slightly to account for imperfect information. If the relative clinic sizes are estimated using the number of initiators per clinic from a previous time period, the sample should be inflated by $DEFF_{info} = 1.10$ or 10% (equivalent to adjusting for the design effect for disproportionate weighting for PPS sampling). If only information on the total number of patients on ART per clinic is available, the sample should be inflated by $DEFF_{info} = 1.50$ (equivalent to adjusting for the design effect for disproportionate weighting for PPPS sampling).

Next, the sample size should be inflated by the amount of expected laboratory success rate (90%), the expected proportion of initiators without prior ARV exposure (75%), and the expected proportion of patients initiating NNRTI-based regimens (if different from 100%). Thus, the actual sample size for the PPS-equivalent design is:

$$k_{act} = \frac{k_{eff} \times 1.10}{0.90 \times 0.75 \times 1.00} \approx 1.63 k_{eff}$$

The actual sample size for the PPPS-equivalent design is: $k_{act} = \frac{k_{eff} \times 1.50}{0.90 \times 0.75 \times 1.00} \approx 2.22 k_{eff}$

The actual sample size is then allocated to the clinics proportional to the number of eligible patients expected to be observed during the survey period. For each clinic, the sample size of that clinic is equal to the total sample size, k_{act}, multiplied by the expected patient accrual at that clinic divided by the expected patient accrual for all clinics included in the survey. For example, if 25% of patients in a country attend a particular clinic, 25% of the sample size should be allocated to that clinic. The per-clinic sample sizes are rounded to the nearest whole number.

8. Data analysis

8.1 Data analysis: Clinic sampling weight and clinic size

Once an appropriate design is identified, clinics will be sampled using either PPS or PPPS systematic sampling. In PPS, clinic size is estimated using prior data on the number of initiators by clinic. In PPPS, clinic size is estimated using prior data on the number of patients on ART by clinic. For clinic i, the estimated clinic size in the sampling frame (from either PPS or PPPS) is denoted as $\widetilde{M}_i$. If the predetermined number of clinics to be selected is n^*, the probability that a clinic is selected is equal to $n^* \widetilde{M}_i$ divided by the total size of all clinics in the sampling frame, $\widetilde{M} = \sum_{j=1}^{N} \widetilde{M}_j$.

Note: llarge clinics may be sampled more than once using this methodology. If a large clinic is sampled twice, this clinic should sample twice as many patients, and so on. In this case, the number of unique clinics selected, denoted n, is fewer than n^*. In this setting, it is necessary to distinguish between n^* and n in the calculations.

Thus, the clinic sampling weight is equal to the following (where SI is the sampling interval defined in Annex 1.1):

$$w_{clinic,i} = \left(\frac{\widetilde{M}}{n^* \widetilde{M}_i}\right) = \left(\frac{SI}{\widetilde{M}_i}\right)$$

If all clinics are included in the survey, the clinic sampling weight is equal to 1 for all clinics. If a stratified survey is conducted, clinic weights should be constructed separately for each sampling frame.

M_i is a count of the number of eligible patients attending clinic i *observed during the six-month survey period.* In the case of those clinics that reach their enrolment quotas before six months, they should continue to count eligible patients for a minimum of three months, and M_i can be estimated as two times the number of eligible patients attending clinic i observed during the three-month period.

8.2 Data analysis: Outcome 1a

Outcome 1a is the overall prevalence of HIVDR among all ART initiators, regardless of prior ARV exposure. Data analysis for this and all additional outcomes is to be conducted in Stata using the SVY utilities[1]. EEven if Stata is not used to conduct the analysis, the Stata SVY manual section on Variance Estimation contains all necessary formulae for calculating the prevalence, variance and 95% confidence interval of each outcome.

The clinic sampling weight is defined in Section 8.1. The patient sampling weight for clinic i is defined as M_i divided by the number of initiators with genotyped data available from that clinic. The overall weight is the product of the clinic and patient sampling weights. All patients with available genotypes are defined as either having HIVDR (binary variable for any HIV drug resistance mutations = 1) or no detected HIVDR (binary variable for any HIV drug resistance mutations = 0). The prevalence is estimated using a ratio. The numerator is an estimate of the total number of initiators in the country with HIV drug resistance mutations during the survey period. The denominator is an estimate of the total number of initiators in the country during the survey period. The variance is calculated using linearization. A 95% confidence interval can be calculated using a standard Wald formula or by a logit transformation (default in Stata).

1 StataCorp. 2013. Stata: Release 13. Statistical software. College Station, TX, StataCorp LP

8.3 Data analysis: Outcome 1b

Outcome 1b is the prevalence of HIVDR among ART initiators without prior exposure to ARVs. Data analysis is conducted using the same sampling weights described for Outcome 1a. The population is restricted to patients without prior exposure to ARVs using the subpopulation command in Stata. All patients without prior ARV exposure with available genotypes are defined as either having HIVDR (binary variable for any HIV drug resistance mutations = 1) or no detected HIVDR (binary variable for any HIV drug resistance mutations = 0). The prevalence is estimated using a ratio. The numerator is an estimate of the total number of initiators in the country without prior exposure to ARVs with HIV drug resistance mutations during the survey period. The denominator is an estimate of the total number of initiators in the country without prior exposure to ARVs during the survey period. The variance is calculated using linearization. A 95% confidence interval can be calculated using a standard Wald formula or by a logit transformation (default in Stata).

8.4 Data analysis: Outcome 1c

Outcome 1c is the prevalence of HIVDR among ART initiators without prior exposure to ARVs initiating NNRTI-based regimens. Data analysis is conducted using the same sampling weights described for Outcome 1a. The population is restricted to patients without prior exposure to ARVs initiating NNRTI-based regimens using the subpopulation command in Stata. All patients without prior ARV exposure initiating NNRTI-based regimens with available genotypes are defined as either having HIVDR (binary variable for any HIV drug resistance mutations = 1) or no detected HIVDR (binary variable for any HIV drug resistance mutations = 0). The prevalence is estimated using a ratio. The numerator is an estimate of the total number of initiators in the country without prior exposure to ARVs starting NNRTI-based regimens with HIV drug resistance mutations during the survey period. The denominator is an estimate of the total number of initiators in the country without prior exposure to ARVs starting NNRTI-based regimens during the survey period. The variance is calculated using linearization. A 95% confidence interval can be calculated using a standard Wald formula or by a logit transformation (default in Stata).

8.5 Data analysis: Outcomes 2a, 2b and 2c

Outcomes 2a, 2b and 2c are the prevalence of "no" prior exposure, "yes" prior exposure and "unknown" prior exposure to ARVs, respectively, among all ART initiators. The clinic sampling weight is defined in Section 8.1. The patient sampling weight for clinic i is defined as M_i divided by the number of initiators with recorded exposure status available from that clinic. The overall weight is the product of the clinic and patient sampling weights. All patients with available exposure status are defined as either having no prior exposure, prior exposure, or unknown prior exposure (categorical variable). The prevalence of each category is estimated using a ratio. The variance is calculated using linearization. A 95% confidence interval can be calculated using a standard Wald formula or by a logit transformation (default in Stata).

Notes

Notes

Notes

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 150719 6