This article describes the development of a novel sequential sampling method for the surveillance of transmitted HIV drug resistance by cross-sectional survey. Two commonly used sequential sampling methods are described and their applicability to the problem of classifying the prevalence of transmitted HIV drug resistance investigated. Both methods are rejected due to insufficient savings in sample size and operational complexity. A novel method is proposed and this is tested using computer-based simulation. This method provides useful sample size savings and operational simplicity and could provide the basis for a rapid and reliable survey method for classifying the prevalence of transmitted HIV drug resistance in circumstances where monitoring HIV drug resistance is an important issue, but resources do not allow full-scale surveillance to be established. The method is currently being used in several such settings.

Introduction

Surveys undertaken during the first decade of antiretroviral treatment (ART) scale-up in resource-rich countries suggest that, despite widespread availability of ART, the prevalence of transmitted HIV drug resistance (HIVDR) remained below 5% for many years [1–3]. Recent surveys from Europe and North America report prevalences of transmitted HIVDR between 5 and 15% [4–8]. In a few areas, there have been reports of prevalences of transmitted resistance to a particular drug class of >15% and some clinicians have recommended that the relevant drug class be excluded from first-line therapy [4–9].

The scale-up of ART in developing countries is not uniform within each country. This means that within-country geographical differences in the prevalence of transmitted HIVDR are likely, especially during the early phases of scale-up. The World Health Organization (WHO) recommends, therefore, stratifying surveillance of transmitted HIVDR by relatively small geographical areas within each country [10].

This article describes the development and technical background to a sequential sampling method that can be used to rapidly and reliably perform surveillance of transmitted HIVDR using repeated small cross-sectional surveys. Discussion is restricted to the application of sequential sampling methods to the problem of classifying the prevalence of transmitted HIVDR using small sample sizes (that is, \( n \leq 50 \)). Other teams were asked by the WHO Global HIV Resistance Surveillance Network to investigate the applicability of pooled testing and Bayesian classifiers to this problem. The sequential sampling method described here was selected by the WHO for the surveillance of transmitted HIVDR in low-resource settings after technical evaluation by the Centers for Disease Control and Prevention (CDC) and other members of the WHO Global HIV Resistance Surveillance Network.

The method described allows classification of the prevalence of transmitted HIVDR in a population without the large expenditure of resources needed for full-scale surveillance. The method classifies the prevalence of transmitted HIVDR into one of three categories consistent with the available historical data [1–9]: low prevalence (\( \leq 5\% \)), moderate prevalence (5–15%) and high prevalence (\( \geq 15\% \)).

As few individuals in each area are likely to be diagnosed with HIV during the period of recent infection, and because the laboratory costs of resistance testing are high, the development of the method described here was primarily informed by the need to make classifications using small sample sizes (that is, \( n \leq 50 \)).
Methods

Sequential sampling

Prevalence surveys commonly rely upon classical statistical approaches in which sample sizes are fixed in advance of data collection according to the expected prevalence and the level of precision or error required. A classical analysis such as estimation is performed using the totality of the collected data [11]. Sequential sampling is an alternative approach in which the sample size is not fixed in advance [12]. Instead, observations are collected individually and, after each observation, the accumulated data are examined to see whether or not a classification can be made. Sequential sampling combines data collection and data analysis into a single process or sampling plan. Sequential sampling plans are formulated to avoid the problems associated with repeated significance testing and multiple comparisons that can occur when using classical statistical approaches. This approach can considerably reduce both the sample size requirements and the data processing overheads of a survey.

Sequential sampling methods are best used in situations where the classification of the prevalence of a condition into categories or classes (for example, high and low prevalence) provides sufficient information on which to base decisions to take specific actions. This approach has been adopted, for example, in many national schistosomiasis and trachoma control programmes. Table 1 suggests actions appropriate to each prevalence category arising from the surveillance of transmitted HIVDR.

Methods

Sequential sampling

Prevalence surveys commonly rely upon classical statistical approaches in which sample sizes are fixed in advance of data collection according to the expected prevalence and the level of precision or error required. A classical analysis such as estimation is performed using the totality of the collected data [11]. Sequential sampling is an alternative approach in which the sample size is not fixed in advance [12]. Instead, observations are collected individually and, after each observation, the accumulated data are examined to see whether or not a classification can be made. Sequential sampling combines data collection and data analysis into a single process or sampling plan. Sequential sampling plans are formulated to avoid the problems associated with repeated significance testing and multiple comparisons that can occur when using classical statistical approaches. This approach can considerably reduce both the sample size requirements and the data processing overheads of a survey.

Sequential sampling methods are best used in situations where the classification of the prevalence of a condition into categories or classes (for example, high and low prevalence) provides sufficient information on which to base decisions to take specific actions. This approach has been adopted, for example, in many national schistosomiasis and trachoma control programmes. Table 1 suggests actions appropriate to each prevalence category arising from the surveillance of transmitted HIVDR.

In its simplest and most frequently used form sequential sampling is used to make binary classifications, but the technique can be extended to accommodate more granular (that is, into three or more classes) classifications.

Classifications made from sample surveys are probabilistic in nature and this will, inevitably, lead to occasional misclassification [13]. The typical behaviour of a two-class sampling plan is summarized in Table 2. The behaviour of a sampling plan is summarized more formally by an operating characteristic

Methods

Sequential sampling

Prevalence surveys commonly rely upon classical statistical approaches in which sample sizes are fixed in advance of data collection according to the expected prevalence and the level of precision or error required. A classical analysis such as estimation is performed using the totality of the collected data [11]. Sequential sampling is an alternative approach in which the sample size is not fixed in advance [12]. Instead, observations are collected individually and, after each observation, the accumulated data are examined to see whether or not a classification can be made. Sequential sampling combines data collection and data analysis into a single process or sampling plan. Sequential sampling plans are formulated to avoid the problems associated with repeated significance testing and multiple comparisons that can occur when using classical statistical approaches. This approach can considerably reduce both the sample size requirements and the data processing overheads of a survey.

Sequential sampling methods are best used in situations where the classification of the prevalence of a condition into categories or classes (for example, high and low prevalence) provides sufficient information on which to base decisions to take specific actions. This approach has been adopted, for example, in many national schistosomiasis and trachoma control programmes. Table 1 suggests actions appropriate to each prevalence category arising from the surveillance of transmitted HIVDR.

In its simplest and most frequently used form sequential sampling is used to make binary classifications, but the technique can be extended to accommodate more granular (that is, into three or more classes) classifications.

Classifications made from sample surveys are probabilistic in nature and this will, inevitably, lead to occasional misclassification [13]. The typical behaviour of a two-class sampling plan is summarized in Table 2. The behaviour of a sampling plan is summarized more formally by an operating characteristic

Table 1. Prevalence categories and appropriate actions for prevalence of HIV drug resistance

<table>
<thead>
<tr>
<th>Category</th>
<th>Prevalence</th>
<th>Appropriate action(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>≤5%</td>
<td>Repeat survey after 2 years.</td>
</tr>
<tr>
<td>Moderate</td>
<td>5–15%</td>
<td>Review antiretroviral therapy programme monitoring data for the relevant geographical area and investigate potential problems with regards to several factors: continuous access to services; drug supply; drug quality; prescribing practice; compliance toxicity and/or adverse events; drug sales by staff and/or patients; drug sharing; treatment failures. Focus on the drugs for which the prevalence of resistance mutations was classified as moderate and take remedial action where necessary. Review specimen handing/laboratory data/laboratory practice to confirm that no contamination occurred. Review eligibility data to confirm that no specimens came from people with previous exposure to antiretroviral drugs. Repeat survey in the following year.</td>
</tr>
<tr>
<td>High</td>
<td>≥15%</td>
<td>As for moderate (above) focusing on the drugs for which the prevalence of resistance mutations was classified as high and taking remedial action where necessary. Repeat survey in the following year and in additional sites in the same year. If prevalence is classified as ≥15% in two consecutive years or in multiple sites in one area in the same year then expert review is required to decide whether changes in guidelines for initial antiretroviral therapy regimes are appropriate and feasible and if HIV drug resistance testing should be performed prior to treatment.</td>
</tr>
</tbody>
</table>

Table 2. Typical behaviour of two- and three-class sampling plans

<table>
<thead>
<tr>
<th>Sampling plan</th>
<th>Classified prevalence</th>
<th>Very low</th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
<th>Very high</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-class</td>
<td>Low</td>
<td>Always</td>
<td>Often</td>
<td>Sometimes</td>
<td>Seldom</td>
<td>Never</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Never</td>
<td>Seldom</td>
<td>Sometimes</td>
<td>Often</td>
<td>Always</td>
</tr>
<tr>
<td>Three-class</td>
<td>Low</td>
<td>Always</td>
<td>Often</td>
<td>Sometimes</td>
<td>Seldom</td>
<td>Never</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>Never</td>
<td>Seldom</td>
<td>Most often</td>
<td>Seldom</td>
<td>Never</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Never</td>
<td>Seldom</td>
<td>Sometimes</td>
<td>Often</td>
<td>Always</td>
</tr>
</tbody>
</table>
Figure 1. Typical behaviours of two- and three-class sampling plans

(A) Operating characteristic curve of a two-class sequential sampling plan. (B) Typical behaviour of a three-class sequential sampling plan.

Sequential sampling for transmitted HIV drug resistance

Candidate methods

Two sequential sampling methods were investigated for their applicability to the problem of classifying transmitted HIVDR: lot quality assurance sampling (LQAS) with multiple sampling plans; and binomial sequential sampling (BSS) for three alternative hypotheses.

LQAS with multiple sampling plans is an extension to the standard LQAS method [16]. LQAS is widely used in the manufacturing industry to judge the quality of a lot (batch) of items. In the industrial context, LQAS is used to identify lots that are likely to contain an unacceptably large number of defective items. In the public health context, LQAS is used to identify populations with low levels of service (for example, vaccination) coverage or high prevalences of disease [17]. LQAS produces data that are easy to analyse. Data analysis is performed as data are collected and consists of counting the number of defects (for example, unvaccinated children or cases of the condition of interest) in the sample and checking whether this exceeds a pre-determined number.

LQAS data are collected and analysed using a sampling plan that specifies a maximum sample size \( n_{\text{max}} \) and the maximum number of defects allowed in the sample \( d \). Sampling plans are developed by specifying a classification system (for example, the levels of prevalence that define high and low prevalence situations) and acceptable probabilities of classification error. Suitable values for \( n_{\text{max}} \) and \( d \) are usually found by performing an exhaustive search of cumulative binomial probabilities for combinations of \( n_{\text{max}} \) and \( d \) that provide acceptable levels of type I and type II error. Different probability distributions or computer-based simulation with empirical distributions may also be used to develop sampling plans. Using an LQAS sampling plan in the field is straightforward. Sampling stops when either the maximum sample size \( n_{\text{max}} \) is met or the number of defects allowed in the sample \( d \) is exceeded. If the maximum sample size \( n_{\text{max}} \) is met without the number of defects allowed in the sample \( d \) being exceeded, the sampled population is classified as having a low prevalence of defects. If the number of defects allowed in the sample \( d \) is
exceeded, the sampled population is classified as having a high prevalence of defects.

LQAS can be extended to provide more granular classifications by using more than one sampling plan: each sampling plan uses the same maximum sample size ($n_{max}$), but different allowable numbers of defects (for example, $d_1$ and $d_2$ where $d_1 > d_2$) [18]. Two sampling plans are required for a three-class problem: $n_{max}, d_1$ provides high prevalence classifications (sampling may stop before $n_{max}$ if $d_1$ is exceeded allowing sample size savings at high prevalences); $n_{max}, d_2$ provides moderate prevalence classifications in samples not already classified by the previous sampling plan ($n_{max}, d_1$). Remaining samples (that is, those in which neither $d_1$ nor $d_2$ are exceeded) are classified as low prevalence. The process for a three-class problem is shown in Figures 2 & 3A.

LQAS with multiple sampling plans has been used successfully for classifying the prevalence of trachoma and schistosomiasis in community and school-based surveys and has proved simple to use and capable of making accurate and reliable classification using small sample sizes [18–20]. LQAS with multiple sampling plans provides sample size savings for high prevalence classifications, but requires the collection of a complete sample for both moderate and low prevalence classifications. As low prevalences of transmitted HIVDR are expected in the early years of use when laboratory costs are likely to be highest, sample size savings in high prevalence situations will not result in significant immediate cost savings. This method was, therefore, rejected as a candidate for the problem of classifying the prevalence of transmitted HIVDR.

BSS is commonly used in agriculture to assess the need for field- or herd-level interventions, such as pesticide treatment of crops or mass treatment of livestock [13]. In its most commonly used form, BSS provides a binary classification. In a BSS sampling plan, the maximum sample size ($n_{max}$) is fixed and the number of defects ($d$) allowed in the sample varies with the size of the sample collected and differs for low and high prevalence classifications. BSS plans have three stopping rules (see Box 1).

A BSS plan can be represented graphically by two parallel lines corresponding to the number of defects allowed in the sample for the upper and lower classification thresholds at different sample sizes [14] (Figure 3B). The two lines in Figure 3B represent $d_{lower\text{_triage\_level}}=bn+b_1$ and $d_{upper\text{_triage\_level}}=bn+b_2$ where $d_{lower\text{_triage\_level}}$ is the threshold value for the number of defects ($d$) allowed in a sample of size $n$ for the lower classification threshold and $d_{upper\text{_triage\_level}}$ is the threshold value for the number of defects ($d$) allowed in a sample of size $n$ for the upper classification threshold. The slope of the two lines ($b$), is calculated as:

$$b=\frac{\ln(q_1/q_2)}{\ln(p_2q_1/p_1q_2)}$$

(1)

The size of sample at any time during sampling is given by $n$, while $b_1$ is the intercept of the stopping rule line for the lower classification threshold. The value of $b_1$ is calculated as:

$$b_1=\frac{-\ln[(1-\alpha)/\beta]}{\ln(p_1q_2/p_2q_1)}$$

(2)

The intercept of the stopping rule line for the upper classification threshold is $b_2$, calculated as:

$$b_2=\frac{-\ln[(1-\beta)/\alpha]}{\ln(p_1q_2/p_2q_1)}$$

(3)

Figure 2. Using two lot quality assurance sampling plans to provide three classifications
The two lines represent the threshold values for the stopping rules to classify a population as either low or high prevalence. A maximum sample size prevents sampling from continuing indefinitely and is calculated using the equation:

$$n_{\text{max}} = \frac{-\ln[1-p_2/\alpha] \times \ln[1-q_2/\beta]}{\ln[p_2/p_1] \times \ln[q_2/q_1]}$$  (4)

Where $p_1$ is the lower prevalence threshold (as a proportion); $p_2$ is the upper prevalence threshold (as a proportion); $q_1 = 1 - p_1$; $q_2 = 1 - p_2$; $\alpha$ is the acceptable type I error; and $\beta$ is the acceptable type II error.

An additional stopping rule defining the minimum sample size ($n_{\text{min}}$) can also be specified. This is chosen to ensure that a decision is possible using the stopping rules at any given sample size. This may be the minimum value of $n$ for which $d_{\text{lower triage level}}$ is greater than $d_{\text{upper triage level}}$.

Figure 3. Graphical representation of the sequential sampling methods discussed in the text

(A) Lot quality assurance sampling (two sampling plans). (B) Binomial sequential sampling (BSS). (C) BSS3 (two BSS sampling plans). (D) Truncated sequential sampling.
than zero. This choice of $n_{\text{min}}$ will, however, tend to increase the sample size required to make high prevalence classifications. Savings in sample size can be achieved by using this value for low prevalence classifications (this cannot be reduced as counts may not be negative) and a lower value for high prevalence classifications. There are, however, no formal methods for deciding appropriate values of $n_{\text{min}}$.

As with the LQAS-derived method, binomial sequential sampling for three alternative hypotheses (BSS3) uses two sampling plans for a three-class problem (Figure 3C) [14,21]. Combining two sampling plans in this way may, however, yield anomalous results at small sample sizes [12]. The use of a third sampling plan has been suggested as a means to overcome this problem [22]. BSS3 provides sample size savings at low levels of prevalence, but is complicated to apply in the field. The use of a third sampling plan, required to improve performance with small sample sizes, adds further complication [12,22]. The operational complexity of the method led to it being rejected as a candidate for the problem of classifying the prevalence of transmitted HIVDR.

A new method: truncated sequential sampling

LQAS with multiple sampling plans was rejected due to a lack of sample size savings in low prevalence situations. BSS3 was rejected due to its operational complexity. It was decided, therefore, that a new method would need to be developed with three attributes: small sample size requirements (that is, $n_{\text{min}} \leq 50$), sample size savings in low prevalence situations and operational simplicity.

Examination of plots of the cumulative number of defects against sample size when using BSS plans in populations with different levels of prevalence suggested that modified BSS plans might be capable of yielding three classifications if the maximum sample size ($n_{\text{max}}$) were to be reduced considerably below that returned by Equation 4. Such truncated sampling plans would, like BSS plans, have three stopping rules (see Box 1). Populations that remain unclassified at the truncated maximum sample size ($n_{\text{max}}$) receive a moderate prevalence classification. This is illustrated in Figure 3D. Such a method would have low sample size requirements, provide sample size savings at low levels of prevalence, and retain the operational simplicity of the original BSS method. Figure 4 shows the behaviour of a truncated sampling plan sampling from populations with true prevalences of 3%, 11% and 18%, respectively.

It should be noted that, at truncated maximum sample sizes, the $\alpha$ and $\beta$ parameters in Equations 2 and 3 are nominal and do not represent the expected levels of type I and type II errors for a particular sampling plan; these parameters act only to specify the intercepts of and distance between the upper and lower classification thresholds of the sampling plan. The behaviour of a sampling plan with respect to the proportion of correct classifications is strongly determined by the minimum sample size ($n_{\text{min}}$) and the maximum sample size ($n_{\text{max}}$), which is likely to be considerably smaller than that required by the LQAS, BSS and BSS3 methods.

Selection and testing of sampling plans for truncated sequential sampling

The suitability of this truncated sequential sampling (TSS) method was tested using computer-based simulation with simulated and real survey data. The performance of a truncated sampling plan depends on the size and shape of the zone of indecision, which is described by the minimum and maximum sample sizes and the slopes and intercepts of the classification thresholds. A set of candidate sampling plans was

---

Box 1. Stopping rules employed in sequential sampling techniques

**Binomial sequential sampling**
- If the number of defects found in the sample falls below a threshold value ($d_{\text{lower triage level}}$) for the current sample size ($n$) then sampling stops and the population is classified as having a low prevalence of disease.
- If the number of defects found in the sample exceeds a threshold value ($d_{\text{upper triage level}}$) for the current sample size ($n$) then sampling stops and the population is classified as having a high prevalence of disease.
- If the maximum sample size ($n_{\text{max}}$) is met and no classification has been made then sampling stops and the population is classified as either having a high or low prevalence of disease depending upon the proximity of the number of defects found to $d_{\text{lower triage level}}$ and $d_{\text{upper triage level}}$ at $n_{\text{max}}$.

**Truncated sequential sampling**
- If the number of defects found in the sample falls below a threshold value ($d_{\text{lower triage level}}$) for the current sample size ($n$) then sampling stops and the population is classified as having a low prevalence of disease.
- If the number of defects found in the sample exceeds a threshold value ($d_{\text{upper triage level}}$) for the current sample size ($n$) then sampling stops and the population is classified as having a high prevalence of disease.
- If the maximum sample size ($n_{\text{max}}$) is met and no classification has been made then sampling stops and the population is classified as having a moderate prevalence of disease.
generated using the following procedure. Firstly, upper and lower prevalence thresholds were fixed at $p_1 = 5\%$ and $p_2 = 15\%$ for all candidate sampling plans and the slope of the classification thresholds calculated using Equation 1. The intercepts of the classification thresholds were then varied by specifying different values (that is, 0.05, 0.025 and 0.01) for the $\alpha$ and $\beta$ parameters and calculated using Equations 2 and 3. Nine subsets of candidate sampling plans were created in this way. Minimum and maximum sample sizes were varied systematically within each of these nine subsets of candidate sampling plans. The minimum sample size was systematically varied between one and 40. The maximum sample size was systematically varied between 25 and 60.

A set of 500 test populations each containing 1,000 data points was generated. The prevalences in the test populations followed the distribution shown in Table 3 with prevalence selected from a uniform distribution within each prevalence band [23]. This distribution is consistent with available historical data [1–9]. No large studies of transmitted HIVDR have reported prevalences >30% with the majority reporting prevalences <20% [1–9]. In some areas and groups in resource-rich countries, however, prevalences of 20–30% have been reported [24,25].

Each of the candidate sampling plans was used to sample each of the 500 test populations 2,000 times using simple random sampling (that is, each candidate plan was used in 1 million simulated surveys). The performance of each candidate sampling plan was evaluated using an extension of the confusion matrix approach that is commonly used to evaluate classification models developed using discriminant analysis [22]. This involved producing a contingency table of true prevalence class versus classified prevalence class and calculating the proportion of correct classifications for the three prevalence classes separately and combined, as well as the proportion of gross misclassifications (that is, true low prevalence populations being classified as high prevalence populations and vice versa). Candidate sampling plans were also evaluated by examination of the probability of classification and average (mean) sample size curves.

The criteria and standards used to select suitable sampling plans are shown in Table 4. The standards for the classification proportions were arrived at in consultation with members of the WHO Global HIV Resistance Surveillance Network. The standard for the maximum sample size (that is, $n_{\text{max}} \leq 50$) was selected after a review of data from HIV seroprevalence surveys conducted in developing countries [26]. This review revealed that, in area-specific surveys sampling between 200 and 10,000 individuals, there were generally fewer than 50 HIV-positive individuals who met a proxy criteria for newly diagnosed, drug-naive, recently HIV-infected persons [27].

**Results**

Systematic evaluation of the candidate sampling plans identified a set of parameters that defined a sampling plan that met the selection criteria and standards (outlined in Table 4) at the smallest maximum sample size: lower prevalence threshold ($p_1$) 5%; upper prevalence threshold ($p_2$) 15%; $\alpha$ 0.010 (nominal); $\beta$ 0.025 (nominal); $n_{\text{min}}$ 14; and $n_{\text{max}}$ 47. Table 4 shows the
Table 4. Criteria and standards used to select suitable truncated sequential sampling plans and the performance of the selected TSS plan in simulated surveys

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Standard</th>
<th>Performance of selected sampling plan*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Simulated data</td>
</tr>
<tr>
<td>Proportion of correct classifications</td>
<td></td>
<td>Overall, %</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low, %</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moderate, %</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High, %</td>
</tr>
<tr>
<td>Gross misclassification†</td>
<td></td>
<td>Overall number/number of surveys</td>
</tr>
<tr>
<td>Maximum sample size</td>
<td></td>
<td>≤50</td>
</tr>
</tbody>
</table>

*The selected plan is defined by \(p_1=5\%\), \(p_2=15\%\), \(\alpha=0.01\), \(\beta=0.025\), \(n_{min}=14\), \(n_{max}=47\). †The proportion of true low prevalence populations being classified as high prevalence populations and vice versa. ‡Average (mean) sample size by prevalence is shown in Figure 6. NRTI, nucleoside reverse transcriptase inhibitor.

Figure 5. Probability of classification curve for the selected truncated sequential sampling plan

Parameters chosen for the truncated sequential sampling plan: \(p_1=5\%\), \(p_2=15\%\), \(\alpha=0.01\), \(\beta=0.025\), \(n_{min}=14\), \(n_{max}=47\).
performance of this sampling plan with regard to the selection criteria and standards.

Figure 5 shows the probability of classification curve for this sampling plan. This curve follows the behaviour outlined in Figure 1B and Table 2 for a three-class sampling plan, although there is a small probability (~1.8%) of the method returning a moderate prevalence classification at prevalences as high as 30%. Gross misclassifications are rare (that is, <5/10,000 surveys).

Figure 6 shows the average (mean) sample size required to make a classification using this sampling plan at different levels of prevalence. The selected sampling plan provides some sample size saving (that is, when compared with the maximum sample size for the sampling plan) at low prevalences and considerable sample size savings at high prevalences.

The computer-based simulations with simulated data confirmed that the truncated method was capable of meeting the requirements of the transmitted HIVDR application. The selected sampling plan was further tested using computer-based simulations using data from 28 cohorts of recently infected and diagnosed cases of HIV infection from the USA and Canada. Testing was undertaken on two variables in these datasets: ‘NRTI direct’ and ‘NRTI all’. NRTI direct included major mutations directly associated with resistance to nucleoside reverse transcriptase inhibitors [28]; the prevalence ranged between 0% and 7.2% with a median prevalence of 2.0%. NRTI all included mutations directly or indirectly associated with resistance to nucleoside reverse transcriptase inhibitors (including T215 revertent mutations) [28]. For this variable, the prevalence ranged between 0% and 18.2% with a median prevalence of 6.5%.

Each cohort was sampled and classified 10,000 times using the selected sampling plan. Results of this further testing were broadly consistent with the results obtained with simulated data (Table 4).

Discussion

The TSS method presented in this article could provide the basis for a rapid and reliable survey method for classifying the prevalence of transmitted HIVDR. Such a method could be used in circumstances in which the monitoring of transmitted HIVDR is an important issue, but where resources do not allow full-scale surveillance to be established. The low sample size requirement means that such a method would be suitable for use in low-resource environments such as developing countries where ART is being scaled up. The method is currently being used in several such settings.

The application of the method in the field does not require the use of computers for data entry and analysis. Data analysis consists of applying stopping rules to collected data. This may be done using graphical devices similar to those shown in Figures 3D&4 or by the use of a tabular data collection and analysis form such as that shown in Figure 7. Application of the sampling plan using a tabular data collection and analysis form involves recording the cumulative number of cases found in the sample against the size of the survey sample taken as data are collected. The cumulative number of cases found is compared with upper and lower limits for the cumulative number of cases found for the size of the survey sample taken. If the cumulative number of cases found in the survey sample falls below the lower classification threshold for the size of the survey sample taken, then sampling stops and the population is classified as low prevalence. If the cumulative number of cases found in the survey sample exceeds the upper classification threshold for the size of survey sample taken, then sampling stops and the population is classified as high prevalence. If the maximum sample size is reached without either classification threshold being crossed then sampling stops and the population is classified as moderate prevalence.

The choice of data collection and analysis device should be informed by the familiarity with the graphical representation of number (that is, the ability to accurately plot points onto graphs and to read values from graphs using the two-dimensional Cartesian coordinate system)
amongst survey staff, which should not be taken for
granted and is likely to vary from setting to setting [29].

The TSS method is not specific to the problem of clas-
sifying the prevalence of transmitted HIVDR and could
be adapted for other applications requiring triage classifi-
cations of prevalence using small sample sizes, although
at present the development of sampling plans requires the
use of computer-based simulation techniques. The use of
computer-based simulation does, however, have the
advantage of allowing sampling plans to be developed
and selected with regard to their behaviour in situations
where complex sampling schemes will be used (for
example, community surveys selecting communities,
households, and individuals within households) and
when information regarding the empirical distribution of
prevalences is available. This is also a limitation and users
of the method should be aware that a sampling plan
developed in this way may be suboptimal when applied
in situations where the sampling scheme or the empirical
distribution of prevalences differ from those specified in
the simulations used to develop the sampling plan.

The development of simple tools (for example,
formulae, algorithms or computer programs) to assist
the development of sampling plans could be the
subject of further work. It is also likely that some
improvements in performance could be achieved by

---

**Figure 7:** Tabular data collection and analysis form for the selected truncated sequential sampling plan

<table>
<thead>
<tr>
<th>Location:</th>
<th>Drug class:</th>
<th>Start date:</th>
<th>End date:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Sample number</th>
<th>Running total of specimens with a listed mutation</th>
<th>Upper limit</th>
<th>Sample number</th>
<th>Running total of specimens with a listed mutation</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ND</td>
<td>ND</td>
<td>25</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>ND</td>
<td>ND</td>
<td>26</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>ND</td>
<td>ND</td>
<td>27</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>ND</td>
<td>ND</td>
<td>28</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>ND</td>
<td>ND</td>
<td>29</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>ND</td>
<td>ND</td>
<td>30</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>ND</td>
<td>ND</td>
<td>31</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>ND</td>
<td>ND</td>
<td>32</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>ND</td>
<td>ND</td>
<td>33</td>
<td>ND</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>ND</td>
<td>ND</td>
<td>34</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>ND</td>
<td>ND</td>
<td>35</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>ND</td>
<td>ND</td>
<td>36</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>13</td>
<td>ND</td>
<td>ND</td>
<td>37</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>ND</td>
<td>5</td>
<td>38</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>15</td>
<td>ND</td>
<td>5</td>
<td>39</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td>ND</td>
<td>5</td>
<td>40</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>17</td>
<td>ND</td>
<td>5</td>
<td>41</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>18</td>
<td>ND</td>
<td>5</td>
<td>42</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>19</td>
<td>ND</td>
<td>5</td>
<td>43</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>20</td>
<td>ND</td>
<td>5</td>
<td>44</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>21</td>
<td>ND</td>
<td>5</td>
<td>45</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>22</td>
<td>ND</td>
<td>5</td>
<td>46</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>23</td>
<td>ND</td>
<td>5</td>
<td>47</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>24</td>
<td>ND</td>
<td>5</td>
<td>STOP</td>
<td>STOP</td>
<td>STOP</td>
</tr>
</tbody>
</table>

ND = no decision - continue sampling
the use of convergent [30] or divergent [31] classification boundaries, which have been demonstrated to introduce sample size savings in applications requiring binary classifications [13], again this requires further investigation.

Acknowledgements

The authors wish to thank Stephano Lazzari (WHO, Lyon, France), Theresa Diaz-Vargas (Centers for Disease Control and Prevention, Atlanta, USA), and Meade Morgan (Centers for Disease Control and Prevention, Atlanta, USA) for their support during the development and testing of the TSS method. We also thank Darwin Minassian (Institute of Ophthalmology, University College London, London, UK), Rebecca Grais (EpiCentre, Paris, France), and Joe Abramson (Hebrew University, Jerusalem, Israel) for reviewing early drafts of this article and the International Trachoma Initiative (New York, USA) for their initial funding of the development of the TSS method. International Medical Press is also acknowledged for their support with publishing this manuscript. Some authors are staff members of the WHO. The authors alone are responsible for the views expressed in this publication and they do not necessarily represent the decisions or stated policy of the WHO.

Disclosure statement

The authors declare no conflicts of interest.

References


---

Accepted for publication 19 September 2007