



Published in final edited form as:

Epidemiology. 2012 March ; 23(2): 293–300. doi:10.1097/EDE.0b013e3182459455.

Multidrug Resistance among New Tuberculosis Cases: Detecting Local Variation through Lot Quality-Assurance Sampling

Bethany Lynn Hedt¹, Frank van Leth^{2,3}, Matteo Zignol⁴, Frank Cobelens², Wayne van Gemert⁴, Nguyen Viet Nhung⁵, Svitlana Lyepshina⁶, Saidi Egwaga⁷, and Ted Cohen^{8,9}

¹Department of Global Health and Social Medicine, Harvard Medical School, Boston ²Department of Global Health, Academic Medical Centre, Amsterdam Institute for Global Health and Development, Amsterdam, The Netherlands ³KNCV Tuberculosis Foundation, The Hague, The Netherlands ⁴Stop TB Department, World Health Organization, Geneva, Switzerland ⁵National Tuberculosis Control Program, National Lung Hospital, Hanoi, Vietnam ⁶Donetsk State Medical University, Donetsk, Ukraine ⁷National Tuberculosis Control Program, Tanzania ⁸Division of Global Health Equity, Brigham and Women's Hospital, Boston ⁹Department of Epidemiology, Harvard School of Public Health, Boston

Abstract

Background—Current methodology for multidrug-resistant TB (MDR TB) surveys endorsed by the World Health Organization provides estimates of MDR TB prevalence among new cases at the national level. On the aggregate, local variation in the burden of MDR TB may be masked. This paper investigates the utility of applying lot quality-assurance sampling to identify geographic heterogeneity in the proportion of new cases with multidrug resistance.

Methods—We simulated the performance of lot quality-assurance sampling by applying these classification-based approaches to data collected in the most recent TB drug-resistance surveys in Ukraine, Vietnam, and Tanzania. We explored three classification systems—two-way static, three-way static, and three-way truncated sequential sampling—at two sets of thresholds: low MDR TB = 2%, high MDR TB = 10%, and low MDR TB = 5%, high MDR TB = 20%.

Results—The lot quality-assurance sampling systems identified local variability in the prevalence of multidrug resistance in both high-resistance (Ukraine) and low-resistance settings (Vietnam). In Tanzania, prevalence was uniformly low, and the lot quality-assurance sampling approach did not reveal variability. The three-way classification systems provide additional information, but sample sizes may not be obtainable in some settings. New rapid drug-sensitivity testing methods may allow truncated sequential sampling designs and early stopping within static designs, producing even greater efficiency gains.

Conclusions—Lot quality-assurance sampling study designs may offer an efficient approach for collecting critical information on local variability in the burden of multidrug-resistant TB. Before this methodology is adopted, programs must determine appropriate classification thresholds, the

Corresponding Author: Bethany L Hedt; Department of Global Health and Social Medicine; 641 Huntington Ave, Boston, MA 02115; (p) 617-432-1056; bethhedt@gmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

most useful classification system, and appropriate weighting if unbiased national estimates are also desired.

In most high-incidence settings, the control of tuberculosis (TB) depends on diagnosis by microscopic examination of sputum, and treatment with standard combinations of antibiotics. Because drug resistant forms of TB are not identifiable by microscopy and often not effectively treated with first-line drugs, the appearance of drug-resistant TB poses a serious challenge for control of this disease. According to the most recent published estimates, there were 440,000 incident cases of multidrug-resistant TB in 2008 (with resistance to at least isoniazid and rifampin).¹

Until 1994, data required to estimate the global distribution of drug-resistant tuberculosis were unavailable. Recognizing the need for improved assessment, the World Health Organization and the International Union Against Tuberculosis and Lung Diseases launched an ambitious project to document the global burden of these infections. To meet this goal, the WHO published survey guidelines² and established an international network of reference laboratories to provide quality assurance for local laboratories. These efforts have resulted in the publication of four reports³ and a recent update¹ documenting the state of resistance to anti-tuberculosis drugs. In total, 114 countries have contributed either surveillance data (for countries with capacity for comprehensive and continuous testing) or survey data (where continuous testing is not feasible and thus sampling must be used).

The most recent international guidelines recommend that all countries begin continuous surveillance for drug-resistant TB among retreatment patients (i.e. patients who had been exposed to at least one month of TB drugs previous to their current diagnosis of TB).² This recommendation is based on both clinical and public health considerations. Because patients requiring retreatment are more likely to have drug-resistant disease than those initiating a first course of TB therapy, drug-susceptibility testing targeted to the retreatment group will yield a higher fraction of detected drug resistance. This should increase the rate at which patients with drug-resistant TB receive appropriate second-line therapy, and therefore may reduce transmission of drug resistance in the community.

While continuous surveillance of retreatment TB cases will continue to be the top priority, surveys among new cases are necessary in places where continuous surveillance is not done, in order to understand the role of the transmission of drug-resistant TB. Drug resistance detected among people who have not previously been treated for TB unambiguously signals transmitted (primary) resistance, whereas resistance detected among cases presenting for retreatment is a mixture of acquired and transmitted resistance. The relative contributions of acquired resistance and transmitted resistance change over the course of an epidemic, and interventions to prevent transmission and acquisition of resistance differ. Accordingly, testing new cases for drug resistance clarifies the role of transmission in a drug-resistant TB epidemic and helps to inform the design of an appropriate health system response.

Broadly, the goal of drug-resistant TB surveys has been to produce point estimates of the fraction of incident TB cases resistant to first-line antibiotics over large geographic areas (i.e. an entire country or large section of a country). Accordingly, survey guidelines specify approaches for collecting representative samples of cases to produce valid regional estimates. This survey approach is usually based on cluster sampling with fixed sample sizes per cluster. The method inherently prioritizes the generation of regional estimates of drug resistance over the identification of local variation because the cluster-level sample sizes are too small to generate stable measurements. While these national estimates provide the data necessary to assess the global burden of resistant disease, they are less useful for helping policymakers determine local variations in the burden of drug resistance. Understanding

patterns of local variation in resistance would allow for focused follow-up investigations of anomalous regions and would facilitate the implementation of targeted interventions and efficient use of limited resources.

We describe an alternative approach for drug resistance surveys among new TB cases that aims to generate actionable information for policymakers at the national and sub-national level. We explore the use of lot quality-assurance sampling to classify the level of resistance into pre-determined categories within particular geographic areas.⁴ Several considerations recommend lot quality-assurance sampling. First, smaller samples are generally enough to classify sites as compared with precise point estimation. Second, the classifications can be linked to recommendations for program response, improving the usefulness of drug-resistant TB surveillance. Finally, we chose to explore lot quality-assurance sampling because of the precedent in using this design for HIV drug-resistance surveillance.⁵ We use data collected for previously reported drug-resistance surveys in both high and low drug-resistant TB settings to simulate the performance of a lot quality-assurance sampling-based surveillance approach. These existing data are useful because they were collected from representative samples of new patients, allowing us to model the application of alternative survey strategies that capture subsets of these patients. Our primary aims are to determine whether the lot quality-assurance sampling methodology can efficiently provide local information about transmitted drug resistance, and to identify the utility of lot quality-assurance sampling in settings with high or low levels of drug-resistant TB.

Methods

We explore the utility of lot quality-assurance sampling classification for drug-resistant TB surveys among new cases using three classification methods: 1) static two-way classification (the classical approach, which classifies sites as high versus low), 2) static three-way classification (allowing for an additional category of moderate), and 3) truncated sequential sampling three-way classification (an adaptive design that also allows for three levels of classification, but potentially with smaller sample sizes). We illustrate these methods with application to data collected as part of drug-resistant TB surveillance activities in Ukraine, Vietnam, and Tanzania. These studies applied published guidelines for the surveillance of drug-resistant tuberculosis, collecting samples from population-representative samples of smear-positive TB patients.²

We focus our analysis and discussion on the classification of multidrug-resistant TB. For each lot quality-assurance sampling method and set of thresholds, we limit our analysis to sites or groups of sites with sufficient samples for classification. For the each site and method, we randomly pull samples without replacement until a classification can be made in accordance to the lot quality-assurance sampling system. If the sample for a site or group of sites is exhausted before classification can be made, that site is dropped and not reported here.

Data

Ukraine: From 1 July 2005 through 30 June 2006, all sputum smear-positive TB cases in the Donetsk oblast were included in a drug-resistant TB survey. Details are published elsewhere.⁶ We limit our analysis to the 999 new cases sampled at 53 treatment sites. Of these sites, 43 (comprising 926 patients) were also sorted into five groups defined by geographical proximity; the excluded sites were very small and not in close proximity to the sites included in this analysis.

Vietnam: From 1 August through 31 October 2001, researchers attempted to enroll 23 consecutively registered new smear-positive TB cases from 40 clusters (i.e. TB diagnostic

and treatment facilities) in the south of Vietnam.⁷ We limit our analysis to 888 new patients who were grouped into 20 provinces in the original survey. These 20 provinces were also aggregated into 5 groups defined by geographical proximity and demographic and socio-economic similarity.

Tanzania: A national drug-resistance survey was conducted between July 2006 and August 2007 with 1167 smear-positive TB cases from 40 clusters randomly selected in the original survey.⁸ We limited the analysis to the 909 new cases with both culture and sensitivity results, combined across 24 sites and 5 groups.

Thresholds

We explore two sets of classification thresholds for this paper. The first set of thresholds had low prevalence of multiple-drug-resistant TB (MDR TB) defined as 2% or less among new cases, and high prevalence defined as 10% or more. The second set of thresholds had low prevalence of MDR TB defined as 5% or less among new cases and high prevalence defined as 20% or more. These choices are intended only to illustrate the method, the selection of threshold values should reflect both the expected local prevalence as well as the types of interventions that will be considered when areas are classified with high levels of resistance. (The selection of thresholds is considered further in the discussion section below.)

Classification approaches

Static Two-Way Classification—The first design we test is the classical two-way lot quality-assurance sampling classification. This design uses a fixed sample size and a corresponding decision rule (d) (namely, d or more cases of resistant disease, is an area classified as high level of resistance; otherwise the area is classified as low level of resistance). To determine the appropriate sample size, we specify an acceptable error rate for classifying an area with (truly) low MDR TB as high MDR TB (p_{lh}) of 10% (Table 1). This is often referred to as the β -error or provider error in classical two-way lot quality-assurance sampling. We also specify the acceptable error rate for classifying an area with truly high MDR TB as low (p_{hl}) of 5%. This is referred to as the α -error or consumer error. We chose a lower α -error than β -error because the consequences of a false negative (missing a truly high MDR TB area) would likely be worse than a false positive (misclassifying a truly low MDR TB area as high).

For the first set of thresholds (optimized to classify areas with MDR TB prevalence below 2% as low and above 10% as high), these choices result in a sample size of 76, and corresponding decision rule of 4. That is, if 4 or more of the 76 persons have MDR TB, then the area is classified as high MDR TB prevalence. If fewer than 4 have MDR TB, then the area is classified as low MDR TB. For the 5%/20% thresholds, the necessary sample size is only 44, with a corresponding decision rule of 5. The operating characteristic curves (OCCs) for this classification scheme are displayed in the eAppendix (<http://links.lww.com>).

Static Three-Way Classification—The second design is a three-way classification scheme with a fixed sample size and two corresponding decision rules (d_l and d_h) such that if fewer than d_l cases of resistant disease are observed, the area is classified as having a low level of resistance, if d_h or more cases are observed, the area is classified as having a high level of resistance, and the rest are classified as having a moderate level of resistance. In addition to specifying p_{lh} and p_{hl} , (set to 10% and 5% as above) we must also specify our desired probability of correct classification at the lower threshold (p_{ll}), upper threshold (p_{hh}), and moderate threshold (p_{mm} , where the moderate threshold, p_m , is the midpoint between the upper and lower thresholds). We set these parameters to 80% ($p_{ll}, p_{hh}, p_{mm} \geq 0.80$).

Under these constraints, for the 2%/10% thresholds we need a sample size of 169, with $d_l=6$ and $d_h=14$. If fewer than 6 of the 169 people have MDR TB, the area is classified as having low levels of MDR TB, and if 14 or more have MDR TB, the area is classified as having high levels. If between 6 and 13 people have MDR TB, the area is classified as having moderate levels. For the 5%/20% threshold, the necessary sample size is 94, with decision rules d_l and d_h equal to 7 and 16, respectively. OCCs are presented in the eAppendix (<http://links.lww.com>).

Truncated Sequential Sampling for Three-Way Classification—The final design is a three-way classification adapted from the truncated sequential sampling design for HIV drug-resistance surveillance.⁹ With this adaptive sampling approach, sampling may be stopped early if the area can be classified into either the high or low categories before the maximum sample size is reached. The truncated sequential sampling design starts with a minimum sample size. If more than d_{hn} have resistance, the area is classified as high; if fewer than d_{ln} have resistance, then the area is classified as low. If the number of resistant samples is between d_{ln} and d_{hn} , then an additional individual is sampled and tested. At this next stage, the number of resistant individuals is compared with new thresholds, namely d_{ln+1} and d_{hn+1} . This iterative process continues until an area is classified as high or low, or until reaching the maximum sample size, at which point the area is classified as moderate.

The process of determining the minimum and maximum sample sizes, and the corresponding decision rules at each stage, are described in detail in the eAppendix (<http://links.lww.com>). Briefly, we use the same constraints as we describe for the static three-way classification, namely $p_l, p_m, p_h \geq 0.80$, $p_{lh}=0.1$, and $p_{hl}=0.05$. We evaluate the performance of the system at the same prevalences as for the static three-way classification, namely ($p_l=0.02$, $p_m=0.06$ and $p_u=0.10$) and ($p_l=0.05$, $p_m=0.125$ and $p_u=0.20$).

The classification thresholds are determined by the lines:

$$d_{hn} = \log(q_l/q_h) / \log((p_h * q_l) / (p_l * q_h)) * n + \log(1 - \text{Beta}) / \text{Alpha} / \log((p_h * q_l) / (p_l * q_h)),$$

and

$$d_{ln} = -\log(q_l/q_h) / \log((p_h * q_l) / (p_l * q_h)) * n + \log(1 - \text{Alpha}) / \text{Beta} / \log((p_h * q_l) / (p_l * q_h)),$$

with $q_l=1-p_l$ and $q_h=1-p_h$. The Alpha and Betas have different meanings here than earlier, but we retain this notation because it reflects previous descriptions of these methods.⁹ A full discussion is available in the online appendix; note that the performance metrics for the truncated sequential sampling system are the same as for our static three-way classification and are comparable to our two-way classification approach (Table 1).

For the 2%/10% thresholds, the minimum sample size is 62, the maximum sample size is 161, and Alpha=0.0001 and Beta = 0.001. For the 5%/20% thresholds, the minimum sample size is 38, maximum sample size is 95, Alpha=0.0001 and Beta = 0.001. Operating characteristic curves and average sample sizes for a range of MDR TB prevalences for both sets of thresholds are provided in the eAppendix, including eTables 1 and 2 (<http://links.lww.com>).

Implementation of the static methods with early stopping—The usual static lot quality-assurance sampling system involves evaluating the entire sample and comparing to the appropriate decision rule. However, it is possible to apply early stopping with these systems, in which no additional samples are collected or tested if the target decision rule is reached. For example, for the two-way classification at the 2%/10% thresholds, if four resistant samples are observed in the first twenty tests, then we can stop early with a “high” classification. Conversely, sampling can also stop early when enough non-cases have been

observed that the decision rule will never be reached. For example, for the two-way classification at the 2%/10% thresholds, if we test 69 samples and none is resistant, then we can classify the site as low and avoid testing the last four because we will never classify the site as high. The results below apply these early stopping rules. Early stopping within the static lot quality-assurance sampling is conceptually distinct from the truncated sequential sampling approach that requires continuous evaluation against changing decision rules that depend on the number of samples tested up to that stage.

Results

We report the performance of lot quality-assurance sampling classification systems in each of the three settings, Ukraine, Tanzania, and Vietnam.

Ukraine

The original survey from the Donetsk oblast of Ukraine estimated 16% (13%–18%) of all new cases in the public sector and 22% (12%–31%) of all new cases in the penitentiary sector had MDR TB.⁶

At the low set of thresholds (2%/10%) none of the classification systems identified regional variability. All sites and groups of sites with enough samples for classification were classified as high (summaries provided in Table 2, with site-specific details provided in the eAppendix, <http://links.lww.com>). As expected for areas with high prevalence, early stopping (for the static designs) or truncated sequential sampling are able to achieve classification well before reaching the maximum sample size.

At the higher set of thresholds (5%/20%) we detected additional variability with some of the classification systems (summaries provided in Table 2, with site-specific details available in the eAppendix, <http://links.lww.com>). For the two-way static lot quality-assurance sampling, most sites and groups of sites with sufficient samples for classification were classified as high; one geographic area was classified as low. The three-way classification systems (static and truncated sequential sampling) detected more variability, with approximately half classified as moderate and the other half classified as high. Both the three-way static and the three-way truncated sequential sampling systems classified the same sites and groups of sites as moderate and high. The average numbers of samples required for classification in Ukraine was again substantially lower than the maximum allowed for each system.

Tanzania

The original survey from Tanzania estimated that 1.1% (0.6%–2.0%)⁸ of all new cases had MDR TB. In this setting where MDR TB is infrequently observed among new TB cases, the classifications of the lot quality-assurance sampling systems for both sets of thresholds were low (Table 2, with site-specific details in the eAppendix, <http://links.lww.com>). Thus, in this setting, the lot quality-assurance sampling systems using these thresholds did not detect any underlying variability in the proportion of MDR TB among new cases. For the static systems, the average numbers of samples required for classification were nearly the maximum, because there is little opportunity for early stopping in a low-prevalence setting. However, fewer samples were required for the three-way truncated sequential sampling classification for both sets of thresholds, because the system is designed to allow for early classification at very low or very high prevalence.

Vietnam

Like Tanzania, the original survey from Vietnam estimated a low fraction of MDR TB among new cases (1.8%, 1.0%–3.3%).⁷ Given this low overall prevalence, we might expect

uniform classification as low for all sites, as we found for Tanzania. However, we found (conditional on a sufficient number of samples) that the Tay Ninh Province was anomalous and was classified as high at both the 2%/10% and the 5%/20% thresholds. Upon further investigation, the original investigators noted that this site might be of particular concern because of its proximity to Cambodia, where MDR TB may be increasing. As in Tanzania, the actual number of samples needed to test to achieve classification was only slightly lower for the static systems. Truncated sequential approaches provided substantial efficiency gains at both sets of thresholds. (As shown in Table 2, with site-specific details available in the eAppendix, <http://links.lww.com>).

Discussion

We explored lot quality-assurance sampling for identifying geographic variation in MDR TB because of the reported value of this technique for surveillance of other diseases (including trachoma, HIV drug resistance, acute malnutrition, and schistosomiasis).^{9–14} In these other contexts, lot quality-assurance sampling has helped identify priority areas for intervention, or helped decide locally appropriate control strategies.

Lot quality-assurance sampling classification methods employed on a sub-regional level may hold some advantages over current survey approaches that produce area-wide estimates of resistance. From a regional policymaker's perspective, it is desirable not to only have some measure of drug resistance for the entire region, but also to be able to identify smaller areas that require further attention or prioritization for interventions. Lot quality-assurance sampling provides a natural framework in which a policymaker can rapidly identify these worrisome areas where extra resources could be employed. For example, if two separate sub-regional areas are classified with high levels of multiple drug resistance in new patients, follow-up studies can be directed at each region to uncover the most likely causal mechanisms based on local knowledge. A single, large undirected study for the entire region will have a lower likelihood of successfully identifying the causal mechanisms. Once local mechanisms are identified, appropriate interventions can be targeted.

We used existing data from the TB drug-resistance surveys in countries with varying levels of MDR TB prevalence to simulate the performance of lot quality-assurance sampling approaches for classifying MDR TB at the sub-regional level. Lot quality-assurance sampling was able to efficiently identify regional variation in the prevalence of MDR among new cases in Ukraine and Vietnam—but not in Tanzania, where prevalence was low. This suggests that lot quality-assurance sampling may allow for rapid identification of areas where transmitted MDR TB exceeds predetermined thresholds. One limitation of the current study is that, because we used previously collected data, many of the sites had insufficient data. These insufficient numbers interfere with a complete demonstration of the efficiency of lot quality-assurance sampling classification system and its ability to detect variability. Opportunities to stop sampling early, under both static and truncated sequential sampling designs, occur most often when MDR TB is at very high levels or very low levels. Detecting variability in moderate MDR TB settings requires sites to have more samples available. Another limitation is that, if population representativeness of cases were not retained in our regional aggregation of diagnostic sites for this re-analysis, our results could be biased. If adopted for prospective MDR TB surveys, the lot quality-assurance sampling system should be designed with sufficient sample sizes and representative sampling at the level to which the data will be grouped for site-level or geographic classification. A related limitation is that any lack of representativeness in the original surveys will be reflected in our re-analyses.

While the lot quality-assurance sampling design holds promise for identifying variation in MDR TB sites, there are several questions that must first be addressed. Foremost is determining appropriate thresholds. We explored two sets of thresholds – 2%/10% and 5%/20% - and applied these to all countries. This approach originated with the idea that a globally endorsed strategy might attempt to harmonize methods across all countries in order to simplify training and minimize the need for in-country technical capacity. A single set of thresholds could also be based on shared epidemiologic motivations. For example, if empiric or modeling evidence suggests that TB epidemics will rise rapidly if the proportion of MDR TB among new cases passes a certain threshold, then that threshold could be applied in lot quality-assurance sampling in all countries.

However, as our results show, a single threshold applied in all countries may not be optimal if the goal of the system is to detect local variability. The challenge in developing a country-specific system is that selection of these thresholds requires an understanding of the local prevalence of disease, technical capacity to develop the survey design, and the commitment of decision-makers to policy changes based on classifications. A priori estimates of drug resistance based on previous national surveys or other programmatic metrics may help policymakers establish reasonable thresholds, and could be incorporated in the design of the system using a Bayes—lot quality-assurance sampling approach.¹⁵

Another consideration is that a specific lot quality-assurance sampling design approach must be chosen: a design with three categories may help identify more variability between sites, and provide more flexibility for intervention if adequate sample sizes can be achieved. In addition to specifying the number of classification strata, a choice must be made regarding simple static, static with early stopping, or truncated sequential sampling. A complicated study with early stopping, while more efficient, may be more difficult to implement. This is particularly true if the classifications are based on samples collected from several sites. A truncated sequential stopping or static with early stopping design would require the ongoing aggregation and monitoring of results by a central coordinator. If one site lags because of lost or compromised samples, the results may be biased. These designs may also not be feasible when there are long delays between sample collection and availability of test results. Recent advances in PCR-based diagnostics (such as Xpert MTB/RIF), which can rapidly detect drug resistance and be placed closer to the point of patient care, greatly improve the appeal of truncated sequential sampling approaches because sequential sputum samples could actually be rapidly tested. Concerns about the possibility of low positive-predictive values for determining rifampin resistance in surveys with low prevalence of MDR (or rifampin monoresistance) will also need to be addressed.¹⁷

If feasible, early stopping can substantially reduce the number of samples needed for collection and testing, saving time and money. If an early stopping three-way system is adopted, then the truncated sequential-sampling design is recommended because it requires fewer samples in both the high and low prevalence setting. Because of the efficiency gains with three-way truncated sequential sampling lot quality-assurance sampling, this approach deserves further attention as rapid-testing technology improves. Future research should also consider truncated sequential sampling designs with convergent or divergent thresholds, as these may provide additional efficiency gains.¹⁸

In summary, the lot quality-assurance sampling design is able to classify the level of MDR TB at the sub-national level, and to detect local variation in resistance in regions with different resistance levels. Ongoing MDR TB studies can apply these techniques to previously collected data as exemplified here. However, this activity would be greatly improved if thresholds and corresponding decisions are specified a priori, and if sites or clusters of sites have sufficient sample sizes to be included in the analysis. The selection of

diagnostic sites to include in the survey is also important.¹⁹ Ideally, all sites would be included, which would allow mapping across a wide-area to understand geographic variability. In this case, the data from each site could be treated as a stratified sample and aggregated to produce regional and national MDR TB estimates. If not all sites can be included, then sites should be randomly selected in order to classify individual sites and also retain the ability to aggregate the data for regional and national estimates MDR TB prevalence.²⁰ To achieve unbiased aggregated estimates, the results must be appropriately weighted to reflect the chosen method for selecting sites for the survey (simple random sampling versus probability proportionate to size). Prospective studies in which lot quality-assurance sampling is used to classify resistance among new TB cases will be needed in order to see how well this approach fulfills its promise for improving our understanding of the distribution and burden of drug resistance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge the efforts of the survey teams in Ukraine, Vietnam and Tanzania involved in the initial data collection.

Sources of Financial Support:

BLH received support from National Institutes of Health Grant R56 EB006195. FvL received funding from the UK Department for International Development (DFID) for his work in the Tanzania survey. TC received funding from National Institute of Health grants DP2OD006663 and U54GM088558.

Financial support for the survey in Donetsk, Ukraine, was provided by the United States Agency for International Development for the survey in Vietnam, from The Netherlands Government Grant for National Strategic Plan of the NTP for the period 2001–2005 and for the survey in Tanzania, from the World Health Organization.

References

1. World Health Organization. Geneva: World Health Organization; 2010. Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 global report on surveillance and response. (WHO/HTM/TB/2010.3)
2. World Health Organization. Guidelines for surveillance of drug resistance in tuberculosis – 4th ed. (WHO/HTM/TB/2009.422)
3. World Health Organization. Geneva: World Health Organization; 2008. Anti-tuberculosis drug resistance in the world, report number 4. (WHO/HTM/TB/2008.394)
4. Dodge, HF.; Romig, HG. Single and Double Sampling. 2nd ed.. New York: John Wiley; 1959. Sampling Inspection Tables.
5. Bennett DE, Bertagnolio S, Sutherland D, Gilks CF. The World Health Organization's global strategy for prevention and assessment of HIV Drug resistance. *Antivir Ther.* 2008; 13 Suppl 2:1–13. [PubMed: 18578063]
6. Dubrovina I, Miskinis K, Lyepshina S, et al. Drug-resistant tuberculosis and HIV in Ukraine: a threatening convergence of two epidemics? *Int J Tuberc Lung Dis.* 2008; 12(7):756–762. [PubMed: 18544200]
7. Huong NT, Lan NT, Cobelens FG, et al. Antituberculosis drug resistance in the south of Vietnam: prevalence and trends. *J Infect Dis.* 2006; 194(9):1226–1232. [PubMed: 17041848]
8. Chonde TM, Doulla B, van Leth F, et al. Implementation of a national anti-tuberculosis drug resistance survey in Tanzania. *BMC Public Health.* 2008; 8:427. [PubMed: 19116022]
9. Myatt M, Bennett DE. A novel sequential sampling technique for the surveillance of transmitted HIV drug resistance by cross-sectional survey for use in low resource settings. *Antivir Ther.* 2008; 13 Suppl 2:37–48. [PubMed: 18575190]

10. Myatt, M.; Limburg, H.; Minassian, D.; Katyola, D. Field trial of applicability of lot quality assurance sampling survey method for rapid assessment of prevalence of active trachoma; Bull World Health Organ. 2003. p. 877-885.
11. Myatt, M.; Mai, HP.; Quynh, NQ., et al. Using lot quality-assurance sampling and area sampling to identify priority areas for trachoma control: Viet Nam; Bull World Health Organ. 2005. p. 756-763.
12. Deitchler M, Valadez JJ, Egge K, Fernandez S, Hennigan M. A field test of three LQAS designs to assess the prevalence of acute malnutrition. Int J Epidemiol. 2007; 36(4):858–864. [PubMed: 17517808]
13. Brooker S, Kabatereine NB, Myatt M, Stothard JR, Fenwick A. Rapid assessment of *Schistosoma mansoni*: the validity, applicability, and cost effectiveness of the Lot Quality Assurance Sampling method in Uganda. Trop Med Int Health. 2005; 10(7):647–658. [PubMed: 15960703]
14. Rabarijaona LP, Boisier P, Ravaoalimalala VE, et al. Lot quality assurance sampling for screening communities hyperendemic for *Schistosoma mansoni*. Trop Med Int Health. 2003; 8(4):322–328. [PubMed: 12667151]
15. Olives C, Pagano M. Bayes-LQAS: classifying the prevalence of global acute malnutrition. Emerg Themes Epidemiol. 2010; 7:3. [PubMed: 20534159]
16. Boehme, et al. Rapid Molecular Detection of Tuberculosis and Rifampin Resistance. N Engl J Med. 2010; 363(11):1005–1015. [PubMed: 20825313]
17. World Health Organization. Rapid implementation of the Xpert MTB/RIF diagnostic test: technical and operational “How-to”; practical considerations. 2011. WHO/HTM/TB/2011.2
18. Armitage S. Restricted Sequential Procedures. Biometrika. 1957; 44(1):9–26.
19. Hedt BL, Laufer M, Cohen T. Drug resistance surveillance in resource-poor settings: current methods and considerations for TB, HIV and malaria. Am J Trop Med Hyg. 2011; 84(2):192–199. [PubMed: 21292884]
20. Biedron C, Pagano M, Hedt BL, Kilian A, Ratcliffe A, Mabunda S, Valadez JJ. An assessment of Lot Quality Assurance Sampling to evaluate malaria outcomes indicators: extending malaria indicators surveys. Int J Epidemiol. 2010; 39(1):72–79. [PubMed: 20139435]

Table 1
Parameters for misclassification errors for calculating sample sizes and determining decision rules for three LQAS designs

	Two-way classification			Three-way classification			Truncated Sequential Sampling for Three-way classification		
	Constraints ^a	Actual ^b (2%/10%)	Actual ^b (5%/20%)	Constraints ^a	Actual ^b (2%/10%)	Actual ^b (5%/20%)	Constraints ^a	Actual ^b (2%/10%)	Actual ^b (5%/20%)
Properties of the system									
P_{Hh} -probability of classifying as high at low prevalence	0.1	0.066	0.067	0.1	<0.001	<0.001	0.1	<0.001	<0.001
P_{Hl} -probability of classifying as low at high prevalence	0.05	0.047	0.044	0.05	<0.001	<0.001	0.05	<0.001	0.001
P_{Ll} -probability of classifying as low at low prevalence	$= 1 - P_{Hh}$	0.934	0.933	0.80	0.875	0.809	0.8	0.806	0.824
P_{min} -probability of classifying as moderate at middle prevalence	N/A	N/A	N/A	0.8	0.804	0.835	0.8	0.804	0.807
P_{Hh} -probability of classifying as high at high prevalence	$= 1 - P_{Hl}$	0.953	0.956	0.80	0.806	0.801	0.8	0.802	0.803
Description of the lot quality-assurance system									
minimum sample size ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A	62	38
maximum sample size	76	44	44	169	94	94	161	161	95
lower decision rule	N/A	N/A	N/A	6	7	7	<i>d</i>	<i>d</i>	<i>d</i>
upper decision rule	4	5	5	14	16	16	<i>d</i>	<i>d</i>	<i>d</i>

^a Constraints that led to specific LQAS system;
^b Actual properties of the selected LQAS system;
^c It is possible to achieve classification before reaching the maximum sample size for the static two-way and three-way classification systems are used if early stopping is allowed.
^d Determined by the lines described in paper with Alpha=0.0001 and Beta=0.001.

Table 2
Lot quality-assurance sampling classifications for the 2%/10% and 5%/20% thresholds

Country	Total Number	Two-way classification					2%/10% Thresholds					Three-way Truncated Sequential Sampling classification				
		Number classified*	Low	High	Average number of samples (max=76)	Number classified*	Low	Moderate	High	Average number of samples (max=169)	Number classified*	Low	Moderate	High	Average number of samples (max=161)	
Tanzania (sites)	24	4	4	0	73.8	0	0	0	NA	1	1	0	0	82.0		
Tanzania (groups)	5	5	5	0	73.8	3	0	0	166.0	4	4	0	0	106.5		
Ukraine (sites)	53	7	0	7	19.1	3	0	3	69.3	5	0	0	5	69.2		
Ukraine (groups)	5	5	0	5	30.4	5	0	0	88.0	5	0	0	5	70.0		
Vietnam (sites)	20	4	3	1	62.5	1	1	0	168.0	1	1	0	0	121.0		
Vietnam (groups)	5	4	3	1	69.8	2	2	0	167.0	3	3	0	0	108.0		
5%/20% Thresholds																
Tanzania (sites)	24	10	10	0	40.4	0	0	0	NA	9	9	0	0	47.0		
Tanzania (groups)	5	5	5	0	40.4	4	4	0	89.0	5	5	0	0	46.4		
Ukraine (sites)	53	7	0	7	24.9	3	0	1	70.3	3	0	1	2	70.7		
Ukraine (groups)	5	5	1	4	31.6	5	0	3	81.6	5	0	3	2	73.2		
Vietnam (sites)	20	9	8	1	39.6	3	2	1	91.0	7	7	0	0	50.0		
Vietnam (groups)	5	5	5	0	41.0	4	4	0	90.5	5	5	0	0	55.4		

* Only sites or groups of sites that could be classified before exhausting the samples available were classified and included in this table.