JOURNAL OF CLINICAL MICROBIOLOGY, Mar. 2008, p. 897–901 Vol. 46, No. 3
0095-1137/08/$08.00+0 doi:10.1128/JCM.01553-07

Sputum Smear Microscopy: Evaluation of Impact of Training, Microscope Distribution, and Use of External Quality Assessment Guidelines for Resource-Poor Settings[▽]

A. Van Rie,[1]* D. Fitzgerald,[2] G. Kabuya,[3] A. Van Deun,[4] M. Tabala,[5] N. Jarret,[1] F. Behets,[1] and E. Bahati[6]

Department of Epidemiology, the University of North Carolina at Chapel Hill, Chapel Hill, North Carolina[1]; Division of Infectious Diseases, the University of North Carolina at Chapel Hill, Chapel Hill, North Carolina[2]; National Reference Laboratory, Tuberculosis Control Program, Kinshasa, Democratic Republic of Congo[3]; Mycobacteriology Unit, Institute of Tropical Medicine, Antwerp, Belgium[4]; Ecole de Sante Publique, University of Kinshasa, Kinshasa, Democratic Republic of Congo[5]; and National Tuberculosis Program, Kinshasa, Democratic Republic of Congo[6]

Received 3 August 2007/Returned for modification 21 October 2007/Accepted 23 December 2007

Sputum smear microscopy is the main and often only laboratory technique used for the diagnosis of tuberculosis in resource-poor countries, making quality assurance (QA) of smear microscopy an important activity. We evaluated the effects of a 5-day refresher training course for laboratory technicians and the distribution of new microscopes on the quality of smear microscopy in 13 primary health care laboratories in Kinshasa, Democratic Republic of Congo. The 2002 external QA guidelines for acid-fast bacillus smear microscopy were implemented, and blinded rechecking of the slides was performed before and 9 months after the training course and microscope distribution. We observed that the on-site checklist was highly time-consuming but could be tailored to capture frequent problems. Random blinded rechecking by the lot QA system method decreased the number of slides to be reviewed. Most laboratories needed further investigation for possible unacceptable performance, even according to the least-stringent interpretation. We conclude that the 2002 external QA guidelines are feasible for implementation in resource-poor settings, that the efficiency of external QA can be increased by selecting sample size parameters and interpretation criteria that take into account the local working conditions, and that greater attention should be paid to the provision of timely feedback and correction of the causes of substandard performance at poorly performing laboratories.

Tuberculosis (TB) is one of the world's leading causes of infectious disease-related morbidity and mortality. The World Health Organization (WHO) estimated that there were 8.9 million new cases of TB in 2004, of which 3.9 million were sputum smear positive (10). Each individual with untreated smear-positive TB infects 10 to 15 persons per year, making the identification of these infectious patients one of the key aspects of TB control (11).

Case detection through quality-assured bacteriology is an essential element of the WHO STOP TB strategy (8). Because of a limited culture capacity, many resource-poor countries rely solely upon sputum smear microscopy for the diagnosis of TB. The quality of smear microscopy depends on a network of local laboratories and external quality assessment (EQA) of these laboratories under the supervision of the national reference laboratory (NRL) (9). EQA of smear microscopy in resource-poor settings most often consists of on-site unblinded review by a laboratory supervisor of positive slides and 10% of negative slides. This method, which has not been validated in the field, is labor-intensive and is often a neglected part of national TB programs in resource-poor countries (7).

In an effort to simplify and standardize EQA activities and to prioritize EQA at national TB control programs (NTPs), a practical EQA guideline was developed by an international working group and endorsed in 2002 (2). These international EQA guidelines recommend the use of three methods for the evaluation of laboratory performance: on-site assessment by the use of a standardized questionnaire, panel testing of technician proficiency by the use of centrally prepared slides, and blinded rechecking of a random sample of routine slides from each peripheral laboratory. To decrease the EQA workload, the blinded rechecking recommendations incorporate statistical sampling methods developed for industrial quality control. To date, few countries have implemented the 2002 EQA guidelines and few studies have evaluated the routine implementation of these new guidelines in resource-poor settings (3–6).

We assessed the impacts of an intervention consisting of a 5-day refresher training course for laboratory technicians and the distribution of new microscopes on the performance of smear microscopy. We assessed the feasibility of the implementation of two components of the new EQA guidelines, blinded rechecking and on-site evaluation, in 13 laboratories in Kinshasa, Democratic Republic of Congo (DRC). Panel testing was not evaluated, as it has limited value in assessing

* Corresponding author. Mailing address: Department of Epidemiology, 2104 McGavran Hall, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7435. Phone: (919) 966-1420. Fax: (919) 966-2089. E-mail: vanrie@email.unc.edu.

▽ Published ahead of print on 3 January 2008.

TABLE 1. IUATLD grading scale for AFB found by Ziehl-Neelsen smear microscopy

Finding	Recording
No AFB in at least 100 fields	Negative
1 to 9 AFB in 100 fields	Exact no./100 fields
10 to 99 AFB in 100 fields	1+
1 to 10 AFB per field in at least 50 fields	2+
>10 AFB per field in at least 20 fields	3+

routine performance and might best be reserved for testing at the end of training sessions (7).

MATERIALS AND METHODS

The study was conducted as part of a program of technical assistance to DRC. DRC is ranked 11th among the TB high-burden countries (10) and has 1,083 sputum microscopy centers, 23 regional laboratories, and one NRL in the capital, Kinshasa. Only the NRL has culture facilities.

Several steps were implemented as part of the study. First, selected sites were visited for on-site evaluation and the collection of slides for blinded rechecking before the intervention. Second, all laboratory staff performing smear microscopy at the selected sites underwent a 5-day refresher training course at NRL, and new microscopes were distributed to each site. Third, 9 months later, slides were again selected for blinded rechecking.

Site selection. Thirteen diagnostic TB centers were selected from among the 89 diagnostic TB centers in Kinshasa, based on the smear microscopy workload in 2003. The 13 laboratories selected performed 45% of all smears for acid-fast bacilli (AFB) in Kinshasa. The median annual workload at the 13 selected sites was 4,371 sputum smears (range, 1,283 to 17,837 sputum smears) for AFB. Carbol fuchsin staining solutions for the entire city were centrally prepared and quality controlled before distribution to the health care centers. Smear microscopy was performed according to DRC national guidelines. The selected sites agreed not to rotate laboratory technicians during the study period.

On-site evaluation. On-site evaluation of each laboratory was performed by using the comprehensive on-site evaluation checklist proposed by the 2002 EQA guidelines and adapted to reflect the DRC NTP guidelines. The evaluation was completed by the same NRL and research staff for all sites and consisted of both open- and closed-ended questions, simple observations, and limited on-site rechecking of slides.

Intervention. All laboratory technicians employed at the study sites underwent a 1-day theoretical and 4-day practical refresher training course in smear microscopy. The participants were administered a standardized pre- and posttraining test addressing theoretical and practical knowledge related to TB and AFB sputum microscopy. Each center received one or two new microscopes following the training, and the technicians were instructed on their use and care.

Blinded slide rechecking. The blinded slide rechecking process consisted of sample size calculation, slide storage, slide collection, blinded slide rechecking at the NRL before and after restaining, rechecking of slides with discrepant results at a supranational reference laboratory (SRL), and the classification of errors.

In compliance with the 2002 EQA guidelines, the smallest possible sample size that allows solid conclusions to be made about the performance of a laboratory was calculated by using the lot QA system (LQAS) method. The 2003 positivity rate and the total number of slides processed at all laboratories in Kinshasa were provided by the NTP. The maximum number of false-negative errors allowed in a sample was set equal to 0 (acceptance number of 0). The expected performance (sensitivity) of the peripheral laboratory technicians compared to the performance of the controllers was set equal to 80%. The specificity of the peripheral laboratory technicians compared to that of the controllers was set equal to 100% (no false-positive results were tolerated). On the basis of these parameters, the annual sample size needed for the blinded rechecking was 57 slides per site.

Each peripheral laboratory was asked to store the 500 most recently collected slides and have them correctly labeled and cleaned with xylene. The 57 slides for blinded rechecking were selected from among these 500 stored samples. To ensure a random, unbiased, and representative sample of slides from each site for blinded review, sample slides were identified from the register by the research staff and collected by the laboratory technician. If a slide was missing, the next slide identified in the register was selected. The slides selected thus represent slides with negative, positive, and scanty positive results in proportion to their occurrence in the laboratory register.

TABLE 2. Classification of rechecking errors encountered upon blinded rechecking of smear microscopy results as part of EQA of peripheral laboratories

Result by peripheral laboratory	Error classification according to the following result of the controller:				
	Negative	1 to 9 AFB	1+	2+	3+
Negative	Correct	LFN	HFN	HFN	HFN
1 to 9 AFB	LFP	Correct	Correct	QE	QE
1+	HFP	Correct	Correct	correct	QE
2+	HFP	QE	Correct	Correct	Correct
3+	HFP	QE	QE	Correct	Correct

Randomly selected slides were reviewed in a blinded fashion at the NRL by using the same standard technique employed at the peripheral laboratory. The results were reported according to the standard IUATLD grading scale (Table 1). Following an initial review at the NRL, all slides were also rechecked after they were restained. Discrepant results were resolved at the Institute of Tropical Medicine, Antwerp, Belgium, an SRL. The technician at the SRL was informed of both results but was blinded to which result was from the peripheral laboratory and which one was from the NRL. The decision given at the SRL was considered final.

All errors were defined as a quantification error (QE), a low-false-negative (LFN) result, a high-false-negative (HFN) result, a low-false-positive (LFP) result, or a high-false-positive (HFP) result according to the international EQA classification (Table 2). EQA results were interpreted by using the most stringent criteria listed in the guidelines, suggesting that any major error (an HFP or HFN result) is unacceptable performance, as well as the least-stringent criteria, suggesting that any HFP result, more than three LFN results, and one or two HFN results define unacceptable performance.

RESULTS

Evaluation of laboratory technicians' knowledge and skills before and after refresher training. The participants had a good understanding of the basic theoretical aspects of TB and sputum microscopy prior to participation in the refresher training course (Table 3). Training resulted in a marked improvement in the practical component (smear preparation, staining, and reading), with median test scores increasing from 70% pretraining to 86% posttraining (chi-square test, $P < 0.01$).

On-site evaluation. The most common problems encountered during the on-site evaluations were shortages of materials (such as distilled water, lens tissue, and disinfectant) and the unavailability or the poor condition of the necessary equipment (including wire loops, staining racks, a biohazard waste bin, and a microscope) (Table 4). Other common problems involved poor microscope care, improper smear preparation,

TABLE 3. Ziehl-Neelsen smear microscopy results pre- and posttraining in 13 public health laboratories in Kinshasa

Training	Score (%)	
	Pretesting	Posttesting
Theoretical	89 (50–100)[a]	92 (72–100)
Practical		
Smear prepn	71 (43–95)	90 (57–100)
Smear staining	73 (30–100)	83 (67–100)
Smear reading	66 (17–83)	100 (75–100)
Practical total	70 (44–92)	86 (73–98)

[a] Values in parentheses are the median (minimum-maximum).

TABLE 4. Most common problems reported or observed during on-site evaluation visits of 13 public health laboratories in Kinshasa

Problem	% of sites
Work space	
No separate area for TB work	57
Electricity unavailable, weak, or intermittent	36
Equipment and reagents	
Wire loops	
Insufficient no.	86
Poor condition	21
Staining racks	
Poor condition	86
Insufficient no.	79
None available	36
No lens tissue available	86
Microscope	
Insufficient no.	50
Insufficient/inadequate light source	50
Objective not cleaned after every slide	79
No routine care	43
Smear prepn	
Incorrect size	86
Incorrect thickness	50
Smears prepared from saliva	43
Smear staining	
Stain allowed to dry on slide	71
Failure to stain with hot carbol fuchsin	50
Destained for less than 3 min	36
Smear reading	
Procedure not followed	57
Results inferior to NTP standards	50
Administrative	
Results not directly recorded on form	50
Elevated workload	43
Forms incomplete	36
Improper report forms used	29
Sanitation and safety	
Insufficient supply of disinfectant	57
No biohazard waste bin	50
Improper use of gloves	43
Disinfectant not used	29
Work areas not washed weekly	29
Supervision	
No NTP feedback	86
Slides not stored	21
Errors on rechecking	14

poor staining or reading techniques, incorrect data recording and slide storage, and a lack of feedback from the NTP.

Participant feedback. In general, the questionnaire administrators believed that the tool was clear, accurately reflected the laboratory conditions, and was important in motivating laboratory workers in their daily work. Some laboratory technicians and the NRL supervisor reported that the on-site evaluation process was too lengthy and time-consuming (1 to 2 h) for routine supervision. Some technicians experienced the new requirements for the storage of slides as being too cumbersome and complex.

Blinded slide rechecking. Among the 741 slides collected from the peripheral laboratories and reviewed by the NRL, there were 77 (10.4%) discrepant results. According to the SRL results, 67 (87%) of these discrepant results were attributed to the peripheral laboratory and 10 (13%) were attributed to the NRL.

Prior to the intervention, major errors (i.e., HFP and HFN errors) were observed in 8 of the 13 (61.5%) laboratories, with HFN results being much more frequent than HFP results (Table 5). Minor errors (i.e., LFP and LFN errors and QEs) were observed in the majority (10 or 77%) of peripheral laboratories. After the intervention, the occurrence of major errors remained high, with one or more major errors being observed in 10 (77%) clinics. HFN results were still more likely to occur than HFP results. The number of major errors decreased at five sites, increased at six sites, and remained unchanged at the remaining two sites. The proportion of laboratories where minor errors were detected also remained unchanged from the preintervention to the postintervention assessment. LFN results also remained more common than LFP results.

Overall, 26% of the discrepancies were detected upon restaining of the slides at the NRL. Most (80%) of the discrepancies detected by restaining were minor errors.

DISCUSSION

This study evaluated the effect of a 5-day refresher training course and the distribution of new microscopes on the quality of smear microscopy at primary health care center laboratories through the implementation of the 2002 EQA guidelines for TB laboratories in resource-poor settings. The on-site evaluation and blinded rechecking of the slides proved to be feasible and complementary aspects of the EQA process, but we could not demonstrate a long-term effect of the training and microscope distribution.

Similar to the findings of studies performed in India, the Philippines, and Mexico, we found blinded rechecking and the LQAS strategy to be operationally feasible (3–6). The blinded rechecking identified those laboratories with a high number of smear microscopy errors and allowed the distinction between those laboratories with optimal and suboptimal performance. Similar to the conclusions of a study performed in Uganda, we conclude that a standardized on-site assessment is a useful tool in quality control and that meaningful changes can be successfully implemented on the basis of the results (1). The on-site evaluation identified potential problems underlying the substandard performance at deficient laboratories and helped with the formulation of corrective actions. On-site evaluation was also a key in motivating laboratory technicians. Implementation of the new EQA method resulted in an important decrease in the number of slides that had to be rechecked, but the use of the comprehensive standardized checklist was time-consuming (1 to 2 h) and many of the deficiencies identified were common, with multiple findings suggesting problems in the same areas. These limitations of the on-site evaluation may be overcome by use of the short checklist included in the 2002 EQA guidelines as a template, and tailoring the questionnaire to the specific problems that a country or a region has identified at sentinel sites may further improve the efficiency of the on-site evaluation.

TABLE 5. Results of slide rechecking at 13 primary health laboratories in Kinshasa

Center	No. of laboratories with the indicated result:									
	Preintervention					Postintervention				
	HFN	HFP	LFN	LFP	QE	HFN	HFP	LFN	LFP	QE
A	3	1	3					2		
B	1							1		1
C			1			1	1	1		
D	1				1			2		
E				1		2				
F	1						1	1		
G		1		1		2	2	1		
H				1		1				1
I		2				1		4		
J				1		1				
K	5		1	1		1			1	
L			3			1				
M	1		1			1		2		

The greatest challenges encountered were related to the interpretation of the results, the provision of timely feedback, and the formulation of corrective actions. The 2002 EQA guidelines state, "When establishing a rechecking program, it will be important for the NTP to establish standards for acceptable performance, as well as recommended investigation steps and appropriate actions to correct problems" (2). Unfortunately, the directions given in the guidelines on interpretation can lead to confusion. The 2002 EQA guidelines suggest three different interpretations. All interpretations have no errors of any type as the target for optimal performance, and the guidelines recommend that any major error and frequent minor errors trigger an evaluation and the taking of corrective action, if it is needed. According to the most stringent interpretation, any major error may indicate unacceptable performance, while according to the least-stringent interpretation, any HFP result or one or two HFN results may indicate unacceptable performance. While HFP errors should not occur (the specificity is set at 100% in the sample size calculation), an isolated HFP result can be due to a clerical error. More than one HFP result suggests serious microscope malfunction, grossly inadequate technique, or an inability to recognize AFB due to inadequate training or high staff turnover. An HFN result may suggest work overload, poor staining reagents or technique, inadequate microscopes, or incorrect reading. Frequent LFN results should also be addressed for these deficiencies but are most often due to careless work.

In our study, most laboratories had one or more major errors, predominantly false-negative errors. The numbers of laboratories with unacceptable performance preintervention were eight (61.5%) when the example in the guidelines with the most stringent criteria was applied and four (31%) when the least-stringent criteria were applied. Until the working conditions in underresourced public health systems are corrected, the use of the least-stringent criteria may be more efficient, as a higher proportion of laboratories with a problem in need of technical correction would be focused on. Under such conditions, one may even consider reduction of the set point of sensitivity from 80% to 70%. This would allow EQA staff to focus on the provision of timely feedback, problem solving, and supportive supervision at those laboratories with the poorest performance.

Even though the 5-day training course resulted in a significant improvement in the scores on the standardized test, refresher training and the provision of new microscopes did not lead to an observable long-term (9-month) improvement in the quality of smear microscopy. This may be because certain factors were not addressed by the intervention, such as a high workload, poor working conditions, and poor staff motivation; because the sample size in the LQAS method is too small to detect relatively small improvements in performance; or because the training did not have a long-lasting effect.

The logistical complexity of involving both an NRL and an SRL resulted in important delays in the provision of feedback of the results to the peripheral laboratories. This experience suggests that the EQA process will be effective and allow the provision of rapid feedback only if it is completed locally, with final controls at the national or even the regional level and continuous evaluation of the controllers. Timely completion and the provision of feedback with probing for problems are probably far more important than the type of laboratory giving the final verdict. Without proper feedback, other efforts are likely to be in vain. On the basis of the results of this study, the DRC NTP has revised its EQA guidelines and has introduced LQAS method sampling via training sessions, with more emphasis on correct technical execution and follow-up.

In conclusion, this study supports the implementation of the 2002 EQA guidelines in resource-poor settings. The use of the LQAS technique leads to a smaller sample size, thus reducing the EQA workload; and the combination of on-site evaluation and blinding of rechecking allows an unbiased and representative evaluation of the quality of sputum microscopy, the identification of underlying problems, the formulation of feedback and corrective actions, and the motivation of laboratory technicians. Prior to implementation of the EQA system, clear interpretation guidelines must be agreed upon by the NTP; and plans must be in place to provide timely feedback, improve working conditions at suboptimally performing laboratories, and correct underlying problems identified during on-site evaluations.

ACKNOWLEDGMENTS

Funding was provided by the Centers for Disease Control and Prevention, Atlanta, GA (UNC-GAP project CDC U62 CCU422).

We acknowledge the work of the laboratory technicians at the participating laboratories and the NRL in Kinshasa. We acknowledge Mourad Gumusboga at the Institute of Tropical Medicine, Antwerp, Belgium, for his work on rereading discrepant slides.

REFERENCES

1. **Aziz, M., and G. Bretzel.** 2002. Use of a standardized checklist to assess peripheral sputum smear microscopy laboratories for tuberculosis diagnosis in Uganda. Int. J. Tuberc. Lung Dis. **6:**340–349.
2. **Aziz, M. A.** 2002. External quality assessment for AFB smear microscopy. Association of Public Health Laboratories, Washington, DC.
3. **Endo, S., M. Trono, A. Fujiki, and N. Macalalad.** 2007. Operational conditions influencing the proficiency of AFB microscopy services in the Philippines. Int. J. Tuberc. Lung Dis. **11:**293–298.
4. **Martinez, A., S. Balnadrano, A. Parissi, A. Zuniga, M. Sanchez, J. Ridderhof, H. B. Lipman, and B. Madison.** 2005. Evaluation of new external quality assessment guidelines involving random blinded rechecking of acid-fast bacilli smears in a pilot project in Mexico. Int. J. Tuberc. Lung Dis. **9:**301–305.
5. **Selvakumar, N., B. N. Murthy, E. Prabhakaran, S. Sivagamasundari, S. Vasanthan, M. Perumal, R. Govindaraju, L. S. Chauhan, F. Wares, T. Santha, and P. R. Narayanan.** 2005. Lot quality assurance sampling of sputum acid-fast bacillus smears for assessing sputum smear microscopy centers. J. Clin. Microbiol. **43:**913–915.
6. **Selvakumar, N., E. Prabhakaran, B. N. Murthy, S. Sivagamasundari, S. Vasanthan, R. Govindaraju, M. Perumal, L. S. Chauhan, T. Santha, and P. R. Narayanan.** 2005. Application of lot sampling of sputum AFB smears for the assessment of microscopy centers. Int. J. Tuberc. Lung Dis. **9:**306–309.
7. **Van Deun, A., and F. Portaels.** 1998. Limitations and requirements for quality control of sputum smear microscopy for acid-fast bacilli. Int. J. Tuberc. Lung Dis. **2:**756–765.
8. **World Health Organization.** 2006. Stop TB strategy: building on and enhancing DOTS to meet the TB-related Millennium Development Goals. World Health Organization, Geneva, Switzerland.
9. **World Health Organization.** 2006. Strategic approach for the strengthening of laboratory services for tuberculosis control, 2006–2009. World Health Organization, Geneva, Switzerland.
10. **World Health Organization.** 2006. Global tuberculosis control: surveillance, planning, financing. WHO report. World Health Organization, Geneva, Switzerland.
11. **World Health Organization.** March 2006. WHO Fact Sheet—tuberculosis. http://www.who.int/mediacentre/factsheets/fs104/en/index.html (verified June 2007). World Health Organization, Geneva, Switzerland.