INT J TUBERC LUNG DIS 1(3):231–238

Quality assurance programme for drug susceptibility testing of *Mycobacterium tuberculosis* in the WHO/IUATLD Supranational Laboratory Network: first round of proficiency testing

A. Laszlo,* M. Rahman,* M. Raviglione,† F. Bustreo,† WHO/IUATLD Network of Supranational Reference Laboratories§

***Laboratory Centre for Disease Control, Health Canada, Ottawa, Canada, †World Health Organization Tuberculosis Programme, Geneva, Switzerland**

SUMMARY

SETTING: Quality assurance of the WHO/IUATLD global tuberculosis drug resistance surveillance programme.
OBJECTIVE: To perform a proficiency test of drug susceptibility procedures within the WHO/IUATLD network of supranational reference laboratories (SRL).
DESIGN: Identical culture panels consisting of 20 clinical isolates of *Mycobacterium tuberculosis* containing both drug susceptible and drug resistant cultures were tested by the 16 laboratories of the network for resistance to streptomycin, isoniazid, rifampicin and ethambutol. The drug susceptibility testing procedures included the proportion, absolute concentration and resistance ratio methods as well as their variants, including the radiometric BACTEC 460 method.
RESULTS: The first round of proficiency testing has shown that the specificity of drug susceptibility testing within the SRL network was significantly higher than its sensitivity. The testing of isoniazid and rifampicin shows a high degree of agreement between the labs, but discordant results can be obtained with streptomycin and ethambutol.
CONCLUSION: Drug susceptibility procedures for the testing of isoniazid and rifampicin, the two anti tuberculosis drugs which define multidrug-resistant tuberculosis, are highly reliable within the SRL network. Procedures for drug susceptibility testing of streptomycin and ethambutol are still in need of standardization.
KEY WORDS: quality assurance; drug susceptibility testing; *M. tuberculosis*; WHO/IUATLD; Supranational Laboratory Network

THE EMERGENCE of *Mycobacterium tuberculosis* resistant to one or more antimicrobial agents has recently received attention, largely due to the outbreaks of multidrug-resistant tuberculosis (MDRTB) in HIV-infected patients in the United States and in Europe.[1–3] The real extent of the problem is, however, difficult to assess due to the lack of reliable bacteriological and epidemiological data. Anecdotal findings suggest that drug resistance could be high in Asia and Africa,[4] where it may become an important barrier to effective tuberculosis control. A World Health Organization (WHO)-sponsored global survey on drug resistance undertaken in the late 1980s in Latin America and the Caribbean[5] showed that although drug resistant tuberculosis had not reached alarming levels, one out of six cases of tuberculosis presented some form of drug resistance. This survey also showed that pockets of high drug resistance occurred throughout the region and were cause for concern, especially in the case of rifampicin (RIF).

Much of the world's literature on tuberculosis drug resistance surveillance lacks uniformity in the presentation of results, uses non-standardized drug susceptibility testing methods, shows sampling bias, and fails to differentiate properly between primary, initial and acquired drug resistance. Other limitations, such as the lack of *M. tuberculosis* culture facilities, deficient anti-mycobacterial drug susceptibility testing in many countries and the absence of longitudinal databases,

§C. Abe, M. Aoki, Research Institute of Tuberculosis, Tokyo, Japan; D. Yala, F. Boulahbal, Institut Pasteur d'Algérie, Algiers, Algeria; K. Feldmann, Kuratorium Tuberkulose in der Welt e.V., Gauting, Germany; P. A. Jenkins, PHLS Mycobacterium Reference Unit, Cardiff, UK; S. Hoffner, The Swedish Institute for Infectious Disease Control, Stockholm, Sweden; I. N. de Kantor, INPPAZ, Buenos Aires, Argentina; S. J. Kim, Korean Institute of Tuberculosis, Seoul, Republic of Korea; A. Laszlo (Co-ordinator), Laboratory Centre for Disease Control, Ottawa, Canada; N. Martin Casabona, CSU 'Vall d'Hebron', Barcelona, Spain; M. F. M. P. Pereira, Instituto Nacional de Saude, Porto, Portugal; F. Portaels, Institut voor Tropische Geneeskunde, Antwerp, Belgium; S. Rüsch-Gerdes, Borstel Institute, Borstel, Germany; V. Sticht-Groh, Armauer Hansen Institut DAHW, Würtzburg, Germany; R. Urbanczik, Fatol GmbH., Schiffweiler, Germany; V. Vincent, Institut Pasteur, Paris, France; K. Weyer, MRC National Tuberculosis Research Programme, Pretoria, South Africa.

Correspondence to: Dr A Laszlo, Laboratory Centre for Disease Control, National Reference Centre for Tuberculosis, EHC building No. 8, Room 230, Tunney's Pasture 08020, Ottawa ON K1A 0L2 Canada.

have rendered the information sometimes uninterpretable, thus preventing international comparisons.

Since clinical drug resistance, of whatever origin, is a man-made phenomenon due to inadequacies of treatment, it may be a useful indicator for monitoring the efficiency of tuberculosis control programmes. Thus, in 1994, the WHO, concerned about this problem, developed a set of guidelines jointly with the International Union Against Tuberculosis and Lung Disease (IUATLD) to assist national tuberculosis programmes in establishing their own policies for TB drug resistance surveillance.[6] This was the first step towards the establishment of a large project on global drug resistance surveillance, whose final objective was to collect and analyse comparable information from a number of countries worldwide.

The strategy to achieve these objectives relies on the implementation of standardized drug resistance surveillance or surveys in a number of countries under the guidance of the network of Supranational Reference Laboratories (SRL). As a first step towards standardization of susceptibility testing, a network of all the laboratories participating in this study was created in 1994. This report deals with the first step in the implementation of strategy, i.e., the adoption of an internationally acceptable laboratory methodology and the maintenance of high levels of drug susceptibility testing proficiency. The results discussed in this paper represent the first round of strains exchanged within the network in late 1994; a second round was carried out in mid 1995 and a third in early 1996.

MATERIALS AND METHODS

Drugs

Isoniazid (INH) Sigma Lot #63H1015 (Sigma-Aldrich Canada Ltd, Oakville, Ontario), rifampicin (RIF) Sigma Lot #123H083625, ethambutol dihydrochloride (EMB) Sigma Lot #24H0118, dihydrostreptomycin sulfate (DSM) ICN Lot #59950 (ICN Pharmaceuticals Canada Ltd, Montréal, Quebec) and streptomycin sulfate (SM) ICN Lot #26148, were tested. The dihydro derivative of streptomycin sulfate is used in egg-containing growth media because it is more stable than SM. Pyrazinamide was not included in the study because of the poor reliability of drug susceptibility testing results obtained when using low pH solid growth media. Antimicrobial base powders were supplied to all sites by the coordinating laboratory to eliminate the source of the drugs as a possible cause for discrepancy in the results. This arrangement will be maintained during the initial rounds of testing until all procedures are properly standardized.

Mycobacterial cultures

Identical sets of ten clinical isolates of *M. tuberculosis* in duplicate (20 cultures) were sent to the laboratories of the network, with the participants' knowledge that the set comprised ten pairs of cultures. Each site received an identifying number known only to the individual laboratory and to the coordinator. The cultures were identified with randomly chosen numbers which varied from site to site. The culture panel included three pairs of drug susceptible isolates, one of which was reference strain H37Rv, and seven pairs of drug resistant isolates encompassing some commonly encountered resistance marker combinations, i.e., two pairs of SM-resistant isolates, two pairs of INH/SM-resistant isolates, two pairs of INH/SM/EMB-resistant isolates and one pair of INH/SM/RIF-resistant isolates.

Drug susceptibility testing methods

In the planning stage of this project, it was decided to allow the participating laboratories to use the method each was most familiar with, in order to eliminate variability due to the disruption induced by changing over to a new testing procedure. The conventional drug susceptibility testing methods being evaluated are those described for *M. tuberculosis*,[7,8] i.e., the absolute concentration method, the resistance ratio method and the proportion method. The proportion method based radiometric BACTEC 460 procedure (Becton Dickinson Microbiology Systems, Sparks, MD) was also included.[9] Laboratories were to adhere strictly to the detailed procedures described in the above-mentioned references. If laboratories so wished, they could conduct parallel testing using their own variation of these standard procedures with the culture panel provided for this round of testing. Laboratories 1, 2, 3, 5, 6, 8, 12, 15 and 16 reported results obtained with the proportion method on Löwenstein Jensen (LJ) medium, while laboratories 10, 13 and 14 reported those obtained with the German Standard Deutsche Industrie Norm (DIN) modification of the proportion method[10] on LJ medium. Laboratory 4 reported results using a modified proportion method on LJ medium where the critical concentration for resistance to SM is 10.0 μg/ml, and laboratory 11 reported those obtained with the proportion method on Middlebrook 7H-10 medium (Difco Laboratories, Detroit, MI, USA). Laboratory 9 reported results obtained with the BACTEC 460 radiometric method and laboratory 7 reported those obtained with the resistance ratio method. This additional testing was to be analysed separately and is reported as 'alternative' testing as opposed to conventional or 'protocol' testing. Laboratories reported their results to the WHO Collaborating Centre for Tuberculosis Bacteriology Research, Ottawa, Canada, for collation and preliminary analysis.

Interpretation of test results

Results were to be interpreted by the participating laboratories as described in the above mentioned refer-

ences.[7–9] Cultures were classified as resistant or susceptible. Actual colony counts or Growth Index (GI) readings were to be kept by each laboratory if needed for evaluation purposes.

Analysis of the data

The sample size of 20 cultures was determined to yield a significance level of $\alpha = 0.05$ to be able to detect a true difference between laboratory methods with a power of 90%. Results from all laboratories were compared against the judicial results, i.e., the agreement of the majority of the participating laboratories was considered the 'gold standard.' Where differences were found in the results of any specific laboratory from those in the group as a whole, contact was made with the director of the laboratory to attempt to determine the possible causes of the discrepancy.

A program for Bayesian analysis designed with Lotus 123 v. 4 software was used to interpret the data. This analysis[11] yields values for sensitivity (ability to detect true resistance), specificity (ability to detect true susceptibility), efficiency (ratio between the number of correct results and the total number of results), and predictive values for resistance (PVR—the rate of true resistances to total resistance) and for susceptibility (PVS—rate of true susceptibility to total susceptibility). Intra-laboratory agreement between duplicate cultures was expressed as percent reproducibility. Analysis of sensitivity, specificity, PVR and PVS was done for resistance to the single drugs INH, RIF, SM and EMB, and as overall mean performance for the four drugs together.

Data were also analysed using the SAS GLM procedure. The analysis of variance (ANOVA) test was used to test the differences between labs and between drug susceptibility results. The confidence intervals (CI) of sensitivity and specificity were calculated with the exact method proposed by Blyth[12] using SAS. The Z (Normal) test was performed to test the differences between protocol and alternative testing based on the assumption that the mean sensitivity and specificity of the protocol testing are the same as those of the alternative testing with constant variance σ^2, which was estimated by the means square error of the ANOVA test for the six laboratories reporting alternative testing.

RESULTS

Protocol testing procedures

STREPTOMYCIN: According to judicial results, the percentage of cultures resistant to SM in the panel sample was 70% (14/20). Table 1 and Figure 1 show that the mean sensitivity value was 88% (46%–100%); the CI of the labs showing 100% sensitivity was 81%–100%, on the other hand labs with sensitivity values as low as 79% reached upper CI values of 100%. Specificity of testing was 100% in all the labs (CI 61%–100%). The mean PVS was 84% (46%–100%), while the PVR was 100% in all labs. The mean efficiency of testing was 92% (63%–100%). The mean reproducibility of results on the ten duplicate cultures was 89% (50%–100%).

ISONIAZID: According to the judicial results, the percentage of cultures resistant to INH in the sample panel was 50% (10/20). Table 1 and Figure 1 show that the mean sensitivity value was 99% (80%–100%), the CI of the labs showing 100% sensitivity was 74%–100% and the specificity was 100% in all laboratories (CI 74%–100%). The mean PVS was 99% (82%–100%), and the PVR was 100% in all laboratories. The mean efficiency of testing was 99% (89%–100%). The mean reproducibility of results on the ten duplicate cultures was 98% (70%–100%).

RIFAMPICIN: According to the judicial results the percentage of cultures resistant to RIF in the sample panel was 10% (2/20). Table 1 and Figure 1 show that the mean sensitivity value was 94% (50%–100%), the CI of the labs showing 100% sensitivity was 22%–100%; the mean specificity was 96% (87%–100%), CI ranged from 64%–100%. The mean PVS was 99% (94%–100%), while the mean PVR was 81% (33%–100%) and the mean efficiency of testing was 96% (88%–100%). The mean reproducibility of results on the ten duplicate cultures was 94% (70%–100%).

ETHAMBUTOL: According to the judicial results, the percentage of cultures resistant to EMB in the sample panel was 20% (4/20). Table 1 and Figure 1 show that the mean sensitivity value was 66%: five laboratories showed sensitivities of 0% (CI 0%–53%), one laboratory showed a sensitivity of 50% (CI 1%–99%), and 11 laboratories showed sensitivities of 100% (CI 47%–100%). The mean specificity was 98%, with 15 laboratories showing values of 100% (CI 83%–100%). The mean PVS was 93% (79%–100%), while the mean PVR was 65% (0%–100%) and the mean efficiency of testing was 91% (75%–100%). The mean reproducibility of results on the ten duplicate cultures was 93% (50%–100%). The difference between the sensitivity of testing ethambutol and that of the three other drugs was statistically significant ($P < 0.05$). There were no statistically significant differences between the specificities obtained for the four drugs tested.

Table 1 shows the analysis of the performance of individual laboratories of the network in the 'protocol' testing of individual drugs. The overall mean performance parameters and 95% CI values for all drugs taken together are graphically represented in Figure 2.

Table 1 also shows the overall performance analysis of the results obtained by the network in the first round of the drug susceptibility proficiency testing for all four drugs taken together. The overall sensitivity was 87% (44%–100%) and specificity was 99% (89%–100%). Predictive values for resistance and suscepti-

Table 1 Overall performance—protocol testing—% values

	Laboratory number																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total
SM																	
Sensitivity	100	100	46	79	79	100	93	57	77	100	100	100	100	100	93	93	88
Specificity	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Predictive Value R	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Predictive Value S	100	100	46	67	67	100	86	50	50	100	100	100	100	100	86	86	84
Efficiency	100	100	63	85	85	100	95	70	81	100	100	100	100	100	95	95	92
Reproducibility	100	100	60	90	90	100	90	60	50	100	100	100	100	100	90	90	89
INH																	
Sensitivity	100	100	80	100	100	100	100	100	100	100	100	100	100	100	100	100	99
Specificity	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Predictive Value R	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Predictive Value S	100	100	82	100	100	100	100	100	100	100	100	100	100	100	100	100	99
Efficiency	100	100	89	100	100	100	100	100	100	100	100	100	100	100	100	100	99
Reproducibility	100	100	90	100	100	100	100	100	70	100	100	100	100	100	100	100	98
RIF																	
Sensitivity	100	100	50	100	100	100	100	100	100	100	100	100	100	100	50	100	94
Specificity	100	100	94	94	100	100	89	100	87	100	89	100	89	100	100	100	96
Predictive Value R	100	100	50	67	100	100	50	100	33	100	50	100	50	100	100	100	81
Predictive Value S	100	100	94	100	100	100	100	100	100	100	100	100	100	100	95	100	99
Efficiency	100	100	89	95	100	100	90	100	88	100	90	100	90	100	95	100	96
Reproducibility	100	100	70	90	100	100	80	100	70	100	100	100	100	100	90	100	94
EMB																	
Sensitivity	100	0	0	100	100	0	100	50	100	100	100	100	100	0	0	100	66
Specificity	100	100	100	100	100	100	100	100	100	100	69	100	100	100	100	100	98
Predictive Value R	100	0	0	100	100	0	100	100	100	100	44	100	100	0	0	100	65
Predictive Value S	100	80	79	100	100	80	100	89	100	100	100	100	100	80	80	100	93
Efficiency	100	80	79	100	100	80	100	90	100	100	75	100	100	80	80	100	91
Reproducibility	100	100	90	100	90	100	100	80	70	100	50	100	100	100	100	100	93
FOUR DRUGS																	
Sensitivity	100	75	44	95	95	75	98	77	94	100	100	100	100	75	61	98	87
Specificity	100	100	99	99	100	100	97	100	97	100	89	100	97	100	100	100	99
Predictive Value R	100	75	63	92	100	75	88	100	83	100	74	100	88	75	75	100	87
Predictive Value S	100	95	75	92	92	95	96	85	88	100	100	100	100	95	90	96	94
Efficiency	100	95	80	95	96	95	96	90	92	100	91	100	98	95	93	99	95
Reproducibility	100	100	78	95	95	100	93	85	65	100	88	100	100	100	95	98	93

SM = streptomycin; INH = isoniazid; RIF = rifampicin; EMB = ethambutol; R = resistance; S = susceptibility.

bility were 87% (63%–100%) and 94% (75%–100%), respectively, and the overall efficiency of testing was 95% (80%–100%). Reproducibility was 93% (65%–100%) and the difference between overall specificity and sensitivity values is statistically significant ($P < 0.05$).

As mentioned above, sensitivity values varied greatly, and because laboratory 3 showed a sensitivity value of 44%, there were significant statistical differences between the sensitivities achieved within the network of laboratories. There seemed to be two clusters of laboratories; in one cluster, consisting of laboratories 2, 3, 6, 8, 14 and 15, the sensitivity of testing was below 90% and in the other cluster, consisting of laboratories 1, 4, 5, 7, 9, 10, 11, 12, 13 and 16, the sensitivity was over 90%. There was no statistically significant difference between the specificity values obtained within the network of laboratories, although in lab 11 the specificity was below 90%. All the other laboratories obtained overall specificities higher than 95%, with laboratories 1, 2, 5, 6, 8, 10, 12, 14, 15 and 16 reaching the 100% level.

Alternative testing procedures

Six laboratories also reported results on alternative testing methods: laboratories 1 and 10 used the BACTEC 460 radiometric method, laboratory 11 the proportion method on Middlebrook 7H11, laboratory 15 a modified absolute concentration method where the results are expressed as proportion of colony forming units resistant to the following drug concentrations: INH 1 μg/ml, SM 20 μg/ml, RIF 50 μg/ml, EMB 5 μg/ml and the critical proportion of resistance is 75%; laboratory 8 used a simplified variant of the proportion method with the critical concentrations INH 0.2 μg/ml, DSM 5 μg/ml, RIF 30 μg/ml, EMB 2 μg/ml and a critical proportion for resistance of 1%, and finally laboratory 5 used its own supply of drugs with the proportion method on LJ medium.

Overall performance analysis

Table 2 shows the overall performance analysis of the alternative testing procedures as reported by six of the

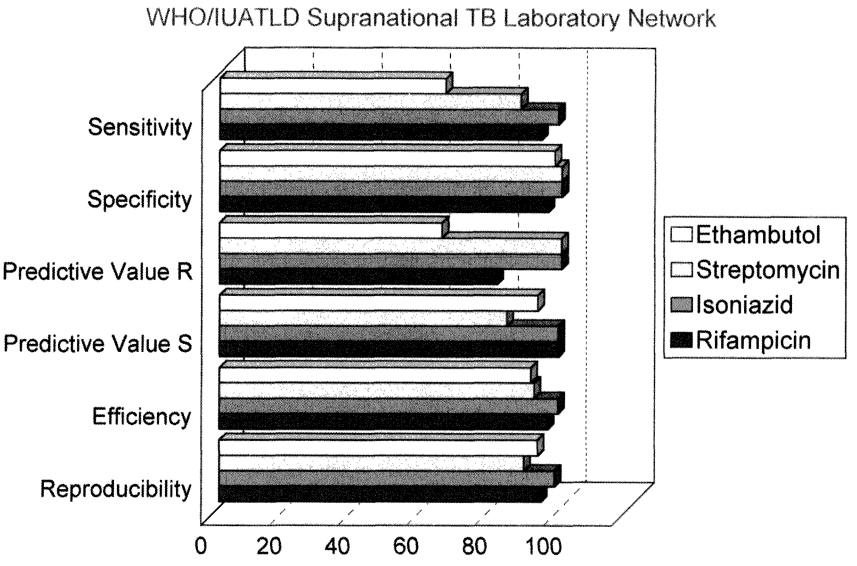

Figure 1 Average 'percentage' results of susceptibility testing. R: resistance; S: susceptibility

laboratories of the network. The results of laboratories 1 and 5 show no differences between protocol and alternative testing; laboratory 8 reports slightly better overall results with the alternative testing, with a sensitivity of 90% versus 77%, a reproducibility of 95% versus 85% and an efficiency of 96% versus 90%. The differences reported are statistically not significant, except in the case of the sensitivity of SM testing ($P < 0.001$). Laboratory 10 reported minor differences in sensitivity (97% versus 100%), reproducibility (98% versus 100%) and efficiency (99% versus 100%) between alternative testing and protocol testing. These differences, however, were not statistically significant. Laboratory 11 reported differences between methods only in the case of EMB, with 100% for sensitivity protocol testing versus 0% for alternative testing, and 69% specificity versus 100%, respectively. These differences are statistically significant ($P < 0.001$). The only difference reported by laboratory 15 was in the testing of SM, which showed a sensitivity of 93% in protocol testing versus 79% in alternative testing. This difference is statistically significant ($P < 0.05$).

DISCUSSION

The aim of this quality assurance programme is not only to standardize drug susceptibility testing procedures but also to provide a solid central reference system against which to compare TB drug resistance surveillance data gathered throughout the world. The test panel of cultures, cross-referenced by the SRL, by itself constitutes a useful tool for the standardization of drug susceptibility testing procedures for *M. tuberculosis*. It is evident that the SRL needs to maintain a high level of proficiency in the testing for drug resistant tuberculosis; to ascertain how high this level can or should be is the ultimate aim of this quality assurance exercise. There is, in clinical bacteriology, a rule of thumb that requires an agreement level of 95% between test results and results obtained with the so-called 'gold standard.' In this study the judicial results were the 'gold standard' against which individual laboratory performance was measured.

There was no obvious overall correlation between the variability of results and the drug susceptibility testing procedure used. Statistical analysis was useful in detecting significant differences between testing, as in the sensitivity EMB testing, which suggested that these differences might be caused by the choice of critical drug concentrations higher (5.0 μg/ml, 7.5 μg/ml) or lower (1.0 μg/ml) than those usually recommended.[6–8,11] Statistical analysis also confirmed one of the most significant findings of this first round of proficiency testing, which is that in most cases specificity values are significantly higher than sensitivity values. The calculation of CI had to take into consideration the fact that the percentage of resistant cultures in the sample varied depending on the drug tested. The CI values obtained were wide because of the limited number of cultures in the panel. In order to tighten

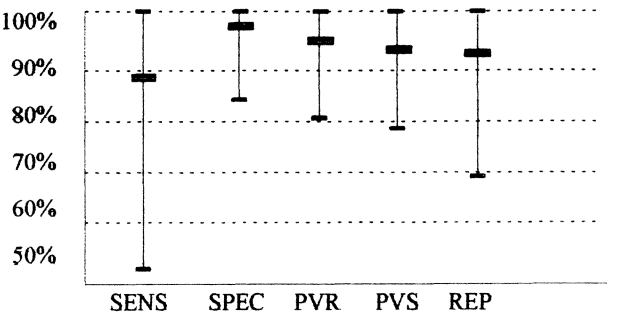

Figure 2 Overall performance analysis protocol testing. Mean percentages and 95% confidence intervals. SENS: sensitivity; SPEC: specificity; PVR: predictive value resistance; PVS: predictive value susceptibility; REP: reproducibility.

Table 2 Overall performance—alternative testing—% values

	Laboratory number						
	1	5	8	10	11	15	Total
SM							
Sensitivity	100	79	93	100	100	79	90
Specificity	100	100	100	100	100	100	100
Predictive Value R	100	100	100	100	100	100	100
Predictive Value S	100	67	86	100	100	67	81
Efficiency	100	85	95	100	100	85	93
Reproducibility	100	100	90	100	100	90	92
INH							
Sensitivity	100	100	100	100	100	100	99
Specificity	100	100	100	100	100	100	100
Predictive Value R	100	100	100	100	100	100	100
Predictive Value S	100	100	100	100	100	100	99
Efficiency	100	100	100	100	100	100	99
Reproducibility	100	100	100	100	100	100	98
RIF							
Sensitivity	100	100	100	100	100	50	94
Specificity	100	100	100	100	89	100	96
Predictive Value R	100	100	100	100	50	100	73
Predictive Value S	100	100	100	100	100	95	99
Efficiency	100	100	100	100	90	95	96
Reproducibility	100	100	100	100	100	90	94
EMB							
Sensitivity	100	100	50	75	0	0	57
Specificity	100	100	100	100	100	100	100
Predictive Value R	100	100	100	100	0	0	100
Predictive Value S	100	100	89	94	80	80	90
Efficiency	100	100	90	95	80	80	91
Reproducibility	100	90	90	90	100	100	96
FOUR DRUGS							
Sensitivity	100	90	90	97	87	73	89
Specificity	100	100	100	100	96	100	99
Predictive Value R	100	100	100	100	93	100	97
Predictive Value S	100	94	94	98	92	86	94
Efficiency	100	96	96	99	93	90	95
Reproducibility	100	98	95	98	100	95	95

SM = streptomycin; INH = isoniazid; RIF = rifampicin; EMB = ethambutol; R = resistance; S = susceptibility.

CI values, the number of cultures to be proficiency tested would have to be substantially increased. Also, it is worthwhile noting that in the case of 100% values, the upper interval has no practical value.

Efficiency values, the percentage of true positivity and true negativity, were highest for isoniazid and rifampicin (99% and 96% respectively) and lowest for streptomycin and ethambutol (92% and 91% respectively). It is worth noting that one pair of cultures was found to be borderline resistant to RIF, a fact that inevitably led to increased discrepancies between laboratory results. Although efficiency values vary according to the relative percentage of resistant and susceptible cultures in the test sample, there is good indication that, at least in this instance and for the two more important drugs, the SRL conformed to the rule of thumb.

Bayesian analysis confirmed that the widest spreads of sensitivity between some of the laboratories occurred in the testing of streptomycin and ethambutol. Conversely, most laboratories showed 100% specificity in their drug susceptibility testing, which in practical terms means that all predictive values for resistance at all prevalence levels were also 100%, i.e., no false resistance. Since the expected prevalence of initial drug resistance in many of the targeted national surveys is approximately 20%, except for the possible case of EMB, even 'low' sensitivity values of 80% would yield predictive values for susceptibility that would meet the generally acceptable level of 95%. The overall sensitivity obtained in laboratory 3 is consistent with possible antimicrobial contamination of the growth medium and can be remedied (personal communication). The reproducibility rate obtained in laboratory 9 was likely due to a poor quality bovine serum albumin used in a batch of BACTEC 460 media[13] which could have caused false susceptibility as well as low reproducibility, a circumstance that does not reflect adversely on the quality of that laboratory's performance. Sensitivities obtained in laboratories 2 and 8 might be due to the fact that results were read at only 28 days, instead of the recommended 42 days. Laboratories 8, 11 and 15 have had a relatively short experience with the methods they re-

ported in protocol testing, a circumstance that might have reflected on some of the results.

Alternative testing results yielded valuable conclusions. The results of laboratory 1 show that there were no differences between the drug susceptibility testing results obtained with the proportion method on LJ medium and those obtained with the radiometric BACTEC 460 method, a fact essentially confirmed by the results obtained by laboratory 10. The results of laboratory 5 show that the source of the four antimicrobials, i.e., those supplied by the coordinating laboratory and those locally available, had no effect on the results of testing. However, laboratory 2 (personal communication) detected differences between the results obtained with the DSM originating from the different suppliers ICN and Sigma. These discrepancies should be further investigated to determine whether manufacturers supply DSM of varying potencies or whether the purported disagreements are due to procedural differences. The results obtained by laboratory 8 seem to suggest that a critical concentration of 5.0 μg/ml of DSM compares better with judicial results than 4.0 μg/ml. These results are difficult to interpret, and could be due to the fact that both sets of results were read at 28 rather than 42 days; they should be investigated further. Conversely, the two sets of RIF results of laboratory 8 are identical, which suggest that there could be no significant difference between critical concentrations of 30 μg/ml and 40 μg/ml of RIF. Protocol testing results obtained with the German Standard (DIN) variant, with a critical concentration of 32.0 μg of RIF, tend to confirm this conclusion. This should be further investigated using more RIF-resistant cultures in the next test panel. The differences in the results of EMB susceptibility testing reported by laboratory 11 prove convincingly that the critical concentration recommended for use in the Middlebrook 7H11 medium (7.5 μg/ml) is not equivalent to that recommended for use in Middlebrook 7H10 (2.0 μg/ml).[14] Results reported by laboratory 15 show that the critical SM concentration of 20.0 μg/ml with a critical proportion of 75% does not compare well with a critical concentration of 4.0 μg/ml and a critical proportion of 1%. However, in the case of the other three drugs, the higher critical concentrations seem to be inexplicably balanced by the choice of a higher critical proportion criterion. Nevertheless, the relatively good agreement between protocol and alternative testing results of laboratory 15 does not translate into good agreement with judicial results in the case of EMB and RIF.

In conclusion, this study shows that drug susceptibility testing within the supranational network of laboratories is reliable for INH and RIF. This is particularly important for the detection of MDR, as MDRTB is defined as TB caused by a strain of the *M. tuberculosis* complex which is resistant to INH and RIF. Some of the discrepancies in this study, especially those detected in the case of SM and EMB, could easily be corrected if the standard critical concentrations, critical proportion and the reading time frames which define drug resistance are adopted. The low specificity of EMB testing leads to under-reporting of drug resistance by almost 50%. Another critical consideration is the fact that some laboratories correct for ethambutol potency and some do not. It is interesting to note, in this context, that the lack of standardization prevailing in the 1960s was the reason why the WHO gathered a group of experts to standardize drug susceptibility testing procedures in 1969.[7]

In a 1977 review of the state of chemotherapy,[15] it was stated 'there are probably very few, if any, laboratories in the world that can perform reliable sensitivity tests.' It would be fair to say, after analysing the results of this study, that this is no longer the case in 1995. There is however, room for improvement by standardization of procedures, especially in the testing of streptomycin and ethambutol. This will be the aim of future proficiency testing rounds. Furthermore, alternative or parallel testing yielded interesting conclusions, and should be continued and refined to yield answers to important questions such as how does the source of drug supply affect the results of testing, and what are the advantages or disadvantages of some of the more popular drug susceptibility testing variant procedures.

Acknowledgements

WHO service contract #T9-181-153 is gratefully acknowledged. The authors wish to thank Mrs. Irena Laszlo for the design of the software needed for Bayesian analysis and the technologists of all sixteen participating laboratories for their excellent technical support.

References

1 Edlin B R, Tokars J I, Grieco M H, et al. An outbreak of multidrug resistance tuberculosis among hospitalized patients with the acquired immunodeficiency syndrome. N Engl J Med 1992; 326: 1514–1521.

2 Pearson M L, Jereb J A, Frieden T R, et al. Nosocomial transmission of multidrug-resistant *Mycobacterium tuberculosis*: a risk to patients and health care workers. Ann Intern Med 1992; 117: 191–196.

3 Angarano G, Carbonara D, Costa D, Italian Tuberculosis Drug Resistance Study Group. Drug resistance of *Mycobacterium tuberculosis* strains isolated from HIV-infected Italian patients: preliminary report from a multicentric study. Microbiologica 1995; 18: 69–72.

4 De Beenhouwer H, Lhiang Z, Jannes G, et al. Rapid detection of rifampin resistance in sputum and biopsy specimens from tuberculosis patients by PCR. Tubercle Lung Dis 1995; 76: 425–430.

5 Laszlo A, de Kantor I N. A random sample survey of initial drug resistance among tuberculosis cases in Latin America. Bull World Health Organ 1994; 72: 603–610.

6 World Health Organization Tuberculosis Programme and International Union Against Tuberculosis and Lung Disease. Guidelines for surveillance of drug resistance in tuberculosis. Document WHO/TB/94.178, Geneva, 1994.

7 Canetti G, Fox W, Khomenko A, et al. Advances in techniques

of testing mycobacterial drug sensitivity and the use of sensitivity tests in tuberculosis control programmes. Bull World Health Organ 1969; 41: 21–43.
8 Public Health Mycobacteriology. A guide for the level III laboratory. 1985. P. T. Kent and G. P. Kubica. U.S. Department of Health and Human Services. Centers for Disease Control, Atlanta, Georgia 30333, pp. 159–184.
9 Siddiqi S H. Bactec TB System. Product and Procedure Manual. 1989. Becton Dickinson, Towson, MD, USA.
10 Medizinische Mikrobiologie Tuberkulosediagnostik. Teil 8: Empfindlichkeitsspufung von Tuberkulosebakterien gegen Chemotherapeutika DIN. 58943-8 September 1994.
11 Toman K. Sensitivity, specificity and predictive value of diagnostic tests. Bull Int Union Tuberc 1981; 56: 18–28.
12 Blyth C R. Approximate binomial confidence limits. JASA 1986; 81: 843–855.
13 Clarridge and Rambo ASM News 1995; 61: 271–272.
14 Ridderhof J, Good R, Muir H, et al. Assessing the performance of drug susceptibility testing for *Mycobacterium tuberculosis* in U.S. laboratories. TB notes Spring/Summer. CDC 1995 pp. 11–12.
15 Fox W. Philip Elman lecture. The modern management and therapy of pulmonary tuberculosis. Proc R Soc Med 1977; 70: 4–15.

RÉSUMÉ

CADRE : Contrôle de qualité du programme de surveillance mondial OMS/UICTMR concernant la résistance aux médicaments.
OBJECTIF : Tester la qualité des procédures d'étude de la sensibilité aux médicaments à l'intérieur même du réseau de laboratoires supranationaux de référence de l'OMS/UICTMR.
SCHÉMA : Un ensemble identique de cultures consistant en 20 isolats cliniques de *Mycobacterium tuberculosis*, comportant à la fois des cultures sensibles et résistantes aux médicaments, a été testé par les 16 laboratoires du réseau en ce qui concerne leur résistance à la streptomycine, l'isoniazide, la rifampicine et l'éthambutol. Les techniques de détermination de la sensibilité aux médicaments comportaient la méthode des proportions, la méthode des concentrations absolues et les méthodes du rapport de résistance, ainsi que leurs variantes, parmi lesquelles la méthode radiométrique BACTEC 460.
RÉSULTATS : Le premier tour testant la compétence a montré que la spécificité de la détermination de la sensibilité médicamenteuse au sein du réseau supranational de laboratoires de référence était significativement plus élevée que sa sensibilité. Ce sont les tests concernant l'isoniazide et la rifampicine qui ont montré un degré élevé de concordance entre les laboratoires, mais des résultats discordants peuvent être obtenus avec la streptomycine et l'éthambutol.
CONCLUSION : Les méthodes de détermination de la sensibilité aux médicaments pour l'isoniazide et la rifampicine, c'est à dire les deux médicaments antituberculeux qui contribuent à la définition de tuberculose multirésistante, sont très fiables à l'intérieur du réseau, les techniques des tests de sensibilité médicamenteuse pour la streptomycine et l'éthambutol nécessitant encore une standardisation.

RESUMEN

MARCO DE REFERENCIA : Control de calidad en el programa OMS/UICTER de vigilancia mundial de la resistencia a los medicamentos.
OBJETIVO : Estudiar la calidad de los procedimientos utilizados par las pruebas de sensibilidad a los medicamentos dentro de la red OMS/UICTER de laboratorios supranacionales de referencia (SRL).
MÉTODO : Conjuntos idénticos de cultivos formados por 20 aislados clínicos de *Mycobacterium tuberculosis*, que contenían cultivos tanto sensibles como resistentes a los medicamentos, fueron sometidos a tests en los 16 laboratorios de la red para comprobar la resistencia a la estreptomicina, la isoniacida, la rifampicina y el etambutol. Los procedimientos de las pruebas de sensibilidad incluían los métodos de proporciones, de concentraciones absolutas, de coeficiente de resistencia, así como sus variantes, entre las cuales el método radiométrico BACTEC 460.
RESULTADOS : La primera etapa de control de calidad mostró que la especificidad de la determinación de la sensibilidad a los medicamentos dentro de la red de SRL era significativamente más alta que su sensibilidad. Las pruebas con respecto a la isoniacida y a la rifampicina mostraron un alto grado de concordancia entre los laboratorios, pero se pueden obtener resultados discordantes con respecto a la estreptomicina y al etambutol.
CONCLUSIÓN : Los procedimientos para determinar la sensibilidad a la isoniacida y a la rifampicina, es decir, los dos medicamentos que definen la tuberculosis multiresistente a los medicamentos, son altamente fiables dentro de la red de SRL. Los procedimientos para determinar la sensibilidad a la estreptomicina y al etambutol necesitan aun una estandarización.