Proficiency Analysis of Drug Susceptibility Testing by National-Level Tuberculosis Reference Laboratories from 1995 to 2003[∇]

Gill-Han Bai,[1] Sang-Jae Kim,[2] Chulhun L. Chang,[1,3,4]* and Members of National and
Regional Tuberculosis Reference Laboratories†

*Korean Institute of Tuberculosis, Seoul,[1] and Department of Laboratory Medicine, School of Medicine,[3] and Medical Research
Institute,[4] Pusan National University, Busan, Korea, and International Union against Tuberculosis and
Lung Disease, Paris, France[2]*

A proficiency review of antituberculous drug susceptibility testing (DST) was undertaken by the regional tuberculosis reference laboratories of the Western Pacific Region of WHO to evaluate the performance of national reference laboratories (NRLs) and to ensure that the results from the participating laboratories are reliable and similar. A panel of 30 *Mycobacterium tuberculosis* strains with various patterns of resistance to isoniazid, rifampin, ethambutol, and streptomycin was sent to the NRLs, and their DST results were analyzed by comparing them with the judicial results. The efficiency scores for each drug were 90 to 99% (mean, 95%) for isoniazid, 77 to 100% (mean, 94%) for rifampin, 82 to 97% (mean, 90%) for ethambutol, and 82 to 98% (mean, 89%) for streptomycin. Significant changes over time in the rates of accordance with the judicial results were observed for rifampin ($P < 0.0001$) and streptomycin ($P = 0.0002$), whereas no changes were observed for ethambutol ($P = 0.0880$). The efficiency score for isoniazid was consistently good throughout the nine rounds. As a whole, NRL02 showed the highest score (95%) in accordance rates for all drugs, while NRL03 (86%) and NRL04 (88%) ranked lowest. Continued proficiency testing with subsequent technical assistance improved the DST quality of participating laboratories, demonstrating the importance of the current WHO/IUATLD external quality assurance program for DST proficiency testing.

In 1994, the World Health Organization (WHO) and the International Union against Tuberculosis and Lung Disease (IUATLD) developed the Global Project on Anti-tuberculosis Drug Resistance Surveillance to determine the magnitude of the problem of drug resistance, which is closely related to the efficiency of treatment programs, and to monitor resistance trends worldwide. Because reliable drug susceptibility testing (DST) is a prerequisite to obtaining accurate and comparable data on drug resistance, a global network of supranational reference laboratories (SRLs) was established to evaluate DST proficiency in countries implementing drug resistance surveillance and to provide technical assistance in improving quality (5, 7–9). Provision of efficient technical assistance was attempted by linking the SRLs to the national reference laboratories (NRLs).

The Department of Microbiology at the Korean Institute of Tuberculosis (KIT), a WHO/IUATLD SRL, has conducted nine rounds of proficiency evaluation of DST between 1995 and 2003 for 16 national or regional laboratories in the Western Pacific Region of WHO. This is the first report presenting the DST proficiency test results from the NRLs in this region that implement the WHO/IUATLD Drug Resistance Surveillance project under the coordination of the SRL network.

MATERIALS AND METHODS

Participating laboratories. Nine rounds of proficiency testing were performed at yearly intervals from 1995 to 2003. The 16 participating NRLs are listed in Acknowledgments. A code was allocated to each laboratory (NRL01 to NRL16). NRL01 to NRL06 participated in six or more rounds. NRL06 discontinued participation as a result of a laboratory change in 1998 to 1999. NRL04 to NRL06 were dropped in round 9, because they were transferred to a new SRL in 2002. These six laboratories were analyzed for their abilities to detect true susceptibility (rate of detection of susceptible strains; specificity) and true resistance (rate of detection of resistant strains; sensitivity), for intralaboratory agreement concerning duplicate cultures (reproducibility), and for their accordance rates (number of correct results divided by total number of results; efficiency). For various reasons, NRL07 to NRL16 participated in four or fewer rounds and were excluded from the statistical analysis.

***Mycobacterium tuberculosis* isolates.** The panel of *M. tuberculosis* strains with various resistance patterns consisted of 15 to 20 clinical isolates. Each round comprised 30 challenges. In the first round, 10 isolates and two identical pairs of another 10 isolates were used. All isolates were collected from Korean patients with or without a history of chemotherapy. From the second to the sixth round, two identical pairs of 15 isolates were used; 10 isolates were provided by the coordinating laboratory (the SRL in Ottawa, Canada), and 5 were selected from strains isolated at KIT or provided by the coordinating laboratory. From the seventh to the ninth round, 10 isolates and two identical pairs of another 10 isolates were used, all of which were provided by the coordinating laboratory (the SRL in Antwerp, Belgium) for the purpose of proficiency testing in that particular year. Shipment of the isolates was subject to the packing and labeling requirements of KIT. For the coordinating laboratory's isolates, the drug susceptibility findings of the majority of the participating laboratories were considered the judicial results. The judicial results for KIT isolates were generated by repetitive testing with the proportion method (PM) in Löwenstein-Jensen medium. The numbers of strains expected to be resistant to each drug are listed in Table 1.

Drug susceptibility tests. The participating laboratories used their own DST methods. Some laboratories used procedures recommended by WHO (1, 2): either the absolute concentration (AC) method, the resistance ratio (RR) method, or the PM. All laboratories used Löwenstein-Jensen medium. Test methods and drug concentrations are given in Table 2. Isoniazid (lot 55HO973; Sigma), rifampin (lot 37HO769; Sigma), ethambutol dihydrochloride (lot

* Corresponding author. Mailing address: Department of Microbiology, Korean Institute of Tuberculosis, Umyon-dong 14, Seoul 137-900, Korea. Phone: 82-2-576-4981. Fax: 82-2-575-3595. E-mail: CCHL@pusan.ac.kr.
† Participating laboratories are listed in Acknowledgments.

TABLE 1. Numbers of drug-resistant strains among 30 challenges tested in each round

Round[a]	No. of strains resistant to:			
	Isoniazid	Rifampin	Ethambutol	Streptomycin
1	18	7	9	21
2	14	14	12	12
3	24	18	14	20
4	24	20	20	24
5	24	14	16	18
6	16	16	12	14
7	18	14	12	12
8	20	11	15	18
9	19	15	8	18

[a] From round 1 (1995) to round 9 (2003).

24HO118; Sigma), and dihydrostreptomycin sulfate (lot 66HO4141; Sigma) were used. Initially, drug powders were supplied, and subsequent requirements for drugs were met by purchase by each laboratory of the same lot numbers from the same manufacturer in order to prevent differences in potency. The NRLs stored the drugs in desiccators at the temperature recommended by the manufacturer and were directed to make solutions according to the manufacturer's instructions.

The DST results were classified as resistant or susceptible. Tests were validated by the susceptibility of *M. tuberculosis* H37Rv, included in the same test series. The SRL (KIT) provided technical assistance through annual supervisory visits and training for laboratory technicians before or after DST proficiency testing.

Data analysis. The sample size of 30 cultures was arrived at, in order to achieve a level of significance of 5% and a power of 90% for detection of a true difference between the results of DST, by using the following equation (6):

$$n = \{Z_{1-\alpha} \sqrt{[P_0(1-P_0)]} + Z_{1-\beta} \sqrt{[P_a(1-P_a)]}\}^2/(P_0 - P_a)^2$$

where P_0 (90%) is the proportion of the test value under the null hypothesis and P_a (70%) is the lowest anticipated accordance range of the DST. The results of the NRLs were compared with the judicial results. The Mantel-Haenszel, or simple chi-square, statistic was adopted for analyzing the meaning of differences (3).

RESULTS

An analysis of the DST results of the six laboratories by rounds is shown in Fig. 1. For isoniazid, agreement values of 77 to 100% for sensitivity, 93 to 99% for specificity, 93 to 100% for reproducibility, and 90 to 99% for efficiency were obtained in all nine rounds. The specificity was poor in rounds 3 through 7. There was no significant change in the accordance rate

throughout the nine rounds ($P = 0.7966$), because the scores were continuously good.

For rifampin, agreement values of 95% or higher were obtained for all four test parameters beginning with the second round (sensitivity, 96 to 100%; specificity, 95 to 99%; reproducibility, 95 to 98%; efficiency, 95 to 100%). In the first round, the specificity (69%), reproducibility (90%), and accordance rate (77%) were poor, whereas sensitivity was good (100%) (Fig. 1). NRL01, NRL03, NRL04, and NRL05 showed significant improvements in their specificities ($P < 0.00095$) and accordance rates ($P < 0.0218$) over time, whereas NRL02 and NRL06 showed no improvement (data not shown). The accordance rate improved greatly throughout nine rounds ($P < 0.0001$) because of the poor performance in round 1 (Fig. 1).

For ethambutol, agreement values of 73 to 100% for sensitivity, 65 to 96% for specificity, 83 to 100% for reproducibility, and 82 to 97% for efficiency were obtained in all nine rounds. The sensitivity in rounds 8 and 9, and thus the efficiency, was poor, even though all other parameter scores increased over time beginning with round 4 (Fig. 1). NRL02 and NRL06 showed significant improvements in reproducibility ($P = 0.0439$ and $P = 0.0061$, respectively), and NRL05 showed significant improvements in sensitivity ($P = 0.0014$) and the accordance rate ($P = 0.0088$). NRL03 was the only laboratory showing deteriorations in sensitivity ($P < 0.0001$) and the accordance rate ($P = 0.0373$) (data not shown). As a whole, no significant improvement in the accordance rate was observed ($P = 0.0880$), because of the poor sensitivity in rounds 8 and 9 (Fig. 1).

For streptomycin, agreement values of 71 to 100% for sensitivity, 84 to 98% for specificity, 78 to 99% for reproducibility, and 82 to 98% for efficiency were obtained in all nine rounds. All parameter scores improved with time beginning with round 5, except for reproducibility in round 8 (Fig. 1). NRL03 and NRL06 improved their specificities ($P = 0.0005$ and $P = 0.0023$, respectively) and accordance rates ($P = 0.0200$ and $P = 0.0238$, respectively). NRL04 also improved its accordance rate ($P = 0.0356$). There was no significant change in the performances of NRL01, NRL02, and NRL05 (data not shown). As a whole, the accordance rate improved greatly throughout the nine rounds ($P = 0.0002$) (Fig. 1).

Table 3 and Fig. 2 show each laboratory's scores on the proficiency parameters. For isoniazid, agreement values of

TABLE 2. Methods and drug concentrations used by participating laboratories

Laboratory	Round(s)	Method	Concn of drug (μg/ml)			
			Isoniazid	Rifampin	Ethambutol	Streptomycin
NRL01	1	AC	0.2	40	2	4
	2–9	PM	0.2	40	2	4
NRL02	1	AC	0.2	40	2	4
	2–9	PM	0.2	40	2	4
NRL03	1	RR	0.2	40	2	4
	2–9	PM	0.2	40	2	4
NRL04	1	AC	1.0	50	5	10
	2–8	AC	0.2	40	2	10
NRL05	1–8	AC	1.0	64	4	32
NRL06	1	RR	0.2	40 (7H10)[a]	2	10
	2, 3, 6–8	PM	0.2	40	2	4

[a] NRL06 used 7H10 (Middlebrook 7H10 agar) for rifampin susceptibility testing only for the first round.
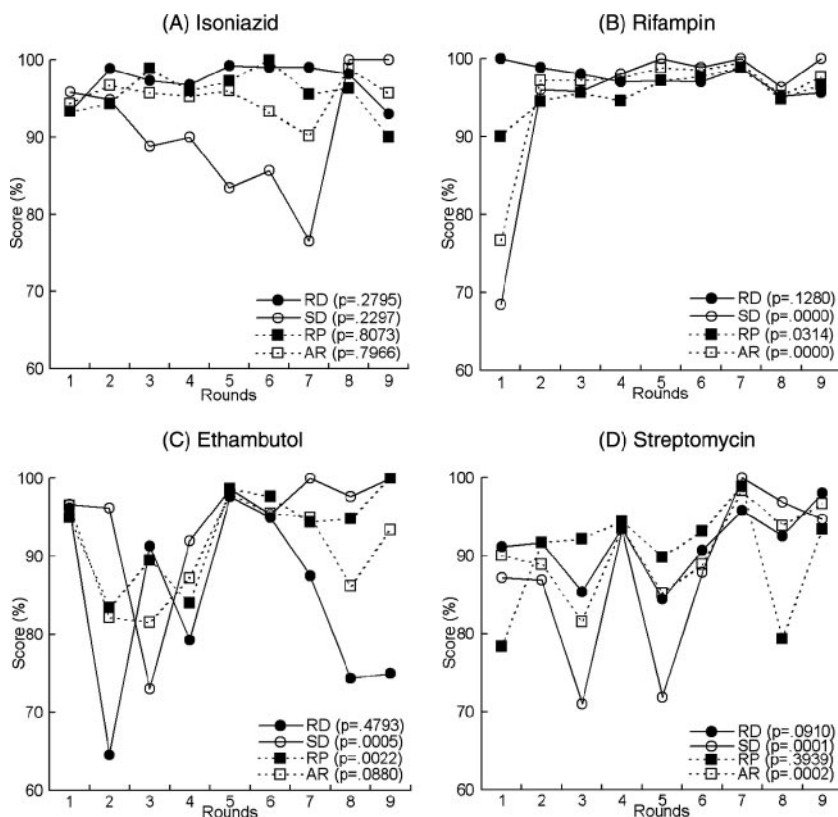
FIG. 1. Performance of six participating laboratories in the Regional Reference Laboratory network DST proficiency testing. The *P* value is a Cochran-Mantel-Haenszel chi-square statistic. RD, rate of detection of resistant strains (sensitivity); SD, rate of detection of susceptible strains (specificity); RP, reproducibility; AR, accuracy ratio (ratio of accordant results to total test results), or efficiency.

90% or higher were obtained by all NRLs, except for reproducibility for NRL03 (88%) and specificity for NRL05 (70%). For rifampin, agreement values of 90% or higher were obtained by all laboratories, except for specificity for NRL06 (87%). For ethambutol, agreement values of 89 to 97% for sensitivity, 71 to 91% for specificity, 89 to 98% for reproducibility, and 87 to 95% for efficiency were obtained by all laboratories. For streptomycin, agreement values of 71 to 96% for sensitivity, 77 to 97% for specificity, 81 to 93% for reproducibility, and 86 to 93% for efficiency were obtained by all laboratories.

The mean efficiency of DST for isoniazid in the six laboratories was 95%, and all but one of the laboratories showed a score the same as or higher than the mean − 1 standard deviation (SD) (92%). The exception was NRL05, which had a 91% efficiency score (Fig. 2A). The mean efficiency of DST for rifampin in the six laboratories was 94%, and all laboratories scored the same as or higher than the mean − 1 SD (91%) (Fig. 2B). The mean efficiency of DST for ethambutol was 90%, and all laboratories had a score higher than the mean − 1 SD (87%) (Fig. 2C). The mean efficiency for streptomycin was 89%, and all laboratories except NRL03 (86%) had a score higher than the mean − 1 SD (87%) (Fig. 2D).

DISCUSSION

The results of DST differ greatly depending on technical precision, especially in the incorporation into the medium of

TABLE 3. Laboratory scores on DST proficiency parameters[a]

NRL	Isoniazid				Rifampin				Ethambutol				Streptomycin			
	SD	RD	RP	AR	SD	RD	RP	AR	SD	RD	RP	AR	SD	RD	RP	AR
NRL01	95.3	96.2	95.2	95.9	92.9	95.7	94.1	93.7	89.4	89.7	92.0	90.0	82.9	91.9	92.8	89.7
NRL02	95.9	95.9	97.4	97.1	97.0	97.0	96.1	95.7	96.8	70.8	85.5	87.4	95.5	77.0	81.1	88.1
NRL03	91.0	94.0	87.8	92.7	90.6	95.6	91.8	91.2	92.3	81.7	90.8	87.8	71.2	96.6	87.0	86.3
NRL04	95.9	95.9	97.4	97.1	97.0	97.0	96.1	95.7	96.8	70.8	85.5	87.4	95.5	77.0	81.1	88.1
NRL05	70.1	100	99.1	90.8	97.6	100	99.1	98.4	96.9	91.0	98.3	95.4	95.0	93.8	91.1	93.3
NRL06	92.5	100	97.8	96.7	86.5	98.0	92.2	91.0	93.0	86.5	91.0	91.0	83.2	94.2	88.2	90.0

[a] All scores are expressed as percentages. SD, rate of detection of susceptible strains (specificity); RD, rate of detection of resistant strains (sensitivity); RP, reproducibility; AR, accuracy ratio (ratio of accordant results to total test results), or efficiency.
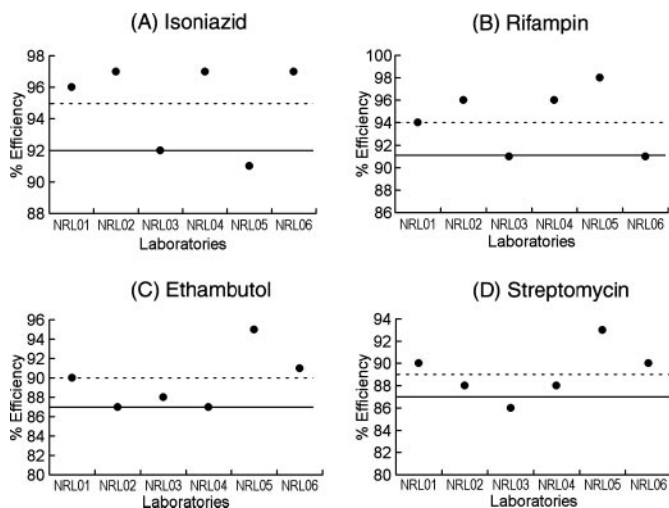
FIG. 2. Laboratory scores for DST performance. Dotted lines, mean efficiency; solid lines, mean − 1 SD.

the required concentration and potency of drugs and in the inoculation of the required number of viable organisms (4). Therefore, improvement in DST is not easily evaluated simply by comparing results on one or two tests within a short period. In the first round of DST, it was recommended that all participating NRLs use the method most familiar to them in order to determine the accuracy of the DST conducted by each NRL and to use the first-round results as a control. We found various technical differences that adversely affected the results and might cause difficulties in proficiency testing, even for laboratories using internationally recommended methods. Thus, after the first round, we advised all laboratories to use the standardized methods taught in supervisory visits.

Among the 16 national or regional laboratories in the Western Pacific Region participating in this testing from 1995 through 2003, only 6 NRLs participated in six or more rounds. These six laboratories are all officially designated reference laboratories, so their results are particularly important, whereas the others were tentatively designated reference laboratories for the purpose of drug resistance surveillance. The mean efficiency scores of the six reference laboratories were 95% for isoniazid, 94% for rifampin, 90% for ethambutol, and 89% for streptomycin. According to Laszlo et al., efficiencies lower than 89% for isoniazid, 95% for rifampin, and 80% for streptomycin and ethambutol are considered substandard performance (5). The efficiency scores were much higher than these limits for ethambutol and streptomycin and somewhat higher for isoniazid. Therefore, we suggest that the acceptance level for efficiency scores for ethambutol and streptomycin be elevated to 89 to 90%. The efficiency levels for rifampin at three NRLs failed to satisfy the limits (91 to 94%), because their performances were poor in the first round. However, performance on all parameters was excellent beginning with round 2.

DST failures at participating laboratories might be caused by several factors, such as inappropriate (heavy or poorly dispersed) inocula or inaccurate concentrations of drugs. NRL03 and NRL05 showed poor performance through several rounds,

whereas NRL02 and NRL06 performed poorly in only one round, indicating that some laboratories are serious about quality improvement, whereas others might not be. The isoniazid resistance found incorrectly by NRL05 even at a high drug concentration (1 μg/ml) might have resulted from a heavy inoculum. The frequent false results in the detection of isoniazid-susceptible strains by NRL03 indicate gross errors in the whole DST procedure, because the results for the other drugs also were poor. The poor performances with rifampin noticed for some laboratories on the first advisory visits probably were attributable to the use of an improper solvent, namely, acid-alcohol instead of dimethylformamide, and reuse of the stock solution. Both of these errors were corrected. The specificity, reproducibility, and accordance rate for ethambutol improved significantly in later rounds. Proficiency testing results for streptomycin showed significant improvement after round 7, whereas wide variations had been observed previously. Taken as a whole, the results with streptomycin showed significant improvement through nine rounds. It is not easy to explain how NRL05 could have maintained an acceptable range of sensitivity with very high critical concentration, but use of a heavy inoculum could be a reason, as could inclusion of a large number of resistant strains in the panels. In addition, inclusion of strains with low levels of resistance to isoniazid or rifampin may lead to poor accordance with the judicial results. Therefore, strains need to be selected carefully for each test panel in order to maintain the same degree of difficulty.

In conclusion, with the technical assistance of supervisory visits and training, most of the participating laboratories significantly improved the quality of their DST through the consecutive rounds of proficiency testing. Thus, the current program of DST should be continued in order to maintain acceptable proficiency of NRLs at drug resistance surveillance and multidrug-resistant tuberculosis management.

REFERENCES

1. **Canetti, G., W. Fox, A. Khomenko, et al.** 1969. Advances in techniques of testing mycobacterial drug susceptibility testing, and the use of sensitivity tests in tuberculosis control programmes. Bull. W. H. O. **41:**21–43.
2. **Canetti, G., S. Froman, J. Grosset, et al.** 1963. Mycobacteria: laboratory methods for testing drug sensitivity and resistance. Bull. W. H. O. **29:**565–578.

3. **Fleiss, J. L.** 1981. Statistical methods for rates and proportions. Wiley-Interscience, New York, NY.
4. **Kim, S. J.** 2005. Drug-susceptibility testing in tuberculosis: methods and reliability of results. Eur. Respir. J. **25:**564–569.
5. **Laszlo, A., M. Rahman, M. Espinal, and M. Raviglione.** 2002. Quality assurance programme for drug susceptibility testing of *Mycobacterium tuberculosis* in the WHO/IUATLD Supranational Reference Laboratory Network: five rounds of proficiency testing, 1994–1998. Int. J. Tuberc. Lung Dis. **6:**748–756.
6. **Lwanga, S. K., and S. Lemeshow.** 1991. Sample size determination in health studies. World Health Organization, Geneva, Switzerland.
7. **World Health Organization.** 2000. Anti-tuberculosis drug resistance in the world. Report no. 2: prevalence and trends. The WHO/IUATLD Global Project on Anti-tuberculosis Drug Resistance Surveillance. WHO/CDC/TB/2000.278. World Health Organization, Geneva, Switzerland. http://whqlibdoc.who.int/hq/2000/WHO_CDS_TB_2000.278_intro.pdf.
8. **World Health Organization.** 2004. Anti-tuberculosis drug resistance in the world: third global report. The WHO/IUATLD Global Project on Anti-tuberculosis Drug Resistance Surveillance, 1999–2002. WHO/HTM/TB/2004.343. World Health Organization, Geneva, Switzerland. http://www.who.int/tb/publications/who_htm_tb_2004_343/en/.
9. **World Health Organization.** 1997. Anti-tuberculosis drug resistance in the world. The WHO/IUATLD Global Project on Anti-tuberculosis Drug Resistance Surveillance, 1994–1997. WHO/TB/97.229. World Health Organization, Geneva, Switzerland. http://whqlibdoc.who.int/hq/1997/WHO_TB_97.229.pdf.