Reliability of skin smear results: experiences with quality control of skin smears in different routine services in leprosy control programmes

L VETTOM & S PRITZE

Armauer Hansen Institute, Leprosy Reference Laboratory, Würzburg, Germany

Accepted for publication 23 February 1989

Summary The quality control of skin smears is an important tool in improving the diagnosis of leprosy. We evaluated the skin smears sent to us by 50 laboratory technicians of 29 projects in Asia, Africa and South America. The skin smears were judged according to taking, staining and reading. The correlation was altogether satisfactory. In reading, a low correlation was found in 11% (42 slides) and it was seen that the highest percentage of low correlations was found in the false negative smears. The evaluation of cases with a low correlation leads to the conclusion that using the new WHO classification of 1988 will not reduce the number of incorrectly classified cases. From 42 slides showing a low correlation of their BI results, 7% led to a different classification (paucibacillary instead of multibacillary or vice versa) according to the WHO definition given in 1982, but 8% according to the 1988 WHO definition.

Introduction

The examination of slit-skin smears—notably in the context of multiple drug therapy (MDT)—is an important tool in leprosy control programmes, namely for the diagnosis and classification of the disease, the assessment of progress, the duration of treatment in multibacillary cases and the diagnosis of relapses.¹

Experiences with the standard of the bacteriological examination techniques are often disappointing. Georgiev & McDougall¹ recently stated that in a considerable number of countries in the main leprosy-endemic areas the standard of work regarding the taking, the fixing, the dispatching, the staining and the reporting of slit-skin smears was deplorably low.

Even though efforts for improvement have been undertaken in different countries, e.g. by intensifying the training of technicians, the overall situation was not suspected to have changed significantly.

In 1987, when we started quality control in our leprosy reference laboratory, the main objective was to get more detailed information on the actual standard of laboratory work in various leprosy control programmes.

188 L Vettom and S Pritze

To date, little literature is available to describe a systematic approach to the quality control of slit-skin smears in leprosy. In 1985 de Rijk *et al.*² published their results on a periodical reexamination of samples of skin smears for leprosy taken in routine services. This quality control was based on the assessment of smearing and staining, as well as the assessment of the differences in BI results of slit-skin smears.

The quality control presented here was based on the criteria proposed by de Rijk *et al.* A detailed written evaluation was given to each participating technician in order to point out reasons which might have lead to a disagreement in the BI results.

Methods

Twenty-nine projects with 50 technicians working in different leprosy service laboratories worldwide were evaluated in this study. Table 1 gives a summary of continents and countries, number of projects and technicians participating in the quality control.

Information sheets and forms were sent to each project (Figure 1).

The comparison between the service laboratory and the leprosy reference laboratory was based on 8 slides with up to 6 smears each taken within the last 2 months. The results of all 8 slides read by one reader in the service laboratory, were reported in the corresponding circles on the form by someone responsible for the organization of the quality control in the respective laboratory.

The form was sent to our reference laboratory, where the smears were examined by a standardized procedure.

Smear taking was judged according to:

The amount of lymph (thin, thick);

Its quality (assessed by the presence of different cell types, e.g. lymphocytes, macrophages, erythrocytes and epithelial cells in the lymph);

Its distribution.

Smear staining technique was judged according to:

The demonstration of acid fastness of bacilli;

The contrast between bacilli and background;

The presence of artefacts.

Continents	Countries	No. of projects	Technicians
Asia	India	19	23
	Pakistan	1	2
	Korea	1	2
	Thailand	1	1
Africa	Ethiopia	1	10
	Nigeria	2	5
	Kenya	1	2
	Uganda	1	2
	Tanzania	1	1
S. America	Brazil	1	2

Table 1.

QUALITY CONTROL

Name and adress of Service Laboratory:	Date of sample collection:
	Samples chosen by:
	Used staining technique:
taken by:	a) Ziehl-Neelson (hot) 🗌
	b) Ziehl-Neelson (cold) 🗌
Name and Position of Examiner:	c) others
	Short description of staining technique used:

Pat. Name	Slide-No.	Date	Results	Any specific comments
			BI BI BI BI BI BI	Service Lab:
			000000	Reference Lab:
			000000	Service Lab:
			000000	Reference Lab:
			000000	Service Lab:
			000000	Reference Lab:
			000000	Service Lab:
-			000000	Reference Lab:
			000000	Service Lab: Reference Lab:
			000000	Service Lab:
			000000	Reference Lab:
			000000	Service Lab:
			000000	Reference Lab:
			000000	Service Lab:
			000000	Reference Lab:
			000000	Service Lab:
			000000	Reference Lab:
			000000	Service Lab:
	-		000000	Reference Lab:

(DAHW208/88)

Figure 1. Form used for quality control: circles under 'results' represent various smears with the number of bacilli found reported as bacteriological index (BI) according to Ridley's logarithmic scale.

190 L Vettom and S Pritze

BI accordance was determined according to the following scheme:

Very good, exact BI accordance is found in 6 out of 8 slides;

Good, half of the 8 slides show exact BI accordance and the rest slight deviations ('slight' deviation: BI difference of ± 1 in smears with BI > 2);

Satisfactory, 3 of 8 slides show exact BI accordance and the rest show slight deviations; Unsatisfactory, less than 3 slides out of 8 show exact BI accordance.

One experienced person served as second reader. Through this procedure a relatively objective assessment was achieved. Double-blind conditions were not put into effect in order to analyse deviations in the results and to explain the reason for them to the technician whenever possible. Besides the evaluation of the reliability of the results an individual feedback was given to the technician from which we hoped to stimulate critical analysis and motivation.

A comment was written on the form, summarizing the assessment of the smears. Suggestions for improvement were given whenever necessary and participation in further quality control was proposed. The completed form was sent back to the service laboratory.

Results

The results of the quality control are summarized in Table 2. The quality of smear taking conducted by 13 technicians was unsatisfactory according to our assessment. In these slides only a small amount of lymph, irregularly distributed and containing lots of erythrocytes was found.

By 15 technicians the quality of taking was satisfactory if the smears showed a sufficient amount of lymph or regular distribution of material or scanty erythrocytes.

By 22 technicians the quality of taking was good, characterized by a sufficient amount of lymph, regularly distributed and with only scanty or no erythrocytes.

By 11 technicians the quality of staining was unsatisfactory under the following conditions: if the smears showed many artefacts predominantly due to unfiltered staining solutions, if an assessment of acid fastness ('pale' bacilli) was difficult and if there was a contrast, which rather impeded the correct judgment of acid fastness.

The quality of staining of 24 technicians was evaluated as satisfactory based on the absence of artefacts or the easy assessment of acid fastness or a good contrast.

The staining conducted by 15 technicians was considered to be good since it showed no artefacts and a good demonstration of acid fastness underlined by a good contrast.

	No. of	Т	akir	ıg	S	taini	ng]	Read	ing	
Continent	technicians	+	±		+	±		++	+	±	
Asia	28	14	4	10	7	13	7	1	12	6	9
Africa	20	8	9	3	8	9	4	0	9	2	9
A. America	2	0	2	0	0	2	0	0	2	0	0
Total	50	22	15	13	15	24	11	1	23	8	18

Table 2. Number of technicians in different continents according to the quality of their performance in taking, staining and reading slit-skin smears

 $+ + \text{very good}; + \text{good}; \pm \text{satisfactory}; - \text{unsatisfactory}.$

Table 3. Staining techniques used

1	Hot carbolfuchsin alk. methylene blue	(15 min); (10 min);	1% HCl in 70% ethanol (2-5 sec) (used by 13 technicians)
2	Hot carbolfuchsin alk. methylene blue	(10 min); (0·5 min);	3% HCl in 70% ethanol (5 min) (used by 4 technicians)
3	Hot carbolfuchsin alk. methylene blue	(15 min); (2 min);	5% H ₂ SO ₄ (1 min) (used by 7 technicians)
4	Hot carbolfuchsin alk. methylene blue	(10 min); (5 min);	20% H ₂ SO ₄ (2 min) (used by 1 technician)
5	Hot carbolfuchsin brilliant green	(10 min); (2 min);	20% H ₂ SO ₄ (1 min); 1% HCl (used by 10 technicians)
6	Cold carbolfuchsin alk. methylene blue	(20 min); (10 min);	1% HClin 70% ethanol(>5 sec) (used by 6 technicians)
7	Cold carbolfuchsin methylene blue	(20 min); (2 min);	5% HCl in 70% ethanol (5 min) (used by 1 technician)
8	Cold carbolfuchsin methylene blue	(20 min); (2 min);	20% HCl in 70% ethanol (1 min) (used by 1 technician)
9	Cold carbolfuchsin methylene blue	(10 min); (2 min);	10% H ₂ SO ₄ (20 min) (used by 5 technicians)
10	Cold carbolfuchsin methylene blue	(20 min); (2 min);	25% H_2 SO ₄ (10–20 min) (used by 2 technicians)

It is noteworthy that in the 29 participating laboratories 10 different staining methods were used, representing the routine stainings. These differences were mainly determined by the choice of hot or cold carbolfuchsin, the employed counterstaining with methylene blue or brilliant green and the agents, concentrations and times used for decoloration.

Summarizing Table 3, 35 technicians (70%) used the hot-staining method and 15 (30%) the cold-staining method.

According to the participating technicians the results of the readings (BI accordance) were satisfactory (Table 4).

When compared to the reference laboratory only one technician had very good results in the readings, 23 technicians had good results, 8 had satisfactory results and 18 had unsatisfactory results in their readings.

When the results of the readings were analysed in regard to the geographic distribution of the technicians, the folowing data were found (Table 5): In fact, the number of participating technicians is too small to draw an objective conclusion from Table 5. If only very good and good results in the readings are taken into account a very small difference in the BI accordance is found between Asian

Table	4. Assessment of reading accord-
ing to	50 participating technicians

BI accordance	No. of technicians
+ +, very good accordance	1
+, good accordance	23
\pm , satisfactory	8
-, unsatisfactory	18

		NL C		BI-accordance			
Continents	Countries	No. of projects	Technicians	+ +	+	±	
Asia	India	19	23	1	9	5	8
	Pakistan	1	2		2		
	Korea	1	2				2
	Thailand	1	1		1		
Africa	Ethiopia	1	10		6	2	2
	Nigeria	2	5		1		4
	Kenya	1	2		2		
	Uganda	1	2			1	1
	Tanzania	1	1 Data Press			2	1
S. America	Brasilien	1	2		2		

 Table 5. Assessment of readings according to 50 participating technicians from 10 countries

very + + good; + good; \pm satisfactory; - unsatisfactory

(BI accordance + +/+ by 28 technicians (46%)) and African technicians (BI accordance + by 20 technicians (45%)).

The choice of regimen and the treatment duration in leprosy is determined by the number of bacilli found in the patient. In 1982 WHO recommended a classification of leprosy in paucibacillary (BI < 2 at any site) and multibacillary (BI \geq 2 at any site) cases, implicating a 6-month treatment with two drugs in paucibacillary patients and a minimum 2-year treatment with three drugs in multibacillary (24).

When, independently of the results obtained by the individual technician, the quality of reading was analysed in all 390 slides examined, full correlation (reading 'same') of the BI was found in 173 slides (45%) (Table 6):

The highest percentage of full correlation was observed in the 106 negative slides $(27 \cdot 5\%)$; in multibacillary slides a full correlation was found in 67 $(17 \cdot 5\%)$. No full correlation was seen in paucibacillary smears. Acceptable correlation ('slight' deviation) was assessed in 343 (89%) of the 390 slides examined. Low correlation in the reading was found in 42 (11%) of the evaluated slides, listed in Table 7.

Service laboratory (BI)	Reference laboratory (BI)	No. of slides
Negative	Negative	106 (27.5%)
Paucibacillary	Paucibacillary	0
Multibacillary	Multibacillary	67 (17.5%)
		173 (45%)

Table 6. Full correlation of BI results in the interobserver comparison

Service laboratory	Reference laboratory	No. of slide		
Negative	Paucibacillary Multibacillary	10 (2·6%) 12 (3·0%)		
Paucibacillary	Negative Multibacillary	5 (1·3%) 8 (2%)		
Multibacillary	Negative Paucibacillary	4 (1%) 3 (0·8%)		
		42 (11%)		

 Table 7. Low correlation in interobserver comparison of slit-skin smears according to negative, paucibacillary and multibacillary BIs.

Discussion

The quality of slit-skin smears is one of the most important factors influencing the reliability of the bacteriological index (BI). The quality itself depends on different parameters. Some of these can easily be evaluated by quality controls, whereas analysing others is difficult or impossible. An unmistakable influence on the quality is evident in the taking and staining of skin smears, the quality of staining dyes and glassware, the fixing of smears and the dispatching of slides.

On the contrary, it is difficult if not impossible, to judge the technician's training, his reliability and his motivation, the correct selection of the sites chosen for taking the smears, the quality and maintenance of microscopes and the availability of good light sources.

Other hindrances encountered are due to the fact that the taking of skin smears and the staining/ reading is often performed by different persons, so that the technician reading the smear is only partly or not at all responsible for its quality. Often the technicians do not get any clinical data, resulting in an isolated analysis without the corresponding context. If they were to have some clinical information they would probably be able to work more conscientiously. Finally, the results of the quality control are usually communicated only to the reader but not to those in charge of taking the smears. This lack of communication makes efficient improvement difficult.

As shown in Table 3, 10 different staining methods were used. There was no clear evidence that the quality of staining was significantly influenced by the method applied. It was not possible to determine the impact of the staining techniques on the number of bacilli found since only stained slides were available; however, in some slides the suspicion arose that the acid fastness of bacilli was not preserved or had disappeared.

Three slides were sent as negative to the reference laboratory, but were suspected to have been positive since very weakly stained bacilli were seen. Slides were restained and proved to be obviously positive. In these cases decoloration was probably inadequate but because of the small number of slides no definite conclusions in regard to the influence of the staining technique used can be drawn.

In 8 slides BI results of the service laboratory were significantly higher than those of the reference laboratory and fading of bacilli was suspected. After restaining the results showed a good correlation. Reasons for fading might be the continuing reaction of the acid if slides have not been washed carefully after decoloration with acid–alcohol, influence of immersion oil, xylene or sunlight, as discussed in the literature.³⁻⁵

In the sample of slides examined there was no tangible evidence that the quality of staining was significantly influenced by the method used. However, it is to be recommended that hydrochloride acid which is less aggressive to the decoloration of the bacilli than sulphuric acid should be used as the agent of choice.

194 L Vettom and S Pritze

The assessment of the BI accordance is a subjective method, and differences in the interobserver comparison have to be accepted to a certain extent.

Even under ideal conditions slight differences may be found mainly due to the irregular distribution of bacilli in the smear and the impossibility of accurately reading the same 100 oil immersion fields assessed by the first reader. On the other hand, especially if the bacteriological index is > 2 the count of bacilli should result in the exact value within the logarithmic scale.

Assessing the BI accordance the single smear results of the service laboratory were compared to the corresponding results determined by the reference laboratory. Intentionally we did not use the average BI/slide in order to identify the slides where patients were incorrectly classified as paucibacillary and multibacillary according to the WHO definition.⁶

This definition describes those patients as paucibacillary where the BI is < 2 at any site and as multibacillary those where the BI is ≥ 2 at any site. If the average BI is chosen as criterium, some multibacillary patients might be dismissed. In a patient with, e.g. a BI result of 1/1/1/1 per smear, the average BI will be one, which leads to the classification as paucibacillary. In a second patient the BI is 2/0/0/2 per smear leading to an average BI of 1, and suggesting a paucibacillary case whereas the patient is multibacillary.

As shown in Table 7 there were relatively few slides read as multibacillary by the service laboratory which were judged to be negative or paucibacillary by the reference laboratory (1.8%). Possible reasons for this might be artefacts which are sometimes difficult to distinguish from acid-fast bacilli (crystals, haemoglobin conglomerates, contaminants).

Two per cent of the slides were judged as paucibacillary by the service laboratory though these were multibacillary according to the reading of the reference laboratory. The most important reason for this was probably the fact that only a few oil immersion fields were counted, often in combination with an uneven distribution of the lymph on the slide. Some of these paucibacillary slides were judged to be negative by the reference laboratory ($1\cdot3\%$). Artefacts mistaken for bacilli were found to be the most common reason.

The highest percentage of low correlation was observed in the false negative smears (5.6%), probably due to the counting of only a small number of oil immersion fields (motivation, lack of time by the technician), bad microscopes and light sources as well as inadequate staining techniques. This is a very interesting finding, especially in the light of the WHO definition published in 1988 that paucibacillary cases are all those with a negative BI result, multibacillary all those with any positive BI result.⁷ Based on our findings the careful conclusion can be drawn that if only the BI result is considered for classification, the number of incorrectly classified patients will not come down significantly. From 42 slides showing a low correlation of their BI results (Table 7), 27 (7%) led to a different classification (paucibacillary instead of multibacillary or vice versa) according to the WHO definition.

De Rijk *et al.*² already raised the question of the possibility of drawing a line between acceptable and unacceptable results. They proposed three criteria, namely the proportion of full correlation when the readings are the 'same' or the difference 0; acceptable correlation when the difference is 1 mark BI to either side; and the calculation of the variance as the sum (Σ) of the square values of the differences (*D*) divided by the number of observations (*N*). Based on these criteria the reliability of the BI results was then judged to be good according to the value of \geq 50% full correlation and/or 50–80% of acceptable correlation and/or a variance < 1 in all samples examined by a reference reader.

Relating to these values our reading showed full correlation in 45% of all smears examined, acceptable correlation in 89% and variance < 1 in 93.8%. Therefore, the conclusion can be drawn that our results are at least satisfactory.

The quality control was well accepted; up to now 50% of the projects which were asked to participate have sent slides and forms. This good response can partly be explained by the fact that in many leprosy control programmes the doctors are in charge of the supervision of the laboratory work, but do not feel able to do the assessment. Sending slides to a reference laboratory is therefore

a welcomed possibility to free themselves of this burden and at the same time give some feedback to the laboratory technician.

In due course we will know more details about the impact of the quality control on the work standard in the laboratory. The feedback we have got so far shows much interest and enthusiasm from the participants. In the future, experiences gained in quality control should stimulate further training, as well as influence the curricula of training and refresher courses.

Acknowledgments

We wish to thank all persons involved in projects of GLRA participating in the quality control and evaluated in this study, particularly with respect to the Leprosy Relief Rural Centre, Chettipatty; Assisi HD Hospital, Talavadi; Vimala Dermatological Centre, Bombay; Pushparani Clinic, Genguvarpatti; Vimukti, Kakinada; Urban Leprosy Treatment & Education Scheme, Cochin; St Mary's Leprosy Centre, Bhimavaram; Damien Institute, Trichur; St Joseph's Convent, Arni; Dinasevanasabha, Pattuvam; IELC Leprosy Control Project, Tamil Nadu; Sivananda Rehabilitation Home, Kukatpally, Hyderabad; Greater Visakha Leprosy Treatment & Health Education Scheme, Visakhaptanam; St Thomas Hospital and Leprosy Centre, Chettupattu; 'GRECALTES', Calcutta; Damien Leprosy Control Centre, Nilakottai; Lok Seva Sangam, Bombay; St Joseph's Leprosy Hospital, Arakkiapuram; Sagayamatha Hospital, Pullambadi—all in India. The Catholic Skin Disease Clinic, Buk-Ku, Eubnaedong, S. Korea; Rawalpindi Leprosy Hospital, Rawalpindi, Pakistan; Daughters of Charity St Vincent de Paul, Niramon Clinic, Bannoi Samliem Khonkaen, Thailand; Buluba Leprosy Centre, Kampala, Uganda; Kagando Hospital, Nairobi, Kenya; Mile Four Hospital, Abakaliki, Nigeria; RTLC, Ruvuma Region, Peramiho, Tanzania; ALERT, Addis Ababa, Ethiopia; Specialist Hospital Ossiomo, Agbor, Nigeria; Centro de Dermatologia Alfredo da Matta, Manaus, Brazil. Thanks also to Mrs I Holzheimer-Colantuoni for secretarial help.

References

- ¹ Georgiev GD, McDougall AC. Skin smears and the bacterial index (BI) in multiple drug therapy leprosy control programs: an unsatisfactory and potentially hazardous state of affairs. *Int J Lepr*, 1988; **56**: 101–103.
- ² de Rijk AJ, Nilson T, Chonde M. Quality control of skin smear services in leprosy programmes: preliminary experience with inter-observer comparison in tourine services. *Lepr Rev*, 1985; **56**: 177–91.
- ³ Bhatia VN. Skin Smear examination in relation to multidrug therapy campaigns. Ind J Lepr, 1987; 59: 75–9.
- ⁴ Kaur S, Kumar B, Ganguly NK, Chahravarti RN. Viability of *M. leprae* under normal and adverse atmospheric conditions. *Lepr India* 1982; 54: 228–32.
- ⁵ Vairamathu Shżri R, Karjat RIHS AJM. Factors affecting accuracy in reporting on smears for AFB. Paper presented before the western regional leprosy workers' conference. Goa, 11/85.
- ⁶ WHO. Chemotherapy of leprosy for control programmes. Report of a WHO Study Group. WHO Technical Report Series 675 (1982).
- ⁷ WHO. WHO Expert Committee on Leprosy. WHO Technical Report Series 768 (1988).

Bibliography

Bhatia VN. Bacterial index: Ridley's vs. Dharmendra's scale. Ind J Lepr, 1987; 59: 80-3.

- Directorate of Health Services, India. Leprosy Division. Report of independent evaluation. New Delhi: National Leprosy Eradication Program, 1986.
- Desikan KV. Correlation of morphology with viability of M. leprae. Lepr India, 1976; 48: 391-7.
- Haidar Abu Ahmed, Belehu A, Stoner G, Touw J, Atlaw T. Selection of sites for slit skin smears. *Lepr Rev*, 1979; **50:** 283–7.
- Hameedullah A, Lal S, Garg BR. Composite skin contact smears in multibacillary leprosy patients. *Lepr India*, 1982; **54**: 605–12.

- Jopling WH, Rees RJW, Ridley DS, Ridley MJ, Samuel NM. The fingers as sites of leprosy bacilli in pre-relapse patients. Lepr Rev, 1979; 50: 289–92.
- Kaur S, Kumar B, Darshan H, Singh S. Choice of skin slit smears for study of bacterial and morphological indices. Lepr Rev, 1980; 52: 540-7.

Kaur S, Kumar B, Darshan H. Choice of sites for study of slit skin smears. Lepr Rev, 1981; 52: 27-33.

- Kumar B, Kaur S, Gupta SK, Rajwanshi A, Darshan H. Acid fast bacilli in lymph node aspirate and smears from ear lobules and fingers in long treated patients. *Ind J Lepr* 1984; **56**: 71–7.
- Kumar B, Kaur S. Selection of sites for slit skin smears in untreated and treated leprosy patients. *Int J Lepr*, 1986; **54**: 540–3.
- Oddinsen O, Nilson T, Humber DP. Viability of *M. leprae*: a comparison of morphological index and fluorescent staining techniques in slit skin smears and *M. leprae* suspensions. *Int J Lepr*, 1986; **54**: 403-8.
- Pariser DM, Caserio RJ, Eaglstein WH. Techniques for diagnosing skin and hair disease. Slit smear. Georg Thieme Verlag, Stuttgart. 1986, 103-7.
- Ridley M, Ridley DS. Stain techniques and the morphology of M. leprae. Lepr Rev, 1971; 42: 88-95.
- Sayer J, Gent R, Jesudasan K. Are bacterial counts on slit skin smears in leprosy affected by preparing slides under field conditions? Lepr Rev, 1987; 58: 271-8.
- Sharma VD, Ramu G, Dutta AK, Katoch K, Ramanathan U. Solid, fragmented and granular index as one of the parameters in drug trials. *Lepr India* 1982; 54: 448–52.
- Wade HW. The examination of skin lesions for bacilli. Int J Lepr, 1963; 31: 240-2.
- Warndorff-van Diepan T. Do the average bacterial and morphological indices reflect the patients' true condition? Int J Lepr 1980; **48**: 441-2.
- WHO. Laboratory techniques for leprosy. WHO (1987); 21-9.