

Validity of the WHO Operational Classification and Value of Other Clinical Signs in the Classification of Leprosy¹

Gift Norman, Geetha Joseph, and Joseph Richard²

ABSTRACT

The objective of this study is to examine the validity of the WHO operational classification using skin smear results as the gold standard and explore the value of additional clinical signs independently and in combination with the WHO classification.

Between 1985 and 2000, 5439 new untreated leprosy patients were registered at the Schieffelin Leprosy Research and Training Center, Karigiri. They were classified according to the Ridley Jopling classification as well as WHO operational classification based on the number of skin lesions.

The sensitivity and specificity of the WHO operational classification tested, using skin smear results as the gold standard, was found to be 88.6% and 86.7% respectively. The Receiver Operator Characteristic (ROC) curve confirms that the best option for sensitivity and specificity is a cut off of 6 and more lesions for MB.

The validity of the number of enlarged nerves and size of the largest skin lesion as independent criteria to classify patients was found to be poor. Addition of three enlarged trunk nerves to the WHO classification improved its sensitivity to 91.4%, while the specificity remained almost unchanged at 85.3%. Addition of the size of the largest skin lesion to the WHO classification reduced its validity considerably.

The study concludes that the WHO recommendation of using six and more lesions for classifying a patient as MB is the best option available at the moment, and calls for further research to identify other clinical criteria that have a better validity and could be easily applied in the field.

RÉSUMÉ

L'objectif de cette étude est d'examiner la validité de la classification opérationnelle de l'OMS, qui utilise l'examen bactérioscopique de frottis de suc dermique comme méthode de référence, et d'explorer l'intérêt d'intégrer des signes cliniques additionnels, indépendamment ou bien en combinaison avec la classification de l'OMS.

Entre 1985 et 2000, 5439 nouveaux patients atteints par la lèpre, encore non traités, furent enregistrés au Centre Schieffelin d'Education et de Recherche sur la Lèpre de Karigiri. Ils furent classés selon la classification de Ridley et Jopling ainsi que selon la classification opérationnelle de l'OMS qui est basée sur le nombre de lésions.

La sensibilité et la spécificité de la classification pratique de l'OMS, testées en utilisant l'examen bactérioscopique de frottis de suc dermique, ont été déterminées être 88,6% et 86,7%, respectivement. La courbe Caractéristique Opérateur Recepteur (ROC) a confirmé que la meilleure option pour la sensibilité et la spécificité est une limite de 6 lésions et plus pour la lèpre multibacillaire (MB).

La validité du nombre de nerfs ayant une taille augmentée ou bien de la taille de la plus grande lésion cutanée comme critère indépendant de classification des patients est faible. L'addition de trois troncs nerveux de taille augmentée à la classification de l'OMS a amélioré sa sensibilité à 91,4%, tandis que la spécificité est restée presque inchangée à 85,3%. L'addition de la taille de la plus grande lésion cutanée à la classification de l'OMS a réduit considérablement sa validité.

¹Received for publication on 16 December 2003. Accepted for publication on 27 May 2004.

²Norman, G., M.D., Senior Specialist and Head, Department of Community Health; Joseph, G. A., M.D., Former Specialist, Department of Community Health; Richard, J., Ph. D., Consultant Biostatistician, Schieffelin Leprosy Research and Training Center, Karigiri, Vellore, India.

Reprint requests to: Gift Norman, Department of Community Health, Schieffelin Leprosy Research and Training Center, Karigiri, Vellore, INDIA; Email: chd@karigiri.org

La conclusion de cette étude est que la classification de l'OMS, qui utilise six lésions ou plus comme critère pour classer un patient comme MB, est la meilleure option pour le moment et appelle à des recherches complémentaires pour identifier d'autres critères cliniques susceptible de présenter une meilleure validité et une application encore plus facile sur le terrain.

RESUMEN

Los objetivos de este estudio fueron (1) examinar la validez de la clasificación operacional de la OMS usando los resultados de los extendidos de linfa cutánea como el estándar de oro, y (2) explorar el valor de los signos clínicos adicionales, usados de manera independientemente y en combinación con la clasificación de la OMS.

Entre 1985 y 2000, 5439 nuevos pacientes con lepra sin tratamiento fueron registrados en el Centro Schieffelin para Investigación y Tratamiento de la Lepra en Karigiri. Los casos se clasificaron de acuerdo a la escala de Ridley-Jopling y también de acuerdo a la clasificación operacional de la OMS basada en el número de lesiones en la piel.

La sensibilidad y especificidad de la clasificación operacional de la OMS utilizando los resultados de las extensiones de linfa cutánea fueron del 88.6% y 86.7%, respectivamente. La curva ROC confirmó que la mejor opción para establecer la sensibilidad y la especificidad es un valor de corte de 6 ó más lesiones para la lepra MB.

La consideración del número de nervios engrosados y el tamaño de la lesión dérmica más grande como criterios independientes para clasificar a los pacientes, arrojó resultados pobres. La adición de 3 troncos nerviosos engrosados a la clasificación de la OMS elevó su sensibilidad hasta el 91.4%, mientras que la especificidad permaneció casi sin cambio a 85.3%. La adición del tamaño de la lesión dérmica más grande a la clasificación de la OMS redujo considerablemente su validez. Este estudio concluye que las recomendaciones de la OMS de usar 6 y más lesiones para clasificar a los pacientes como MB es la mejor opción que se tiene hasta el momento y hace énfasis en la necesidad de identificar otros criterios clínicos que tengan una mejor validez y que puedan ser fácilmente aplicados en los estudios de campo.

The classification of leprosy determines the type and duration of treatment a patient receives. The World Health Organization (WHO) Study Group on the Chemotherapy of Leprosy in 1993 recommended that clinical criteria might be required for classification of the disease where facilities for the bacteriological examination of skin smears are either unreliable or unavailable⁽¹²⁾. The WHO Expert Committee concluded that patients could be classified according to the number of skin lesions into three groups namely, paucibacillary single lesion (PB), paucibacillary leprosy (PB) (2 to 5 skin lesions), multi bacillary leprosy (MB) (6 or more skin lesions)⁽¹¹⁾.

Several studies have been published on the validity of the clinical criteria, such as number of skin lesions or body area affected, using skin smears and/or histopathological diagnoses as the gold standard. Although the criteria for classification (both clinical and bacteriological) vary among the studies, the sensitivity of the clinical classification ranged from 85% to 93% and specificity from 39% to 88%^(2, 3, 4, 7, 9). There is a concern however, that by classi-

fying patients solely on the number of skin lesions, a small but significant number of MB patients may wrongly receive PB treatment and a fairly large number of PB patients would be treated unnecessarily with the MB regimen⁽⁴⁾. Therefore, there is good reason to identify and test other simple clinical signs that would complement and improve the validity of the present clinical classification. This paper reviews the validity of the WHO operational classification, and explores whether additional clinical signs that are simple and could be applied in the field would improve the validity of this classification using skin smear examination as the gold standard.

MATERIALS AND METHODS

The Schieffelin Leprosy Research and Training Center (SLR & TC) Karigiri, has been implementing the National Leprosy Eradication Program (NLEP) in the Gudiyatham Taluk since 1962 (Pop. in 1961: 400,000). Leprosy services to the area were provided through mobile clinics to different villages in the control area, supported by leprosy paramedical workers and non-

medical supervisors. All persons with symptoms and signs of leprosy were referred to a medical officer who made the diagnosis based on the presence of the cardinal signs. Patients were classified based on clinical features that would correlate with the six groups of the Ridley Jopling classification⁽⁸⁾. Subsequently, they were also classified into MB and PB based on the number of skin lesions as recommended by WHO⁽¹¹⁾. Patients had a body chart that had a record of the number, distribution, the size of the largest skin lesion, and the peripheral nerves that were enlarged.

Slit-skin smears were part of the routine diagnostic procedure. Standardized methods were used for the preparation, staining, reading, and grading of smears throughout the reference period. Skin and/or nerve biopsies were done only for confirmation of diagnosis or for research purposes. A review of all new previously untreated leprosy patients registered from the control area from 1985 to 2000 was used for this study. The average Bacteriological Index (BI) is used as the gold standard. The sensitivity, specificity, and predictive values were calculated using standard formulae⁽¹⁾. The Receiver Operator Characteristic (ROC) curve is drawn by plotting the sensitivity on the Y axis and 1-specificity or false positives on the X axis. The point on the curve, which is nearest to 100% sensitivity, determines the best cut off point.

RESULTS

A total of 5439 new untreated leprosy patients were registered between 1985 and 2000. Information on the number of skin lesions and skin smear results was available only for 5165 (95.0%) patients. Those with numerous skin patches were included in the group of more than 9 lesions. Persons with no skin lesions, but having other signs of multibacillary leprosy, such as infiltration of the skin, thickened earlobes or nodules, were also included in this group.

Table 1 shows that the sensitivity of the current WHO operational classification that classifies a patient with 6 and more skin lesions as MB is 88.6%, while the specificity is 86.7%. The positive predictive value (PPV) is 39.1% and the negative predictive value is 98.7%. The Receiver Operator Characteristic (ROC) curve (The Figure) il-

TABLE 1. *Validity of the WHO operational classification using skin smears as the gold standard.*

No. of skin lesions	Slit skin smear examination		
	POS	NEG	Total
>5 lesions	403	626	1029
1-5 lesions	52	4084	4136
Total	455	4710	5165

Sensitivity: $403 / 455 \times 100 = 88.6\%$

Specificity: $4084 / 4710 \times 100 = 86.7\%$

False positive rate: $626 / 4710 \times 100 = 13.3\%$

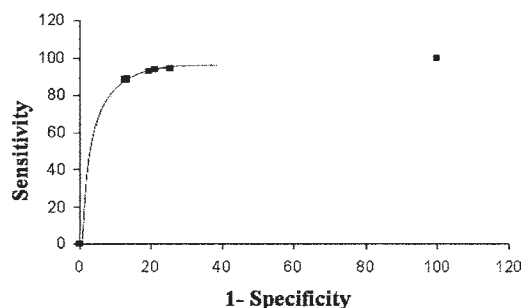
False Negative rate: $52 / 455 \times 100 = 11.4\%$

Positive Predictive value: $403 / 1029 \times 100 = 39.1\%$

Negative Predictive value: $4084 / 4136 \times 100 = 98.7\%$

lustrates that the best option is at the cut off for 6 and more lesions for classifying a patient as MB (which is the point nearest to 100% sensitivity).

There were 52 smear positive patients who presented with 1 to 5 skin lesions and would have wrongly been classified as PB according to the WHO classification, a false negativity rate of 11.4%. On further analysis, of these 52 patients, 7 (13.5%) had a high BI of 3.00+ or more. Among the 4710 patients who were skin smear negative (Gold Standard), 626 patients were found to have 6 and more skin lesions and hence would have been classified as MB, a false positivity rate of 13.3% (Table 1).



THE FIGURE. The points on the ROC Curve represent sensitivity versus 1-specificity for different cut off points (in terms of the number of skin lesions) used for determining MB cases; the points plotted represent 8 different cut off points ranging from 2 to 9 skin lesions. The point on the curve closest to the point where sensitivity is 100% and 1-specificity is 0 determines the best cut off point, which in this case is 6 lesions.

TABLE 2. *Validity of using size of largest skin lesion or enlarged trunk nerves as independent criteria using skin smear as the gold standard.*

Clinical criteria		Total no.	SS + ve with positive clinical criteria (TP)	SS - ve with positive clinical criteria (FP)	SS + ve with negative clinical criteria (FN)	SS - ve with negative clinical criteria (TN)	Sensitivity	Specificity
Largest skin lesion	≥5 cms	652	9	643	8	1630	52.9	71.7
	≥8 cms	267	4	263	13	2010	23.5	88.4
No. of enlarged trunk nerves	1 nerve	946	282	664	66	2638	81.0	79.9
	2 nerves	511	240	271	108	3031	68.9	91.8
	5 nerves	70	46	24	302	3278	13.2	99.2

* SS = Skin smear TP—True positives FP—False positives TN—True negatives FN—False negatives

The validity of independent additional signs such as size of the largest skin lesion and the number of enlarged trunk nerves to classify leprosy was tested. The findings are summarized in Table 2. The number of skin lesions, the size of the largest lesion and the skin smear results were available for 2290 patients. Using the size of the largest skin lesion of 5 centimeters (cm) or more as an independent criterion for classification as MB, the sensitivity was found to be fairly low at 52.9% and a specificity of 71.7%, the sensitivity decreasing further as the size of the lesion increased (Table 2). On combining the size of the largest skin lesion (5 cm or more as MB) to the WHO operational classification, the sensitivity improved only slightly to 58.8%, while the specificity remained almost the same at 70.4%. When the size of the largest skin lesion is more than 8 cm, the sensitivity decreased considerably to 35.3% while the specificity increased to 86.4% (Table 3).

The validity of using the number of enlarged trunk nerves as an independent criterion for classification as MB was tested (Table 2). The number of skin lesions, the number of enlarged nerves and the skin smear results were available for 3650 patients. The sensitivity was found to be 81.0% and the specificity was 79.9% for one or more enlarged nerves. It is observed that the sensitivity steadily decreases as the number of enlarged nerves increased, while on the other hand, the specificity increased considerably to 91.8% when two or more nerves are enlarged, further increasing to

99.2% when five or more nerves are enlarged.

On adding the number of enlarged nerves to the WHO operational classification, a good balance of sensitivity and specificity was observed when enlargement of three nerves or more were added to the WHO classification, with a sensitivity of 91.4% and a specificity of 85.3%. This gives a fairly low false positivity rate of 8.6%, but a false negative rate of 14.7%. The positive predictive value is 39.5% and the negative predictive value is 98.9% (Table 4).

DISCUSSION

Using skin smear examination as the gold standard has its limitations. It has been reported that using skin smear alone may not be as sensitive as using skin smears in combination with skin and/or nerve biopsies (^{4,7}). The reliability of skin smears has been reported to be low (^{6,10}). However, at laboratories with good standardization of techniques

TABLE 3. *Validity of adding size of the largest skin lesion to the WHO operational classification standard.*

More than 5 skin lesions and size of largest lesion (cms) among those with 1-5 skin lesions	Sensitivity	Specificity
>2 cm	82.4	24.5
>5 cms	58.8	70.4
>8 cms	35.3	86.4
>10 cms	29.4	91.2

TABLE 4. *Validity of adding the number of enlarged nerves to the WHO operational classification.*

More than 5 skin lesions and no. of nerves involved among those with 1 to 5 skin lesions	Sensitivity	Specificity
1	93.4	74.9
2	92.5	83.0
3	91.4	85.3
4	90.2	85.7
5	89.1	86.0

and quality control, skin smears could be considered to have a fairly high reliability.

Two important points have to be considered while discussing validity of test criteria such as classification based on clinical criteria into MB/PB leprosy. One is the ability of the criteria to correctly identify the MB patients (true positives or sensitivity) and thereby minimize the number who would be wrongly classified as PB, while in fact they are MB patients (false negatives). Wrongly classifying a MB patient as PB is a cause for concern due to the increased risk of relapse and reactions in MB patients, as well as the inadequate treatment they would receive.

The second issue is the ability of the criteria to be able to correctly identify those who are not MB (true negatives or specificity) and consequently minimize the number of PB patients who would be diagnosed as MB (false positives). The findings in our analysis compares well with similar studies that have reported a sensitivity of 85.0% and 89.0% and a specificity of 81.0% and 88.0%, respectively (3, 4). Croft, *et al.* and Dasananjali, *et al.* have reported similar results of 11% and 12% as "missed" (false negative) MB cases (4, 5).

With the addition of more than two enlarged nerve trunks to the WHO operational classification increased the sensitivity to 91.4% without altering the specificity much (85.3%). Consequently, the number of false negatives and false positives are fairly low, bringing down the number of MB patients who would be under treated.

Considering that nerve enlargement is one of the cardinal signs for diagnosis of leprosy, its inclusion in the classification should not be very difficult. However, ap-

plication under field conditions for use by general health workers might be difficult, since identifying nerve enlargement per se is subjective, with a fairly high inter and intra observer variation. The WHO recommended that a patient with enlargement of more than one nerve trunk could be classified as MB (10). It was observed that the sensitivity of using such a criterion independently was 68.9%, which is rather low, while the specificity is fairly high at 91.8%.

The addition of the size of the largest skin lesion to the WHO operational classification showed considerable decrease in the sensitivity and the specificity. However, it should be noted that the among the smear positive patients, only a small proportion had a record of the size of the largest skin lesion (4.9%). Since the measurement of the largest skin lesion is fairly simple and could be applied in the field, it may be worthwhile to recheck its validity using another data set.

The classification of leprosy using clinical criteria has its benefits and limitations. While it is easy and simple to apply in the field by general health workers, there is always the risk of a certain proportion of patients being under or over treated. Based on the WHO operational classification this study found that about 11% of the patients would be under-treated, while 13% would be over-treated. The validity of using number of enlarged trunk nerves and the size of the largest skin lesion as independent criteria was low. Addition of enlarged trunk nerves to the operational classification improved its sensitivity with no significant change to the specificity; on the contrary, addition of the size of the largest skin lesion to the operational classification lowered its validity considerably.

It appears from this study, that the WHO recommendation of using six lesions for classifying a patient as MB is the best option available at the moment and calls for further research to identify other clinical criteria that have a better validity and can be applied in the field.

Acknowledgment. The authors wish to thank Mr. P. Samuel, Leprosy Supervisor, Mr. Raja Samuel, Programmer, and Mr. P. Dorairaj Medical Records Technician for all the help rendered in the collation and analysis of the data. The authors appreciate the Department of Laboratories and the smear technicians for their efforts in maintaining high standards in examination of skin smears.

REFERENCES

1. BEAGLEHOLE, R., BONITA, R., and KJELLSTROM, T. *Basic Epidemiology*. Geneva: World Health Organization, 1993. pp. 94–96.
2. BECX-BLEUMINK, M. Allocation of patients to paucibacillary or multibacillary drug regimens for the treatment of leprosy—a comparison of methods based mainly on skin smears as opposed to clinical methods—alternative clinical methods for classification of patients. *Int. J. Lepr. Other Mycobact. Dis.* **59** (1991) 292–303.
3. BUHRER-SEKULA, S., SARNO, E. N., OSKAM, L., KOOP, S., WICKERS, I., NERY, J. A., VIERA, L. M., DE MATOS, H. J., FABER, W. R., and KLASTAR, P. R. Use of ML dipstick as a tool to classify leprosy patients. *Int. J. Lepr. Other Mycobact. Dis.* **68** (2001) 456–463.
4. CROFT, R. P., SMITH, W. C. S., NICHOLLS, P., and RICHARDUS, J. H. Sensitivity and specificity of methods of classification of leprosy without use of skin smear examination. *Int. J. Lepr. Other Mycobact. Dis.* **66** (1998) 445–450.
5. DASANANJALI, K., SCHREUDER, P. A., and PIRAYAVARAPORN, C. A study on the effectiveness and the safety of the WHO/MDT regimen in the Northeast of Thailand: a prospective study, 1984–1996. *Int. J. Lepr. Other Mycobact. Dis.* **65** (1997) 28–36.
6. GEORGIEV, G. D., and MCDUGALL, A. C. Skin smears and the bacterial index (BI) in multiple drug therapy leprosy control program: an unsatisfactory and potentially hazardous state of affairs. *Int. J. Lepr. Other Mycobact. Dis.* **56** (1988) 101–103.
7. GROENEN, G., SAHA, N. G., RASHID, M. A., HAMID, M. A., and PATTYN, S. R. Classification of leprosy cases under field conditions in Bangladesh. I. Usefulness of skin smear examinations. *Lepr. Rev* **66** (1995) 126–133.
8. RIDLEY, D. S., and JOPLING, W. H. Classification of leprosy according to immunity. A five group system. *Int. J. Lepr. Other Mycobact. Dis.* **34** (1966) 255–273.
9. VAN BRAKEL, W. H., DE SOLDENHOFF, R., and MCDUGALL, A. C. The allocation of leprosy patients into paucibacillary or multibacillary groups for multidrug therapy, taking into account the number of body areas affected by skin, or skin and nerve lesions. *Lepr. Rev.* **63** (1992) 231–246.
10. WHO EXPERT COMMITTEE ON LEPROSY. Sixth Report. Geneva: World Health Organization, 1988. Tech. Rep. Ser. 768.
11. WHO EXPERT COMMITTEE ON LEPROSY. Seventh Report. Geneva: World Health Organization, 1998. Tech. Rep. Ser. 874
12. WHO STUDY GROUP. Chemotherapy of leprosy for control programs. Geneva: World Health Organization, 1994. Tech. Rep. Ser. 847.