

## Confiabilidad y concordancia de dos escalas de lectura de baciloscopias para clasificación y seguimiento del tratamiento con múltiples medicamentos de los pacientes con lepra

Claudia Lucía Colorado, Guillermo Sánchez, Martha Inírida Guerrero, Clara Inés León

Centro Dermatológico Federico Lleras Acosta, E.S.E., Bogotá, D.C., Colombia

**Introducción.** El diagnóstico de la lepra es clínico pero requiere métodos de clasificación para establecer el tratamiento y el pronóstico de los pacientes, acordes con su carga bacilar; ésta se detecta en extendidos de piel y se establece mediante la escala logarítmica de Ridley. En Colombia se emplea una escala semicuantitativa.

**Objetivo.** Establecer la reproducibilidad entre observadores para las dos escalas y establecer el nivel de correlación de la concordancia entre el índice bacilar obtenido con la escala-colombiana y con la de Ridley, para evaluar cuánto se pueden intercambiar.

**Materiales y métodos.** Se estandarizó la lectura de las baciloscopias por dos lectores, con posterior evaluación del acuerdo entre observadores en un ensayo ciego. Cada lector cuantificó la carga bacilar de las muestras, usando la escala colombiana y la internacional. El grado de concordancia entre observadores se evaluó con el coeficiente kappa ponderado. El nivel de correlación de la concordancia entre las mediciones del índice bacilar, se estableció con el coeficiente de Lin.

**Resultados.** El coeficiente kappa ponderado entre observadores fue de 0,83, y de 0,85 para las escalas colombiana y la de Ridley, respectivamente. El coeficiente de Lin fue de 0,96 para la evaluación de correlación de la concordancia de los índices bacilares obtenidos con ambas escalas

**Conclusiones.** La concordancia entre observadores obtenida para ambas escalas fue excelente, al igual que la correlación de la concordancia de los índices bacilares calculados con los dos métodos. Con los puntos de corte establecidos, se obtuvo un nivel de concordancia excelente, lo que garantiza que las escalas son intercambiables a la hora de establecer si la carga bacilar es alta o baja.

**Palabras clave:** lepra/clasificación, terapéutica, escalas, Colombia.

### Reliability and agreement of two smear reading scales for classification and monitoring of multidrug therapy in leprosy patients

**Introduction.** After the clinical diagnosis of leprosy, classification methods are necessary to define a treatment and prognosis of patients consistent with bacterial load. Bacteria are detected in skin smear, and bacterial load typically is established by the internationally used Ridley's logarithmic scale. However, in Colombia an alternative semiquantitative scale is used.

**Objective.** The interobserver reproducibility was established for the Ridley and Colombia scales, and the level of correlation-matching was identified between the bacillary indices obtained in order to assess the degree of interchangeability.

**Materials and methods.** Standardization was attained by a reading of the smears by 2 readers with subsequent, blinded evaluation of inter-observer agreement. Each reader quantified the bacterial load of for each sample (n=325) using the Colombian and the Ridley scales. The degree of interobserver agreement was assessed with weighted kappa coefficient. The level of correlation and agreement between the measurements of the bacillary index was established with coefficient of Lin.

**Results.** The interobserver weighted kappa coefficient was 0.83 for the Colombia scale and 0.85 for the Ridley scale. The Lin coefficient was 0.96 for the correlation-matching of bacillary indexes.

**Conclusions.** Interobserver agreement obtained for both scales was excellent as the correlation-matching bacillary indices determined with both methods. With the cut-off points yielded a good level of agreement, ensuring interchangeability between the scales defining the high or low bacterial load.

**Key words:** leprosy/classification, therapeutics, scales, Colombia.

La lepra continúa siendo un problema de salud pública en Colombia, mayor en algunas regiones del país, y aunque se consideró una enfermedad eliminada, en los últimos años ha aparecido un número similar de casos nuevos, lo cual indica que la transmisión de la enfermedad aún se mantiene activa (1,2).

El diagnóstico de lepra es eminentemente clínico, pero requiere métodos complementarios de histopatología y bacteriología para clasificar la enfermedad, definir un adecuado tratamiento y tener un mejor pronóstico. La Organización Mundial de la Salud (OMS) recomendó en 1988 la clasificación bacteriológica operativa (3), en la cual todo paciente con bacilos ácido-alcohol resistentes en el extendido de piel debería ser tratado como paciente multibacilar, mientras que aquel sin dichos bacilos debería tratarse como paucibacilar.

Se acepta internacionalmente que en los pacientes que no han iniciado el tratamiento se deben tomar muestras de varios sitios, como mínimo del lóbulo de las orejas y dos lesiones activas, y que los mismos sitios deben utilizarse para el seguimiento del tratamiento (4). La muestra de moco nasal no se recomienda, porque su toma es molesta, las micobacterias no tuberculosas que no son patógenas se encuentran, generalmente, en la nariz de sujetos saludables y, durante el curso del tratamiento, *Mycobacterium leprae* puede desaparecer de la mucosa nasal antes que de las lesiones cutáneas (5).

Para determinar la carga bacilar en las muestras de piel, se usa la escala logarítmica internacional de Ridley (6), que va de 0 a 6+, según el número promedio de bacilos ácido-alcohol resistentes presentes por campo microscópico observado, cuyo denominador varía según la carga bacilar, ya que, en las muestras en las que es alta, el número de campos observados es menor, y en aquellas con baja carga bacilar, es necesario revisar un número mayor de campos. Además, siempre se calcula el índice bacilar que corresponde a la suma del número de cruces de cada una de las muestras, dividida por el número de muestras examinadas (6).

Correspondencia:

Martha Inírida Guerrero, Oficina de Docencia e Investigación, Centro Dermatológico Federico Lleras Acosta ESE, Avenida 1 N° 13A-61, Bogotá, D.C. Colombia.

Teléfono: 242 8160, extensión 145; fax: 242 8160, extensión 141 marthainiridag@yahoo.com

Recibido: 25/11/10; aceptado:07/06/11

En Colombia se determina el índice bacilar utilizando cinco muestras: dos de los lóbulos de las orejas y dos de los codos, y una de moco nasal (7). En los casos que presentan lesiones, las muestras de los codos son reemplazadas por las de estas lesiones. Para cuantificar, leer e informar los índices bacilares, se utiliza una escala semicuantitativa (7) diferente a la usada en el resto del mundo (6), por lo cual las categorías no son equiparables. La escala usada en Colombia —que para efectos de este documento se denominará “escala colombiana”— cuantifica la presencia de bacilos en cada una de las cinco muestras examinadas, incluyendo el moco nasal, entre 0 y 3+, por lo cual los índices bacilares varían entre 0 y 3, mientras que la escala de Ridley las cuantifica entre 0 y 6+, y sus índices bacilares varían entre 0 y 6.

Frente a la dificultad encontrada al tratar de equiparar la escala de lectura de baciloscopias utilizada en Colombia y la internacional de Ridley, empleadas para la clasificación y el seguimiento del tratamiento de los pacientes con lepra, el presente estudio buscó establecer la concordancia entre observadores en las lecturas de las dos escalas, identificó el nivel de correlación de la concordancia existente entre el índice bacilar obtenido con las dos escalas y, finalmente, evaluó la posibilidad existente de intercambio entre las dos metodologías, todo esto con el fin de poder confrontar la información de los pacientes colombianos con los del resto del mundo, en cuanto a carga bacilar.

### Materiales y métodos

Se llevó a cabo un estudio analítico para evaluar la concordancia entre observadores y la posibilidad de intercambiar las dos escalas empleadas para la clasificación y el seguimiento de la poliquimioterapia en pacientes con lepra. Según la Resolución 8430 de 1993 del Ministerio de Salud, el presente estudio se clasifica como una investigación sin riesgo, cuyo protocolo fue avalado por el Comité Científico institucional.

Con base en la fórmula propuesta por Machin (8) para los estudios de concordancia, asumiendo un error de tipo I de 5 %, un error de tipo II de 20 %, un coeficiente kappa de 0,6 para la hipótesis nula y un coeficiente kappa mayor a 0,9 para la hipótesis alternativa, se estimó un tamaño de muestra de 65 observaciones.

Esta muestra se completó incluyendo todas las láminas con muestras de piel y moco nasal de 65 pacientes con lepra, pertenecientes al banco

biológico del laboratorio del Centro Dermatológico Federico Lleras Acosta, procesadas en septiembre de 2009, todas las cuales habían sido tomadas por una sola persona y coloreadas mediante la coloración estándar de Ziehl-Neelsen en calor (7) por un solo operario encargado de este procedimiento.

De forma preliminar, se estandarizaron las lecturas de las baciloscopias en un ensayo ciego, con posterior evaluación de la concordancia entre observadores de las lecturas de los índices bacilares. Para tal efecto, se eligieron dos lectores microbiólogos, con 10 años de experiencia en lectura de baciloscopias para lepra y con el reconocimiento de los estándares de calidad de laboratorio.

Durante el estudio, cada uno de los lectores por separado cuantificó la carga bacilar de las muestras incluidas, usando tanto la escala colombiana como la escala internacional de Ridley. Con una cuadrícula de 10 x 10 diseñada para registrar el número de bacilos observados por campo, se efectuaron las lecturas de cada una de las cinco muestras de las 65 láminas, para un total de 325 muestras y, posteriormente, utilizando las dos escalas, se calculó el número de cruces para cada una de las muestras y, finalmente, los índices bacilares de las 65 láminas incluidas en el estudio.

Los resultados de las lecturas registrados en una cuadrícula por cada muestra y el resultado del índice bacilar respectivo para la escala colombiana y la de Ridley, fueron entregados por los lectores al responsable del análisis de los resultados.

Para conocer el grado de acuerdo existente en las lecturas adelantadas para cada una de las escalas, entre los dos observadores, se empleó el coeficiente estadístico kappa ponderado (9), con su respectivo intervalo de confianza del 95 %; para la clasificación cualitativa del nivel de acuerdo, se utilizó la escala de interpretación propuesta por Landis y Koch (10).

El grado de correlación de la concordancia entre las mediciones del índice bacilar, obtenidas por los dos observadores y con las dos escalas empleadas, se estableció a partir del coeficiente de correlación de la concordancia propuesto por Lin (11).

Las dos escalas evaluadas buscan establecer la carga bacilar. Sin embargo, no fue posible compararlas bajo un mismo parámetro, dado que su cálculo proviene de diferentes niveles de medición (semicuantitativa *versus* logarítmica), aunque existía la posibilidad de separar los

pacientes en portadores de alta y de baja carga bacilar. Para la escala de Ridley, se tomó un punto de corte de tres, en el que los pacientes con valores de índices bacilares iguales o mayores a este se consideraron como de alta carga bacilar (12). Asumiendo que la escala internacional se comporta como una prueba estándar, se empleó el método de curvas ROC (*Receiver-Operating Curve*) (13), para establecer el punto de corte ideal (porcentaje de sujetos correctamente clasificados), en el cual la escala Ridley y la colombiana se pueden intercambiar en cuanto a la carga bacilar alta o baja. Una vez establecido el punto de corte para la escala colombiana, se evaluó la concordancia de los resultados obtenidos empleando el coeficiente kappa (9).

## Resultados

### **Concordancia entre observadores de la escala colombiana**

La distribución del acuerdo entre observadores de los resultados de las lecturas de baciloscopias con la escala colombiana, se presenta en el cuadro 1. A partir de estos resultados, se calculó el nivel de concordancia más allá del azar. El valor global del coeficiente kappa ponderado fue de 0,83, para un nivel de concordancia considerado como excelente, según la clasificación de Landis y Koch. Se calcularon los niveles de concordancia específicos por sitio de muestra con los resultados del coeficiente kappa ponderado, que oscilaron entre 0,71 y 0,89. Los detalles se presentan en el cuadro 2.

### **Concordancia entre observadores de la escala internacional de Ridley**

La distribución del acuerdo entre observadores de los resultados de las lecturas de baciloscopias con la escala internacional, se presenta en el cuadro 3. Se obtuvo un coeficiente kappa ponderado global, de 0,85, para un nivel de concordancia considerado como excelente, según la clasificación

**Cuadro 1.** Concordancia de la lectura de baciloscopias, escala colombiana

		Observador 2				Total
		Lectura	0 +	+	++	
Observador 1	0 +	209	3	0	0	212
	+	28	28	1	0	57
	++	0	3	14	4	21
	+++	0	0	1	21	22
	Total	237	34	16	25	312

**Cuadro 2.** Concordancia entre observadores, escala colombiana

Tipo de muestra	Acuerdo absoluto (%)	Kappa ponderado	IC <sub>95%</sub>		p	Nivel de acuerdo Landis y Koch
Concordancia global	95,7	0,83	0,83	0,84	0,00	Excelente
Moco	96,6	-0,01	0,05	0,02	0,55	No significativo
Muestra de oreja derecha	91,8	0,71	0,70	0,80	0,00	Bueno
Muestra de oreja izquierda	94,5	0,80	0,74	0,82	0,00	Bueno
Muestra de lesión derecha	96,1	0,88	0,78	0,91	0,00	Excelente
Muestra de lesión izquierda	96,7	0,89	0,84	0,94	0,00	Excelente
Muestra de otros sitios	96,7	0,89	0,82	0,94	0,00	Excelente

**Cuadro 3.** Concordancia de la lectura de baciloscopias, escala internacional de Ridley

Lectura	Observador 2							Total
	0 +	+	++	+++	++++	+++++	++++++	
Observador 1								
0 +	210	0	2	0	0	0	0	212
+	23	4	1	0	0	0	0	28
++	4	7	17	1	0	0	0	29
+++	0	0	3	14	3	0	0	20
++++	0	0	0	1	4	1	4	10
+++++	0	0	0	0	0	0	0	0
++++++	0	0	0	0	0	0	13	13
Total	237	11	23	16	7	1	17	312

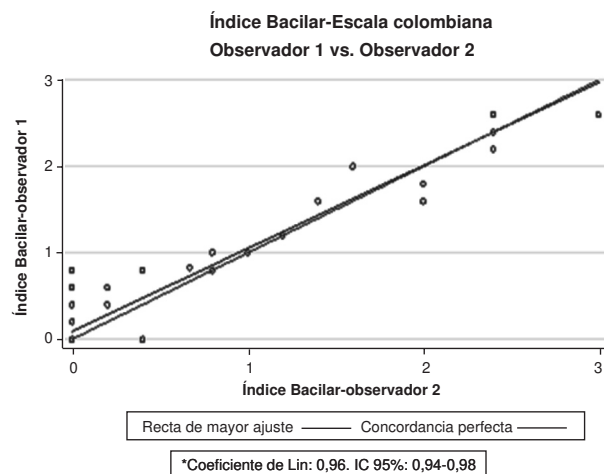
**Cuadro 4.** Concordancia entre observadores, escala internacional de Ridley

Tipo de muestra	Acuerdo absoluto (%)	Kappa ponderado	IC <sub>95%</sub>		p	Nivel de acuerdo Landis y Koch
Comparación general	96,7	0,85	0,79	0,86	0,00	Excelente
Moco	97,5	-0,01	0,19	0,17	0,55	No significativo
Muestra de oreja derecha	93,7	0,77	0,70	0,78	0,00	Bueno
Muestra de oreja izquierda	95,7	0,84	0,79	0,87	0,00	Excelente
Muestra de lesión derecha	96,4	0,86	0,77	0,85	0,00	Excelente
Muestra de lesión izquierda	97,1	0,90	0,83	0,84	0,00	Excelente
Muestra de otros sitios	100	1,00	1,00	1,00	0,00	Excelente

de Landis y Koch. Los niveles de concordancia por sitio de muestra oscilaron entre 0,77 y 0,90. En las muestras de moco, los resultados no fueron significativos desde el punto de vista estadístico. Los detalles se presentan en el cuadro 4.

#### **Análisis de las lecturas del índice bacilar con la escala colombiana**

Desde el punto de vista descriptivo, las distribuciones de los resultados del índice bacilar obtenidos por los dos evaluadores no presentaron diferencias clínicamente importantes. El índice de correlación de la concordancia de Lin para la escala colombiana, fue de 0,96, que puede interpretarse como un nivel de acuerdo casi perfecto (14). En la figura 1 se pueden observar el nivel de acuerdo y la correlación existente entre las observaciones de los dos evaluadores.

**Figura 1.** Coeficiente de correlación de la concordancia de Lin: índice bacilar, escala colombiana.

**Análisis de las lecturas del índice bacilar según la escala internacional de Ridley**

Desde el punto de vista descriptivo, las distribuciones de los resultados del índice bacilar obtenidos por los dos evaluadores no presentaron diferencias clínicamente importantes. El índice de correlación de la concordancia de Lin para la escala internacional, fue de 0,96, que puede interpretarse como un nivel de acuerdo casi perfecto (14). En la figura 2 se pueden observar el nivel de acuerdo y la correlación existente entre las observaciones de los dos evaluadores.

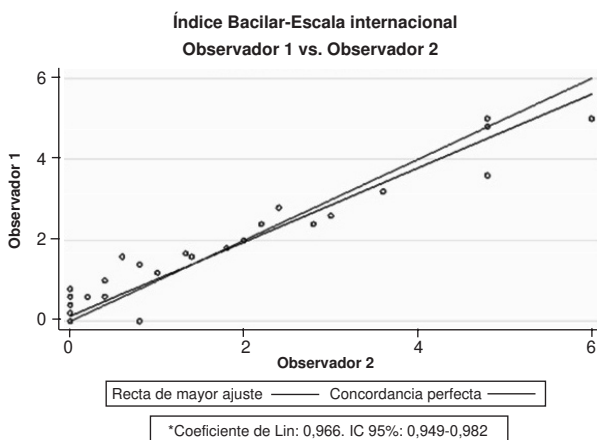
**Comparación de las escalas de lectura colombiana e internacional de Ridley**

Con base en los resultados de las curvas ROC, el punto de corte en el cual se consigue el mayor nivel de sensibilidad y especificidad para la escala colombiana, fue de dos. Es decir que un índice bacilar mayor o igual a dos obtenido con la escala colombiana, indica un nivel de alta carga bacilar. Los resultados se presentan en la figura 3.

La concordancia entre los resultados de alta y baja carga bacilar entre las dos escalas (Ridley y colombiana) a partir del punto de corte definido, fue excelente, con un coeficiente kappa de 0,9992 (IC<sub>95%</sub>: 0,9972-1,000).

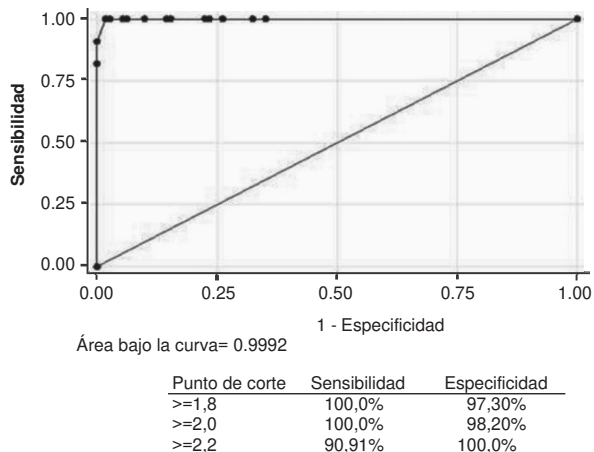
**Discusión**

Se sabe que, aun en condiciones ideales, pueden encontrarse diferencias en las lecturas de las baciloscopias, principalmente debidas a la distribución irregular de los bacilos en el extendido y a la imposibilidad de leer exactamente los mismos 100 campos leídos por el primer observador,



**Figura 2.** Coeficiente de correlación de la concordancia de Lin: índice bacilar, escala de Ridley

**Curva Roc - Índice Bacilar- Escalar R&J Vs - Escala colombiana**



**Figura 3.** Curva ROC y punto de corte de la escala colombiana

especialmente en las muestras con escaso número de bacilos en las que se puede encontrar una alta proporción de falsos negativos.

En el presente estudio se eliminó la variabilidad en los aspectos técnicos, por lo cual, al realizar la comparación general del acuerdo entre observadores, fue posible obtener resultados del coeficiente kappa ponderado de 0,83 y 0,85 para la escala colombiana y la internacional de Ridley, respectivamente, los cuales obtuvieron clasificación de bueno a excelente, según la de Landis y Koch. Comparativamente, estos resultados fueron superiores a los pocos reportes existentes en la literatura científica mundial (15).

Por otra parte, el alto nivel de acuerdo podría obedecer a la utilización sistemática de la cuadrícula de registro de los cien campos recomendada que, unida con una buena coloración de Ziehl-Neelsen, permite una fácil diferenciación de los bacilos por parte de los observadores, quienes siguen el protocolo estándar de lectura e informe de estas baciloscopias (16).

El análisis de concordancia discriminando por tipo de muestra, fue clasificado entre bueno y excelente, salvo en las muestras de moco nasal, en las cuales, a pesar de obtener un alto acuerdo absoluto en las observaciones (97,5 %), el acuerdo esperado por azar no mostró una diferencia significativa, probablemente debido al escaso número de muestras positivas para esta medición.

Se observó que alrededor de 50 % de las discrepancias encontradas en las lecturas de los



dos observadores utilizando la escala colombiana o la de Ridley, se concentraron en las muestras de índices bacilares bajos, es decir, inferiores a 1,0, en las cuales existe una mayor probabilidad de que los pocos bacilos presentes no sean coincidentalmente observados por los dos lectores (cuadros 1 y 3).

En Colombia, se ha requerido la comparación del comportamiento de los casos de lepra del Centro Dermatológico Federico Lleras Acosta, al igual que los del resto del país, con los casos del mundo. Esto no ha sido posible debido a que desde hace mucho tiempo se viene utilizando la escala semicuantitativa, la cual no se usa en el resto del mundo; por lo tanto, es importante resaltar que, aunque las escalas de medición y los supuestos matemáticos que sustentan la medición con las dos escalas son diferentes, en este trabajo fue posible establecer una alta correlación entre los puntos de corte que definen el estado de alta y baja carga bacilar con las dos escalas.

Este hallazgo es importante en la medida en que permite intercambiar las escalas para este propósito, abriendo una puerta para la interpretación y el análisis de los resultados obtenidos entre una y otra escala. Por lo tanto, podemos extrapolar los hallazgos de la literatura científica mundial con respecto al riesgo y el pronóstico de recaídas (17,18), reacciones (19) y desarrollo de discapacidades (20), reportada entre pacientes con altas cargas bacilares determinadas con la escala de Ridley.

Sin embargo, es claro que en Colombia se requiere empezar a usar la escala internacional de Ridley, hecho que vemos posible ya que, mediante el presente ejercicio, se pudo evidenciar que se puede obtener un alto grado de acuerdo entre lectores entrenados y porque las dos escalas no son homologables entre sí, por tratarse la colombiana de una escala semicuantitativa mientras que la de Ridley es una escala logarítmica.

El empezar a utilizar en Colombia la escala internacional de Ridley, permitirá contrastar la situación epidemiológica y los impactos logrados por el programa de lepra, con los del resto del mundo, en todos los aspectos en los que se requiera conocer los índices bacilares.

Teniendo en cuenta que en nuestro país alrededor de 60 a 70 % de los casos de lepra son pacientes multibacilares y que se cuenta con una red de bacteriología de lepra, creemos que la baciloscopia

se debe continuar utilizando para hacer una buena clasificación de los pacientes; además, se puede continuar utilizando en el seguimiento del tratamiento de la enfermedad, así como en la detección de recaídas, sin olvidar que se debe pensar en herramientas más sensibles para el seguimiento de los pacientes (21,22).

La OMS propuso utilizar una clasificación clínica operativa que tiene en cuenta el número de lesiones presentes, aduciendo que la sensibilidad de los extendidos de piel es baja, como se ha reportado en algunos estudios (23,24); además, haciendo eco de otros estudios que han demostrado que el uso de los extendidos y biopsias de nervio pueden superar la sensibilidad de los extendidos de piel (25,26). Sin embargo, Norman, *et al.*, consideran que los laboratorios con técnicas bien estandarizadas y con un control de calidad funcional, hacen que los extendidos sean confiables (21).

Dicha clasificación clínica operativa tiene la desventaja de necesitar personal con mucha experiencia clínica para realizarla correctamente y que los pacientes reciban el tratamiento adecuado. Esto no es fácil, ya que médicos que no son especialistas con esta habilidad, no se encuentran en todas las zonas del país. Por lo tanto, acogerse a esta recomendación de la OMS no sería apropiado en Colombia, donde se cuenta con una red de bacteriología de la lepra operante que tiene metodologías estandarizadas y control de calidad periódico sobre los laboratorios que realizan este procedimiento, lo cual es una debilidad en otros programas de control (22).

En conclusión, podemos decir que la comparación entre observadores de las lecturas para las dos escalas mostró una buena concordancia y, dado que estas dos escalas no son homologables, es necesario implementar la escala de Ridley en Colombia para poder comparar los resultados de la clasificación y del seguimiento de nuestros pacientes durante el tratamiento, de la efectividad de los esquemas de tratamiento actuales y de los nuevos, con los estudios realizados en otras partes del mundo, al igual que comparar nuestros datos epidemiológicos y los indicadores del impacto del programa de control.

Los resultados del presente estudio permiten establecer el grado de confiabilidad de dos escalas utilizadas para la lectura de baciloscopias, la de Ridley y la nacional, ambas de utilidad para efectos de clasificación y seguimiento de la poliquimioterapia de los pacientes con lepra. Los

resultados son alentadores en la medida en que se pudieron corroborar altos niveles de concordancia entre los observadores con las dos escalas, y un alto nivel de correlación de la concordancia entre los índices bacilares obtenidos por dos diferentes observadores.

### Agradecimientos

Los autores agradecen a Gloria Iris Pinilla y Sandra López, por su contribución en la parte operativa de toma, coloración y archivo de las láminas del banco biológico del laboratorio del Centro Dermatológico Federico Lleras Acosta, E.S.E., utilizadas en este estudio.

### Conflicto de interés

Los autores declaran que no tiene conflicto de interés.

### Financiación

Centro Dermatológico Federico Lleras Acosta, E.S.E., dentro del proyecto de investigación 4000-16.2K.

### Referencias

- Castiblanco C.** Situación de la lepra en Colombia, 2007. *Inf Quin Epidemiol Nac.* 2008;13:129-40.
- World Health Organization.** Global leprosy situation, 2009. *Wkly Epidemiol Rec.* 2009;84:333-40.
- World Health Organization.** Expert Committee on Leprosy. *6th Report*, Technical Report Series 768. Geneva: WHO; 1988.
- ILEP Federation.** Improving skin smears & the reading of the bacteriological index in MDT leprosy control programmes. Guidelines and Practice Technical Bulletin. Issue No 3. 1990. Fecha de consulta: 18 de junio de 2010. Disponible en: <http://www.ilep.org.uk/technical-advice/guidelines-practice/>
- Thangaraj RH, Yawalkar SJ.** La lèpre pour les médeciens et le personnel paramédicale. Troisième édition révisée. Balê, Suisse: Ciba-Geigy SA; 1988. p. 16-23.
- Ridley DS.** Bacterial indices. In: Cochrane RG, Davey TF, editors. *Leprosy in theory and practice*. Bristol: John Wright and Sons Ltd.; 1964. p. 620-2.
- Guerrero MI, León CI, Naranjo N, Orozco LC, Giraldo E, Camargo D, et al.** El laboratorio en lepra: bacteriología y patología. Red Nacional de Laboratorios. Bogotá: Instituto Nacional de Salud; 1992.
- Machin D, Campbell M, Fayers P, Pinol A.** Sample size tables for clinical studies. Second edition. Oxford: Blackwell Science; 1997.
- Cohen J.** A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20:37-46.
- Landis JR, Koch GG.** The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-74.
- Lin LI-K.** A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989;45:255-68.
- World Health Organization.** Shortening duration of treatment of multibacillary leprosy. *Wkly Epidemiol Rec.* 1997; 72: 125-32.
- Zweig MH, Campbell G.** Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem.* 1993;39:561-77.
- National Institute of Water and Atmospheric Research.** Lin's concordance. Fecha de consulta: 30 de mayo de 2011. Disponible en: <http://www.niwa.co.nz/our-services/online-services/statistical-calculators/lins-concordance>.
- Awofeso N.** Inventory of skin smear practices in 6 leprosy control programmes in Nigeria. *Lepr Rev.* 1993;64:150-6.
- de Rijk AJ, Nilsson T, Chonde M.** Quality control of skin smear services in leprosy programmes: Preliminary experience with inter-observer comparison in routine services. *Lepr Rev.* 1985;56:177-91.
- Balagon MF, Cellona RV, Cruz E, Burgos JA, Abalos RM, Walsh GP, et al.** Long-term relapse risk of multibacillary leprosy after completion of 2 years of multiple drug therapy (WHO-MDT) in Cebu, Philippines. *Am J Trop Med Hyg.* 2009;81:895-9.
- Ferreira SM, Ignotti E, Senigalia LM, Silva DR, Gamba MA.** Recurrence of leprosy cases in the State of Mato Grosso, Central-West Brazil. *Rev Saude Publica.* 2010;44:650-7.
- Teixeira MA, Silveira VM, França ER.** Characteristics of leprosy reactions in paucibacillary and multibacillary individuals attended at two reference centers in Recife, Pernambuco. *Rev Soc Bras Med Trop.* 2010;43:287-92.
- van Veen NH, Meima A, Richardus JH.** The relationship between detection delay and impairment in leprosy control: A comparison of patient cohorts from Bangladesh and Ethiopia. *Lepr Rev.* 2006;77:356-65.
- Norman G, Joseph G, Richard J.** Validity of the WHO operational classification and value of other clinical signs in the classification of leprosy. *Int J Lepr Other Mycobact Dis.* 2004;72:278-83.
- Teixeira AC, Cruvinel DL, Roma FR, Luppino LF, Resende LH, Sousa T, et al.** Evaluation of the agreement between clinical and laboratorial exams in the diagnosis of leprosy. *Rev Soc Bras Med Trop.* 2008;41(Suppl.2):48-55.
- World Health Organization.** Expert Committee on leprosy. 7th report. Technical Report Series, 874. Geneva: WHO; 1998.
- Georgiev GD, McDougall AC.** The bacteriological examination of slit-skin smears in leprosy control programmes using multiple drug therapy: A plea for radical changes in current operational methodology. *Indian J Lepr.* 1987;59:373-85.
- Croft RP, Smith WC, Nicholls P, Richardus JH.** Sensitivity and specificity of methods of classification of leprosy without use of skin smear examination. *Int J Lepr Other Mycobact Dis.* 1998;66:445-50.
- Groenen G, Saha NG, Rashid MA, Hamid MA, Pattyn SR.** Classification of leprosy under field conditions in Bangladesh. I. Usefulness of skin-smear examinations. *Lepr Rev.* 1995;66:126-33.