

Prueba piloto para determinar la reproducibilidad de la carta de Snellen al evaluar la agudeza visual en sujetos jóvenes

Alejandro León Álvarez* / Jorge Mario Estrada**

RESUMEN

La agudeza visual es la principal característica evaluada para conocer cómo ve una persona. Al medirla se utilizan Cartas de Agudeza Visual, de las cuales la más empleada en el medio colombiano es la carta Snellen. **Objetivo:** el propósito de esta investigación fue determinar la reproducibilidad ínter e intraobservador de la carta de letras serif Snellen para evaluar la agudeza visual (AV). **Materiales y métodos:** en un muestreo por conveniencia, dos evaluadores midieron en dos sesiones la AV con la carta de Snellen tipo serif en 110 estudiantes universitarios (220 ojos) entre 15 y 30 años de edad, con ametropía corregida, visión binocular normal y sin patologías del segmento anterior o posterior. Se repitieron las pruebas en los pacientes induciendo una miopía de 1,50 D. Los valores de AV fueron convertidos de la escala fraccional a la logarítmica. **Resultados:** los resultados mostraron que la AV promedio (LogMAR) fue -0,06 (20/20⁺⁵); con la miopía inducida, 0,51 (20/63).

La prueba de Wilcoxon pareada mostró diferencias significativas para la primera sesión entre los dos evaluadores ($p = 0,0001$; $p = 0,007$) y con el defecto inducido interevaluadores en las dos sesiones ($p = 0,03$; $p = 0,001$). El coeficiente de correlación y concordancia de Lin (Pc) indicó pobre concordancia (Pc = 0,9) ínter e intraobservador, sin defecto inducido y con éste. Los límites de concordancia de Bland-Altman muestran variaciones de AV de 3 líneas en los sujetos corregidos y de 5 a 6 líneas de visión con la miopía inducida. **Conclusiones:** la carta de Snellen presenta concordancia baja y variabilidad en los resultados moderada-alta, por lo que se sugiere la realización de estudios que evalúen la validez de la prueba en las poblaciones sana y con alteraciones visuales.

Palabras clave: agudeza visual, carta Snellen, reproducibilidad, LogMAR.

* Optómetra y maestrante en Ciencias de la Visión, Universidad de La Salle, docente investigador del grupo Salud Visual, Línea de Investigación en Salud de la Fundación Universitaria del Área Andina, Seccional Pereira. aleon@funandi.edu.co

** Terapeuta respiratorio, especialista en Epidemiología, Estudiante de Matemáticas de la UTP, docente de la Facultad Ciencias de la Salud, investigador del grupo GIBIBIO, línea de investigación en Salud de la Fundación Universitaria del Área Andina, seccional Pereira.

Pilot trial to determine the reliability of the snellen chart for assessment of the visual acuity in young people

ABSTRACT

The visual acuity is the main characteristic evaluated to know how a person sees. For the visual acuity measuring it is used the letter chart, in Colombia the most used it is the Snellen Chart. **Objective:** To determinate the reliability inter e intraevaluator of Snellen Chart to test visual acuity. **Methods:** With convenience sampling two evaluators tested visual acuity with serif Snellen Chart to 110 subjects (220 eyes) who were university students between 15-30 years old, full distance refractive correction, normal binocular vision, and no pathologies of anterior and posterior of eye. A refractive error (myopia 1.50D) was induced for all the subjects and the trials were carrying out.

Results: Visual acuity means was -0.06 ($20/20^{+5}$), with induced myopia 0.51 ($20/63$). Wilcoxon sign rank test showed difference between inter evaluators ($p=0,0001$; $p=0,007$) in first trial and wearing plus lens $+1.50D$ between inter evaluators in both trials ($p=0,03$; $p=0,001$). **Conclusions:** Correlation Coefficient and Concordance of Lin showed a poor concordance and a moderated-high variability in the outcomes; then, Snellen chart has a poor concordance and moderate variability, so we suggest achieving researches about the validity of this test in the population with normal and subnormal vision.

Keywords: visual acuity, Snellen chart, reliability, coefficient of correlation and concordance of Lin, LogMAR.

zontal como vertical, no es igual, lo cual hace que la interacción de contorno y, por tanto, la visibilidad sean mayores en algunos casos. Esta carta ha sido empleada como test estándar para evaluar la agudeza visual clínicamente y, en muchos casos, para la investigación. Sin embargo, no se han realizado estudios correspondientes a la validez de la prueba y solo se encontró un estudio donde se evalúa la correlación de la prueba. Se encontró que era alta, pero presentaba una marcada dispersión de los resultados, por lo que se recomendaba el empleo de la carta LogMAR en vez de la Snellen (Lovie-Kitchin, 1988). Esta dispersión tiene una importancia clínica grande; si se sigue la evolución de un problema visual (ambliopía, ametropía, patologías oculares, etc.), una variabilidad grande haría que en ocasiones se sobrevalore o se subvalore la efectividad del tratamiento o la corrección, lo que llevaría a la toma de decisiones clínicas erróneas que podrían afectar la salud visual del sujeto o a gastos de dinero innecesarios de este o del profesional.

Con este trabajo se pretende determinar la reproducibilidad e intraobservador de la prueba de Snellen, por medio de un estudio piloto para sentar un precedente acerca de la variabilidad en la toma de la agudeza visual con esta cartilla la más utilizada en Colombia, y realizar estudios posteriores sobre la validez de esa técnica.

MATERIALES Y MÉTODOS

Se realizó una prueba piloto con el propósito de generar evidencia descriptiva y analítica que sirva de soporte para determinar, en un estudio posterior, la validez de la carta de Snellen.

PACIENTES

Fueron seleccionados 110 estudiantes (220 ojos) de pregrado de la Fundación Universitaria del Área

Andina, Seccional Pereira, convocados de manera voluntaria a participar en el estudio en octubre de 2008. Los participantes presentaban agudeza visual en VP de 0,50M, estereopsis igual a 40 segundos de arco (Titmus Test), defecto refractivo corregido entre +1,50 D y -1,00D, astigmatismo no mayor de 1,00 D; polo anterior y posterior normales. El rango de edad estuvo entre 15 y 30 años. Fueron excluidos aquellos sujetos que presentaran estrabismo o ambliopía.

PROCEDIMIENTOS

El estudio, realizado siguiendo lineamientos establecidos (Consilium Ophthalmologicum Universale, 1988; Benjamin y Borish, 2006), fue llevado a cabo en dos sesiones (una semana de diferencia) por dos evaluadores del programa de optometría, entrenados previamente en la evaluación de la agudeza visual. Se empleó la carta de letras serif de Snellen, ubicaba a seis metros del sujeto. El test tenía una luminancia de 110 cd/m². Se ocluyó primero el OI (con parche pirata) y se pidió al examinado que comenzara a leer el optotipo de la primera línea hasta la última letra que fuese capaz de reconocer. La prueba culminaba cuando el sujeto no podía leer más de 50% de las letras del nivel de agudeza visual examinado. Luego se continuó con el OD y ambos ojos. A continuación se colocó un lente positivo de +1,50D sobre la corrección (si la tenía) del paciente y se repitió el procedimiento.

Se dejó descansar cinco minutos al sujeto y el segundo evaluador procedió a repetir el procedimiento. Los resultados obtenidos por un examinador fueron desconocidos por el otro durante la investigación. Una semana después se realizó la segunda sesión.

Con el propósito de cuantificar de forma continua la agudeza visual, se transformó la fracción Snellen en unidades logarítmicas, asignándole un valor a cada letra, dependiendo del nivel en el cual se encontrara.

ANÁLISIS ESTADÍSTICO

Se realizó un estudio exploratorio descriptivo de los datos mediante análisis de tendencia central y dispersión y algunas variables demográficas; se usó un nivel de confianza $p < 0,05$ en todas las pruebas y se empleó la prueba de Wilcoxon pareada para obtener la diferencia de medias. Para el análisis de la concordancia se usó el coeficiente de correlación-concordancia de Lin (ρ_c) con sus respectivos componentes: la exactitud llamada por algunos autores sesgo (C); para la precisión se empleó el coeficiente de co-

relación de Pearson (ρ). Se utilizaron gráficas LOA para mostrar las diferencias entre medias intraobservador e interobservador, acompañadas de los límites de acuerdo de Bland y Altman (2003). El análisis fue ejecutado en STATA intercooled, versión 9.0

ANÁLISIS DE RESULTADOS

La agudeza visual promedio estuvo entre -0,06 ($20/20^{+5}$) y -0,072 ($20/20^{+6}$). Con el defecto inducido, fue similar para ambos evaluadores 0,5 ($20/63$). Los valores obtenidos por cada evaluador en cada sesión se observan en la tabla 1.

Tabla 1. Valores promedio para los diferentes evaluadores y las sesiones realizadas. La agudeza visual está expresada en unidades LogMAR.

Variable	Media	IC (95%)
SNA1 Evaluador A primera sesión	-0,06 ($20/20^{+5}$)	-0,071 -0,04 ($20/20^{+6}$ a $20/20^{+3}$)
SNA2 Evaluador A segunda sesión	-0,06 ($20/20^{+5}$)	-0,075 -0,045 ($20/20^{+6}$ a $20/20^{+4}$)
SNB1 Evaluador B primera sesión	-0,078 ($20/20^{+6}$)	-0,095 -0,062 ($20/20^{+7}$ a $20/20^{+5}$)
SNB2 Evaluador B segunda sesión	-0,07 ($20/20^{+6}$)	-0,09 -0,06 ($20/20^{+7}$ a $20/20^{+5}$)
SNA1M Evaluador A primera sesión defecto inducido	0,52 ($20/63^{-1}$)	0,5 0,55 ($20/63$ a $20/63^{-2}$)
SNA2M Evaluador A segunda sesión defecto inducido	0,52 ($20/63^{-1}$)	0,5 0,55 ($20/63$ a $20/63^{-2}$)
SNB1M Evaluador B primera sesión defecto inducido	0,5 ($20/63$)	0,48 0,53 ($20/63^{+1}$ a $20/63^{-1}$)
SNB2M Evaluador B segunda sesión defecto inducido	0,5 ($20/63$)	0,46 0,51 ($20/63^{+3}$ a $20/63$)

La prueba de Shapiro-Wilk mostró que ninguna de las variables presentaba una distribución normal ($p: 0,000$), ya que la aplicación de la t-pareada no es precedente. Se empleó la prueba pareada de Wilcoxon.

Se obtuvieron cambios estadísticamente significativos entre evaluadores cuando se midió la AV en sujetos normales y con el defecto inducido (tabla 2). Sin embargo, desde el punto de vista clínico las diferencias son insignificantes (2 ó 3 letras), lo que no alcanza a ser siquiera una línea de agudeza visual.

Tabla 2. Diferencia de medias intraobservadores e interobservadores.
Las diferencias están expresadas en unidades LogMAR.

Variables	Diferencia de medias	IC (95%)	Valor p*
SNA1-SNA2	0,005	-0,005 0,015	0,4526
SNB1-SNB2	-0,005	-0,017 0,006	0,6439
SNA1-SNB1	0,023	0,012 0,034	0,0001**
SNA2-SNB2	0,013	0,003 0,023	0,0075**
SNA1M-SNA2M	-0,005	-0,027 0,017	0,8682
SNB1M-SNB2M	0,014	-0,005 0,034	0,1772
SNA1M-SNB1M	0,019	-0,001 0,039	0,0301**
SNA2M-SNB2M	0,038	0,014 0,062	0,0014**

* Prueba Pareada de Wilcoxon ($P \leq 0.05$) ** Diferencias significativas

Desde el punto de vista estadístico, la concordancia inter e intraobservadora sería pobre (tabla 3). El coeficiente de correlación de concordancia de Lin (ρ_c) estipula que valores por debajo de 0,90 deben

ser considerados faltos de fuerza en la concordancia (Camacho, 2008); sin embargo, si se tienen en cuenta los valores expresados para el coeficiente de kappa, la concordancia encontrada sería moderada (NIWA, 2008).

Tabla 3. Concordancia intra e interobservadora del test de Snellen empleando el coeficiente de correlación de concordancia de Lin (Pc).

Variables	Pc	IC 95%	Pearson (r)	S	Cs	Valor p
SNA1-SNA2	0,783	0,726 0,829	0,784	1,035	0,999	0,000
SNB1-SNB2	0,776	0,718 0,823	0,778	0,956	0,998	0,000
SNA1-SNB1	0,737	0,671 0,791	0,752	0,944	0,980	0,000
SNA2-SNB2	0,809	0,759 0,849	0,821	0,872	0,985	0,000
SNA1M-SNA2M	0,675	0,596 0,741	0,675	1,020	1,000	0,000
SNB1M-SNB2M	0,683	0,606 0,748	0,573*	1,015	0,997	0,000
SNA1M-SNB1M	0,710	0,639 0,769	0,717	1,109	0,990	0,000
SNA2M-SNB2M	0,560	0,464 0,642	0,573*	1,103	0,976	0,000

IC: Intervalo de confianza de Pc. S: Pendiente. Cs: Factor de corrección de sesgos

* Variables con menor concordancia

Clínicamente se observa que la concordancia sería moderada para la evaluación en sujetos normales. La diferencia apreciada fue aproximadamente 3 líneas de agudeza visual, mientras que cuando se indujo el defecto los límites fueron mayores, entre 5 y 6 líneas,

en los cuales se obtuvieron (tabla 3) los valores más bajos de r: (0,573).

Al evaluar los límites de concordancia en (intra) y entre (inter) los evaluadores, se evidencia que los va-

lores de AV oscilan de +/- 2 a 3 líneas de visión y se observa la presencia de mayor cantidad de valores extremos hacia una menor AV en la primera sesión

(figuras 1A y 1C). Para la segunda sesión, los datos marginales se sitúan hacia una mejor AV (figura 1B y 1D).

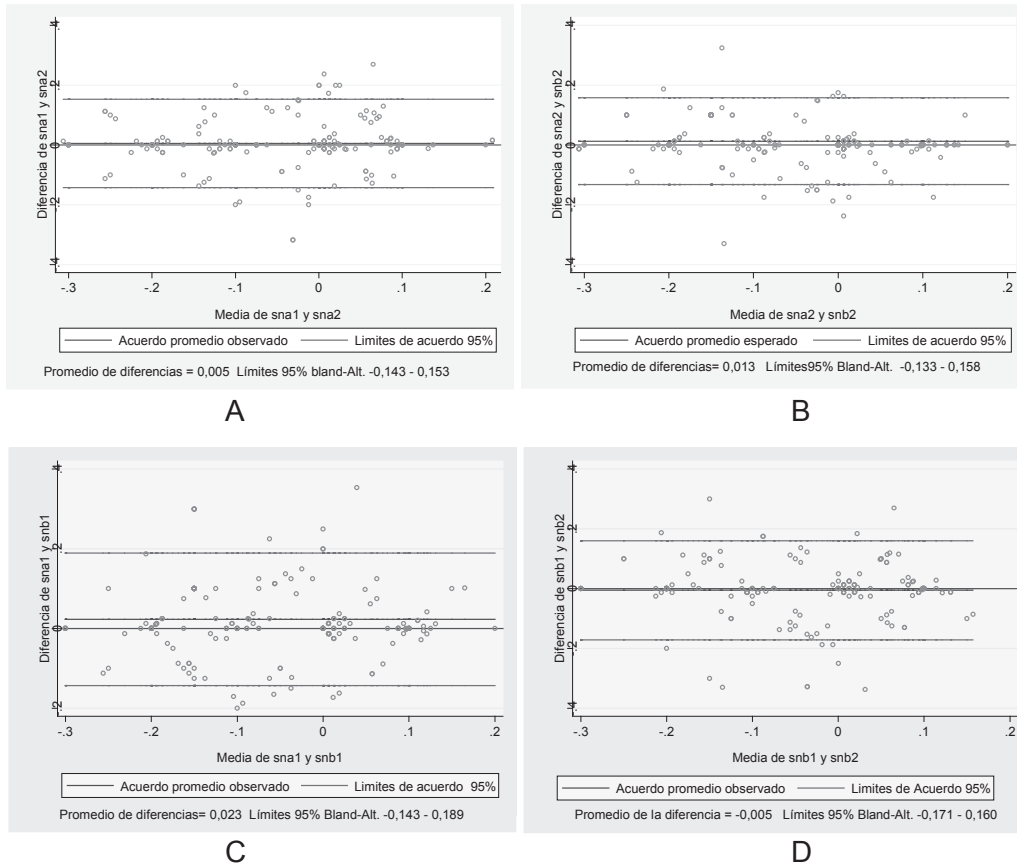


Figura 1. Límites de concordancia intra e interevaluadores para la agudeza visual medida en sujetos sanos visualmente. Los límites oscilan en +/- 3 líneas de agudeza visual.

Al inducir el defecto refractivo, los límites de concordancia se amplían en casi el doble (5 a 6 líneas de AV) (figura 2); sin embargo, los valores extremos no

muestran un patrón de ubicación (mayor o menor visión) igual al presentado cuando se evaluó la visión sin el lente de +1,50D.

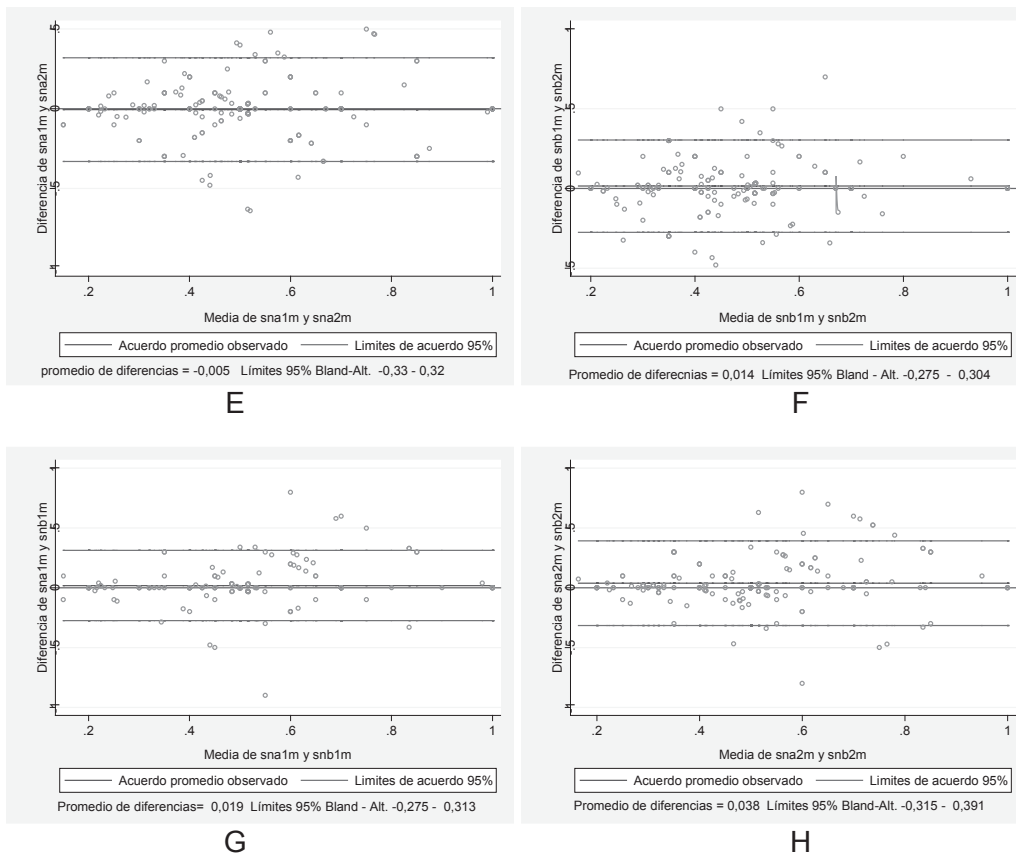


Figura 2. Límites de concordancia para la agudeza visual con un defecto miópico de 1,50D inducido intra e interobservador.

DISCUSIÓN

Los métodos para evaluar concordancia entre sistemas de medición clínica se clasifican en tres tipos: herramientas descriptivas, índices de acuerdo no escalados e índices de acuerdo escalados (Huiman *et ál*, 2007); entre los no escalados están los límites de acuerdo (LOA) de Bland y Altman (1986). Este método LOA se enfoca básicamente en la relación gráfica de las diferencias entre los dos métodos contra la media ± 2 desviaciones estándar, lo cual muestra qué tanto se alejan las diferencias de un promedio verdadero y empieza a dar una idea del grado de acuerdo de las pruebas, además de que provee de una estimación por intervalos llamados límites, de

acuerdo con 95% de confianza, que da una idea definitiva de si las pruebas son intercambiables.

Como índice de acuerdo escalado se tiene el coeficiente de correlación-concordancia de Lin (ρ_c) (Huiman *et ál*, 2007), el cual presenta enormes ventajas frente a métodos usados convencionalmente para evaluar el grado de acuerdo entre dos sistemas de medición inter e intraobservador. En el artículo original de Lin (1989) se explican las dificultades de otros métodos como la prueba t-pareada, la correlación de Pearson, el coeficiente de correlación intraclass, la prueba t-student pareada, que busca la diferencia entre la mediciones pero no realiza un análisis sobre el grado de acuerdo de las mediciones, es decir, la caída

de los datos sobre una línea de 45°. Otro caso interesante es el coeficiente de correlación de Pearson, que aunque muestra la relación lineal entre los datos, y aun observándose una alta correlación no implica directamente acuerdo entre ellos, caso posible solo si los puntos caen sobre una línea igual de 45°; por el contrario, la correlación se cumple con solo ver que caen sobre cualquier línea. Otros sustentos de por qué el coeficiente no mide concordancia (acuerdo) se encuentran en el artículo original de Bland y Altman (1986). Esta argumentación sirvió a Lin para proponer el conocido Coeficiente Concordancia Correlación (ρ_c) que, a través de su formulación (Lin 1989), plantea un coeficiente que evalúa el grado en que los pares de las observaciones caen sobre la línea de 45° o su desviación teniendo en cuenta una pendiente de 1 y un intercepto 0. Así cualquier desviación produciría un coeficiente $\rho_c < 1$, pero esta nueva medida viene acompañada de componentes que proporcionan mayor información sobre el acuerdo entre las pruebas: el sesgo o exactitud (C), que toma valores posibles $0 \leq C \leq 1$, con valores de 1 indicando la no desviación de la mejor línea de 45°, y la precisión, dada por el coeficiente de Pearson (ρ), que evalúa el grado de alejamiento de los datos de la línea de 45°. En Colombia la carta de letras Snellen es empleada por la mayoría de profesionales del sector salud para evaluar la agudeza visual en la población de 5 años en adelante. Generalmente, los sitios donde se emplea no presentan las dimensiones requeridas para realizar la prueba, que se lleva a cabo casi siempre en una distancia de 3 metros. Este hecho hace que la agudeza visual pueda ser subestimada debido a que se sitúa el límite de resolución en 20/20 (0,0), pero, como se muestra en este artículo el promedio se sitúa en un punto mayor que este valor. Esta medida es inferior en una línea a la encontrada por De Haan en 1862 (citado por Colenbrander 2001) y por Elliot *et ál.* (1995). El primero empleó la carta diseñada por Snellen; el segundo, la diseñada por Bailey y Lovie (1976). La diferencia con el primer estudio estaría en los optotipos empleados por De Haan, ya que las

figuras usadas no tenían el mismo grado de visibilidad entre ellas. Así, algunas serían más fáciles de identificar, lo que llevaría a una sobreestimación de la agudeza visual. Las cartas actuales de Snellen emplean letras tipo serif con el mismo grado de visibilidad; sin embargo, algunos autores han criticado este tipo de letras debido a que podrían ocasionar mayor dificultad en identificarlas (Green, 1868; Grimm *et ál.*, 1994; Rabbetts, 1998). Elliot *et ál.* emplearon una carta LogMAR. Estas tienen la ventaja de contar con un grado de visibilidad similar entre las diferentes letras, emplear optotipos no serif y generar una interacción de contorno similar en los diferentes niveles. Estos elementos harían que la agudeza visual no fuese sobreestimada ni subestimada, como podría ocurrir con la carta Snellen.

La comparación de medias mostró diferencias clínicas no significativas entre los evaluadores y en ellos, lo que llevaría a concluir que la carta de Snellen es un test con buena reproducibilidad. No obstante, el empleo de la diferencia de medias para evaluar la concordancia y reproductibilidad de una prueba diagnóstica ha sido objetado por algunos autores. Ruiz y Morillo (2004) mencionan que esta prueba puede generar resultados equivocados por diferencias sistemáticas entre los métodos de evaluación y, por tanto, producir resultados errados, independientemente de que exista o no concordancia. Szklo y Nieto (2003) advierten que los valores extremos afectan el resultado y que si el tamaño de la muestra es pequeño podría no haber significancia estadística así hubiese concordancia (Martínez *et ál.*, 2006). Cuando se emplea la diferencia de medias se puede obtener el “mismo” resultado sin una diferencia significativa entre ellos, lo que conduce a obtener poca información de la exactitud de la prueba (Altman, Bland, 1983).

Los coeficientes de concordancia de Lin mostraron que la prueba tiene una pobre reproductibilidad y los límites de concordancia evidenciaron que las

medidas oscilaron entre 3 líneas de agudeza visual, cuando se evaluaron sujetos emétopes, pero la variabilidad aumentó hasta alcanzar 5 ó 6 líneas al inducir el defecto refractivo.

No se encontraron estudios en que se evalúe la concordancia de la carta Snellen. Lovie-Kitchin (1988) evaluó la validez y fiabilidad de las mediciones en la agudeza visual empleando la carta de Flom, LogMAR y Snellen, y demostró que la correlación test retest era muy buena para esta última ($r: 0,94$). Sin embargo, la dispersión de los resultados era mucho mayor que la presentada por las otras cartas. Al tomar en cuenta los límites de confianza, halló que podría haber una diferencia de hasta 0,26 unidades LogMAR; por tanto, al medir la AV con esta carta y obtener 20/20, con el re-test se podría hallar 20/10. Estos resultados, en cuanto a los límites, son similares a los hallados en este estudio y aunque Lovie-Kitchin no empleó la gráfica de Bland-Altman para los límites de concordancia, su análisis lleva a concluir que el test de Snellen tiene limitaciones en la exactitud de la medida de la, por lo que sería recomendable el empleo de la carta LogMAR.

Los cambios aleatorios en los niveles de agudeza visual, así como el número de optotipos por nivel, pueden generar mayor o menor concordancia de la carta Snellen. Bailey *et ál.* (1991) hallaron discrepancias en la concordancia cuando se empleaba un criterio de diferencia entre cambios de líneas o de letras de AV. La concordancia era mayor para el primer criterio, lo que conduciría a que el clínico no pudiera apreciar cambios significativos en la visión. Cuando se empleó el segundo criterio, se pudieron apreciar mejor los cambios en la visión, por lo que concluyen que la escala diseñada por Bailey y Lovie es la mejor para comparar los cambios en la agudeza visual de las personas. En este trabajo encontramos resultados de concordancia diferentes a los planteados por estos autores; sin embargo, esto podría explicarse de la siguiente manera: la carta Snellen presenta dife-

rente cantidad de figuras por nivel, en los más bajos existen menor cantidad de optotipos, que tan solo vendrían a evaluar cambios gruesos de visión, mientras que en los niveles altos el número de optotipos es mayor, lo que lleva a evaluaciones más finas. Pero estas diferencias en los niveles bajos y altos llevaría a que la visión se sobreestime para los primeros, puesto que con ver tan solo dos letras podría considerarse que está en el nivel de AV que se evalúa, cuando posiblemente tan solo sea parcial. Asimismo, el ver una sola letra (de tres) indicaría que no está en el nivel de visión y por tanto llevaría a subestimar la AV. Estas variaciones harían que la concordancia y los límites de ésta fuesen muy variables, como se muestra en este estudio. Para los niveles más altos sucedería algo similar. El mayor número de letras sobrevaloraría o subvaloraría la agudeza visual en estas líneas, puesto que los cambios ocurren en una escala logarítmica de aproximadamente 0,0125, mientras que el cambio necesario para saber si existe variación o no de la visión es 0,05 logMAR. Esto también lleva a subvalorar la concordancia.

Al evaluar la reproductibilidad de la prueba induciendo un desenfoque óptico de +1,50 D, la concordancia es más baja aun y se aprecia una dispersión mucho mayor. Rosser *et ál.* (2004) examinaron la influencia del desenfoque sobre la repetitividad de la agudeza visual empleando una carta ETDRS (LogMAR) y encontraron que cuanto mayor era el desenfoque inducido, eran mayores los límites de concordancia (con 0,00 D 0,10 logMAR; +0,50 D 0,18 logMAR; +1,00 D 0,24 logMAR), casi 2 ½ líneas de visión. Estos datos concuerdan con el aumento en la variabilidad de los resultados en este estudio, mas no en la magnitud, puesto que aquí el aumento se evidenció en 5 líneas de visión. Esto podría ser el resultado de los mismos problemas inherentes de la carta de Snellen, ya mencionados (McGraw *et ál.* 1995). Además si observamos la diferencia de medias como indicador de reproductibilidad, los valores de Rosser y colaboradores son similares a los hallados por no-

sotros, es decir, de 2 a 3 letras de diferencia entre las pruebas, por lo que nuevamente esto llevaría a una conclusión errónea de la repetitividad del test.

CONCLUSIÓN

El estudio muestra que la carta de Snellen tiene una reproducibilidad entre pobre y moderada, que la variabilidad de los resultados es clínicamente amplia, principalmente ante la presencia de errores refractivos, que el empleo de la diferencia de medias podría generar errores de interpretación en cuanto a si la técnica puede ser válida o no. Esto podría contribuir a que la agudeza visual sea subestimada o sobreestimada en los sujetos evaluados. La carta de Snellen presenta una concordancia baja y una variabilidad en los resultados moderada-alta; en consecuencia, se

sugiere la realización de estudios que evalúen la validez de la prueba en la población general sana y en la que tiene alteraciones visuales.

RECOMENDACIÓN

Realizar un estudio con la carta de Snellen empleando una población más amplia en los rangos de edad, así como en sujetos que presenten diferentes tipos de ametropía, defectos de la visión binocular, ambliopía y patologías de segmento posterior para hallar la validez de este test en la práctica clínica.

AGRADECIMIENTOS

A los estudiantes de Neurofisiología de la Visión por su colaboración en la obtención de los pacientes y los datos para la realización de este trabajo.

BIBLIOGRAFÍA

- Altman, D., Bland J. (1983). Measurement in medicine: The analysis of method comparison studies. *The Statistician*, 32, 307-317.
- Bailey, I., Bullimore, M., Raasch, T., & Taylor, H. (1991). Clinical grading and the effects of scaling. *Investigative Ophthalmology and Vision Science*, 32, 422-432.
- Bland, J., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1: 307-310.
- Bland, J. & Altman, D. (2003). Applying the right statistics: Analyses of measurement studies. *Ultrasound in Obstetrics and Gynecology*, 22, 85-93.
- Bailey, I. & Lovie, J. (1974). New design principles for visual acuity letter charts. *American Journal of Optometry and Physiological Optics*, 53, 740-745.
- Benjamin, W. & Borish, I. (2006). Borish's clinical refraction. St. Louis, MO: Butterworth-Heinemann Elsevier.
- Camacho-Sandoval, J. (2008). Coeficiente de concordancia para variables continuas. *Academia Mexicana de Ciencias*, 50(4), 211-212.
- Colenbrander, A. (2001a). *History of visual acuity measurement*. Meeting of the Cogan Society for Ophthalmic History. San Francisco, United States. Extraído el 15 de noviembre del 2008 desde http://www.ski.org/Colenbrander/Images/History_VA_Measurement.pdf

- Colenbrander, A. (2001b) Measuring vision and vision loss. Duane's Clinical Ophthalmology. San Francisco, United States. Extraído el 16 de noviembre del 2008 desde http://www.ski.org/Colenbrander/Images/Measuring_Vis_Duane01.pdf
- Consilium Ophthalmologicum Universale International Council of Ophthalmology (1988). Visual Acuity Measurement Standard. *Italian Journal of Ophthalmology*, 1, 1-15.
- Currie, Z., Bhan, A., & Pepper, I. (2000). Reliability of Snellen Charts for testing visual acuity for driving: Prospective study and postal questionnaire. *British Medical Journal*, 321, 990-2.
- Elliot, D., Yang, K., & Whitaker, D. (1995). Visual acuity changes throughout adulthood in normal, healthy eyes: Seeing beyond 6/6. *Optometry and Vision Science*, 72, 186-191.
- Green, J. (1868). On a new series of test-letters for determining the acuteness of vision. *Trans Am Ophthalmol Soc*, 1(4-5), 68-71.
- Grimm, W., Rassow, B., Wesemann, W., Saur, K., & Hiltz, R. (1994). Correlation of optotypes with the Landolt Ring-A fresh look at the comparability of optotypes. *Optometry and Vision Science*, 71, 6-13.
- Huiman, X., Barnhart, M., & Lin, L. (2007). An overview on assessing agreement with continuous measurements. *Journal of Biopharmaceutical Statistics*, 17, 529-569.
- Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45, 255-268.
- Lovie-Kitchin, J. (1988). Validity and reliability of visual acuity measurements. *Ophthalmic and Physiological Optics*, 8, 363-370.
- Martínez-González, M., Sánchez-Villegas, A., & Faulín, F. (2006). *Bioestadística Amigable* (2ª ed.) España: Díaz de Santos.
- McGraw, P., Winn, B., & Whitaker, D. (1995). Reliability of the Snellen Chart. *British Journal of Ophthalmology*, 310, 1481-1482.
- National Institute of Water and Atmospheric Research. *Lin's Concordance*. Extraído el 11 de noviembre de 2008 desde <http://www.niwa.cri.nz/services/free/statistical/concordance>
- Pandit, J. (1994). Testing acuity of vision in general practice: Reaching recommended standard. *British Medical Journal*, 309, 1408.
- Rabbetts, R. (1998). Bennett and Rabbett's Clinical Visual Optics. Butterworth-Heinemann.
- Rosser, D., Murdoch, I., & Cousens, S. (2004). The effect of optical defocus on the test-retest variability of visual acuity Measurements. *Investigative Ophthalmology and Vision Science*, 45, 1076-1079.
- Ruiz, A., Morillo, L. (2004). *Epidemiología clínica. Investigación clínica aplicada*. Bogotá: Panamericana.
- Szklo, M., Nieto, J. (2003). *Epidemiología intermedia, Conceptos y aplicaciones*. Madrid: Díaz de Santos.