

¿Cuánto entrenamiento necesitan los residentes de radiología para una interpretación estandarizada de la TC craneal?

Ignacio Herves Escobedo¹, Irene García Tuells¹,
Gloria Pérez Hernández¹, Marina Lozano Ros¹, Juana
María Plasencia Martínez¹, Jose María García-
Santos¹.

¹Hospital General Universitario Morales Meseguer,
Murcia/ES.

Email: nacho_herves_1993@hotmail.com

Objetivo:

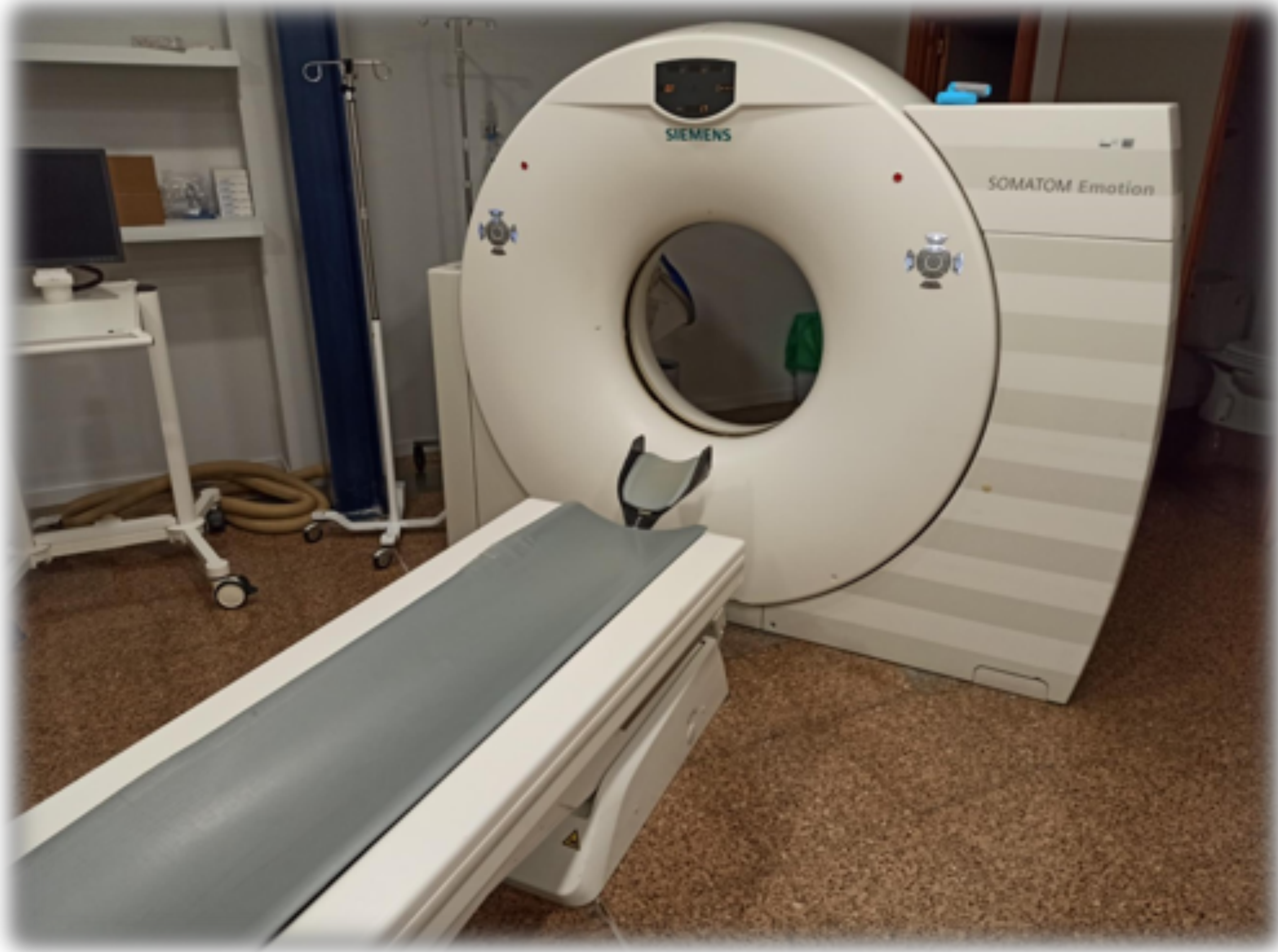
La TC craneal urgente es uno de los estudios más frecuentemente realizados durante las guardias, y su uso aumenta exponencialmente¹.

Son escasos los estudios sobre el grado de acuerdo de su interpretación con un radiólogo experto y su estandarización. Se han llevado a cabo algunos estudios de concordancia previos aunque no específicamente para residentes y centrándose en patologías específicas, como la isquemia^{2,3}. Esos estudios han reconocido una variabilidad interobservador sustancial³.

Nuestro objetivo ha sido cuantificar el grado de entrenamiento necesario por los residentes para estandarizar la lectura de la TC craneal urgente. Creemos que el conjunto de variables analizadas en este estudio puede ser un modelo útil para futuros estudios en el campo.

Material y Método:

Selección de muestra:



106 TC craneales

Adecuado para nuestro objetivo⁴

Somatom Emotion CT

- 16 filas de detectores.
- Posición decúbito supino.
- Grosor corte: 2,5 mm.
- Separación: 2,4 mm.
- Reconstrucción SOFT y BONE.

Se excluyeron:

- estudios de calidad diagnóstica subóptima
- estudios con contraste IV

92 secuenciales

80 mA / 120 Kv

14 helicoidales

270 mA / 130 Kv

No colaborativos.

Motivo de consulta:
traumatismo

Material y Método:

Entrenamiento de lectores:

12 residentes de radiología se reunieron con un neurorradiólogo con 28 años de experiencia para un primer taller teórico-práctico de una hora de duración.



Después del taller, los 12 residentes evaluaron de forma independiente la muestra. Estaban cegados a los datos demográficos.



Tras completar esta evaluación inicial, pasaron una hora de entrenamiento directo diario (segundo taller) con el mismo neurorradiólogo durante cinco días consecutivos.



Dos semanas después del final del segundo taller, nueve de los doce residentes evaluaron la misma muestra bajo las mismas condiciones.

Prueba oro: coincidencias del neurorradiólogo experimentado después de dos lecturas consigo mismo.

Un residente de segundo año, uno de tercero y uno de cuarto año no estaban disponibles en el momento de la segunda lectura.

Material y Método:

Entrenamiento de lectores:

| | |
|---|--|
| <p>Atrofia cortical global (ACG): <i>retracción cortical.</i></p> | <p>0: (normal): ausencia de surcos en la mayoría de los cortes del estudio.</p> <hr/> <p>1: (retracción incipiente): es posible dibujar los surcos bruscamente en la mayoría de los cortes del estudio.</p> <hr/> <p>2: (retracción moderada): surcos claramente abiertos.</p> <hr/> <p>3: (retracción severa): surcos muy distendidos y circunvoluciones adelgazadas.</p> |
| <p>Ventrículos dilatados (VD). Valoración subjetiva. El índice de Evans no fue empleado.</p> | <p>0: Normal.</p> <hr/> <p>1: Dilatado.</p> |
| <p>Enfermedad de la sustancia blanca (ESB): <i>isquemia de vasos pequeños.</i></p> <p>Para evaluar áreas confluentes, se debía excluir la hipodensidad periventricular</p> | <p>0 (normal): ausencia de focos hipodensos.</p> <hr/> <p>1 (carga isquémica leve): focos hipodensos > 3 mm y hasta 1 cm o encefalopatía vascular periventricular SIN otras lesiones focales aisladas</p> <hr/> <p>2 (carga isquémica moderada): parches de > 1 cm y < 2 cm.</p> <hr/> <p>3 (carga isquémica grave): hipodensidad difusa en la sustancia blanca, áreas afectadas de más de 2 cm de diámetro máximo.</p> |
| <p>Lesión específica (LE)</p> | <p>0: Ausencia.</p> <hr/> <p>1: Presencia.</p> |

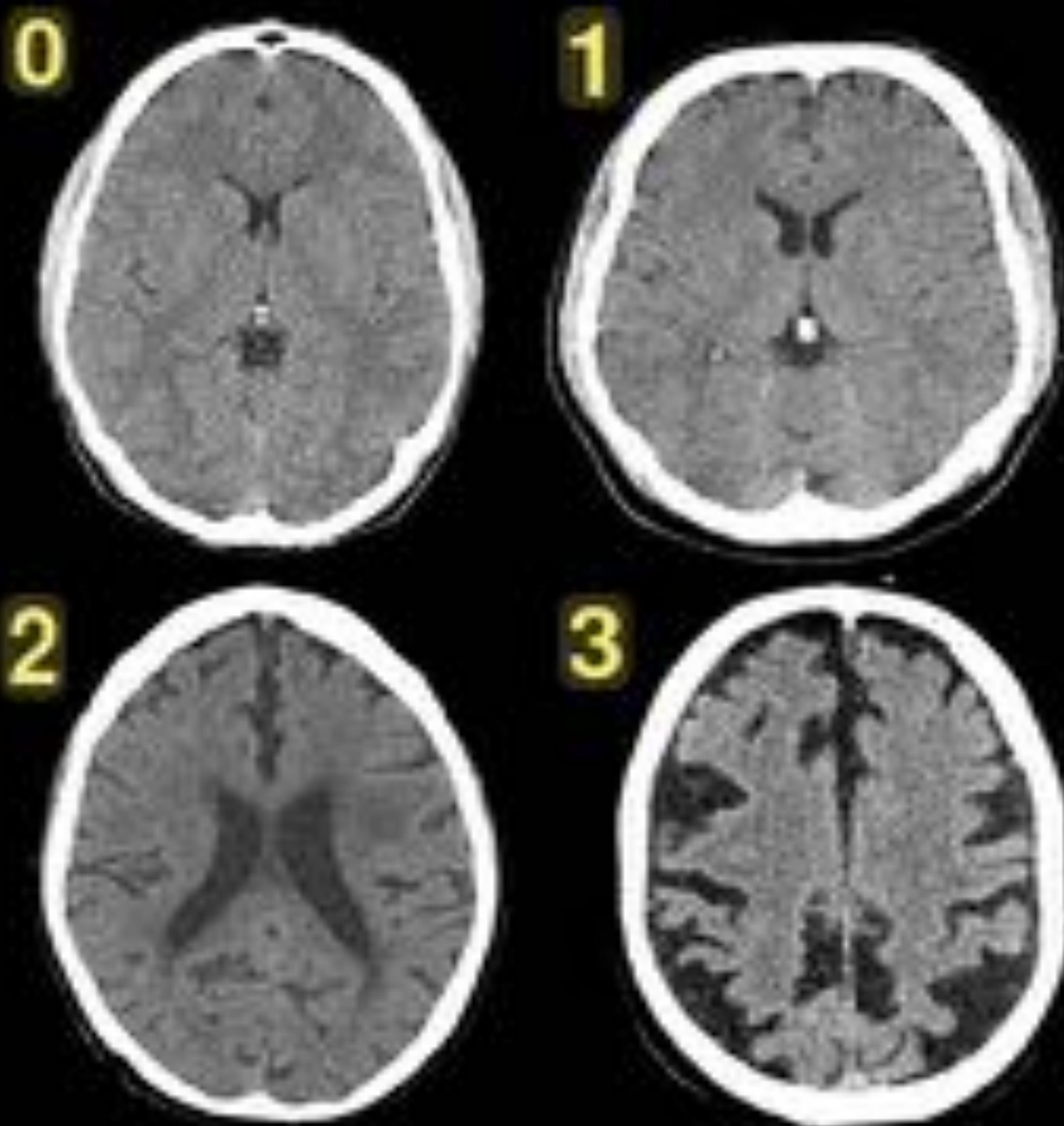
Las variables a estudio aparecen en la Figura 1 junto con ejemplos de cada una en las Figuras 2, 3, 4 y 5

Figura 1

Material y Método:

Entrenamiento de lectores:

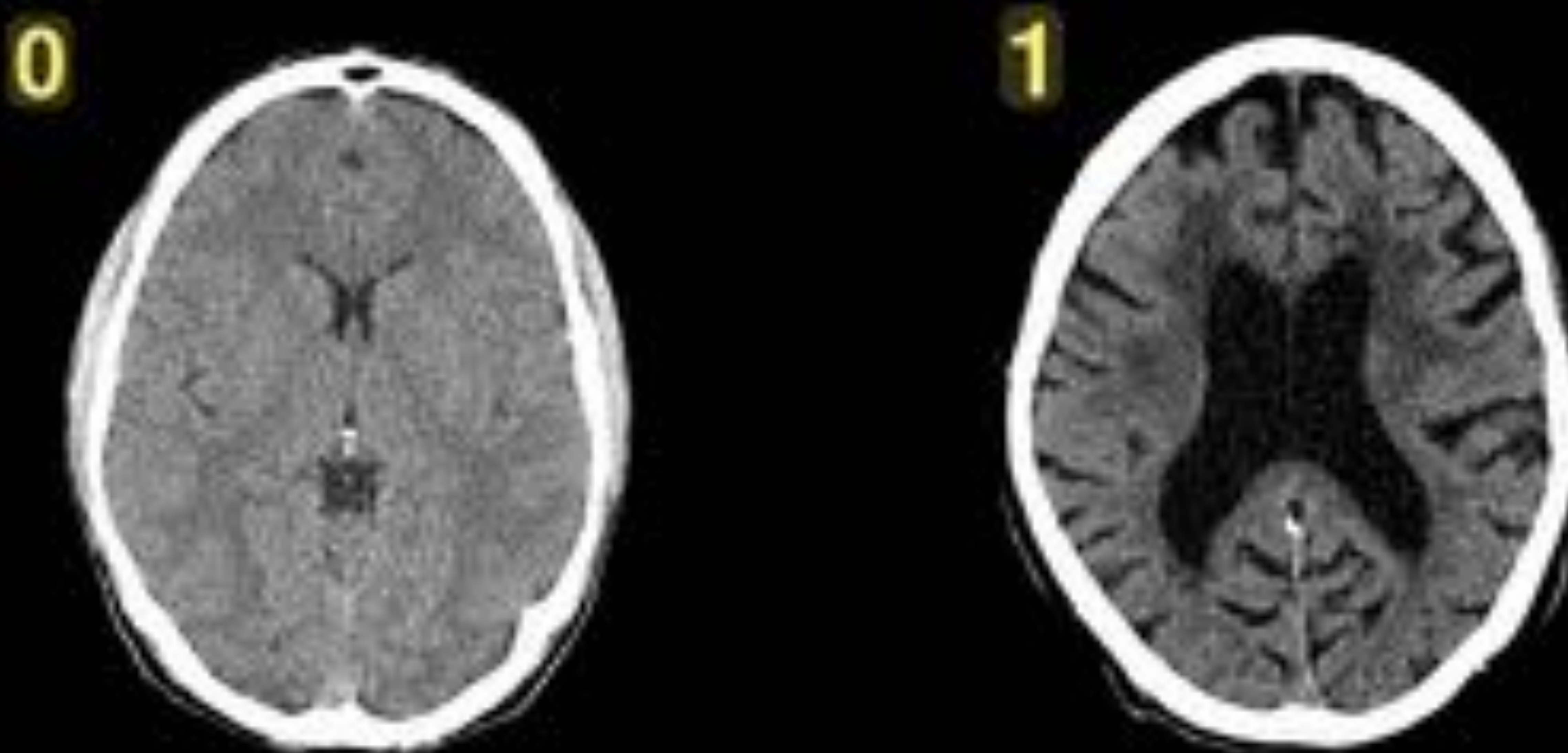
Atrofia cortical global



0: normal.
1: atrofia leve.
2: atrofia moderada.
3: atrofia grave.

Figura 2

Dilatación ventricular



0: normal.

1: ventrículos dilatados.

Figura 3

Material y Método:

Entrenamiento de lectores:

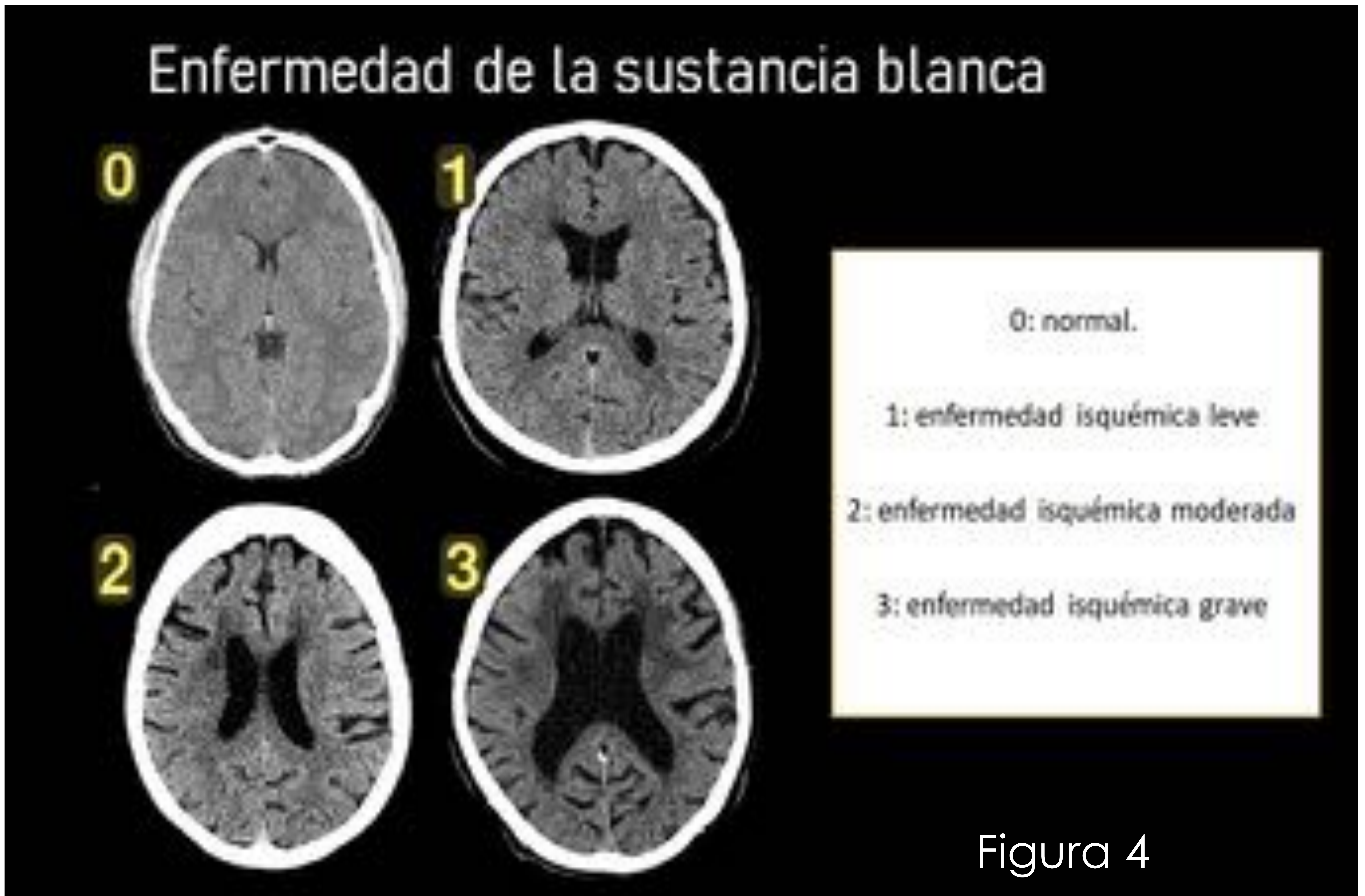


Figura 5 Lesión específica



Material y Método:

Análisis estadístico:

El análisis de acuerdo se evaluó mediante el índice Cohen-Kappa (K). Los resultados se presentaron como rango K, K medias y sus intervalos de confianza del 95%. Los niveles de acuerdo considerados fueron: (a) ningún acuerdo para ≤ 0.20 , (b) débil para 0,21-0,40, (c) moderado para 0,41-0,60, (d) bueno para 0,61-0,80 y (e) excelente para $\geq 0,81$.

También hicimos la evaluación del acuerdo dicotomizando las calificaciones de GCA y WMD como 1 (puntajes 0 y 1: sin enfermedad o leve) y 2 (puntajes 2 y 3: enfermedad grave).

La sensibilidad y la especificidad también se calcularon.

Para el análisis estadístico se emplearon el software IBM SPSS Statistic versión 20 y el software MedCalc versión 12.7.0.0.

Resultados:

Descripción de la muestra:

106 TC craneales consecutivos del departamento de Urgencias

Periodo: entre el 18 de marzo de 2016 y el 11 de mayo de 2016

43 hombres y 63 mujeres, con edades entre 16 a 98 años.

Hallazgos en los estudios de TC:

82 pacientes (77%) tenían una puntuación ≥ 1 para atrofia cortical global

75 pacientes (71%) tenían una puntuación ≥ 1 para enfermedad de sustancia blanca

29 pacientes (27%) presentaban ventrículos dilatados.

46 pacientes (43%) presentaban al menos una lesión específica.

Resultados:

Concordancia:

Previo al primer taller, el rango del K fue de:

ACG: 0,579-0,865

DV: 0,504-0,763

ESB: 0,561-0,839

LE: 0,498-0,723

(Figuras 6 y 7)

Después del primer taller, el K fue buena o excelente ($\geq 0,61$) en:

11/12 residentes para valorar ACG

6/12 residentes para valorar VA

9/12 para valorar ESB

4/12 residentes para valorar LE

(Figura 6)

El nivel de concordancia alcanzado tras el primer taller no estuvo influenciado por el año de residencia, con los siguientes K medios para el 1º, 2º, 3º y 4º año de residencia, respectivamente, de:

-ACG: 0,752, 0,723, 0,718 y 0,738

-DV: 0,609, 0,647, 0,606 y 0,640

-ESB: 0,626, 0,589, 0,717 y 0,737

-LE: 0,598, 0,622, 0,548 y 0,601

(Figura 6)

Resultados:

Concordancia:

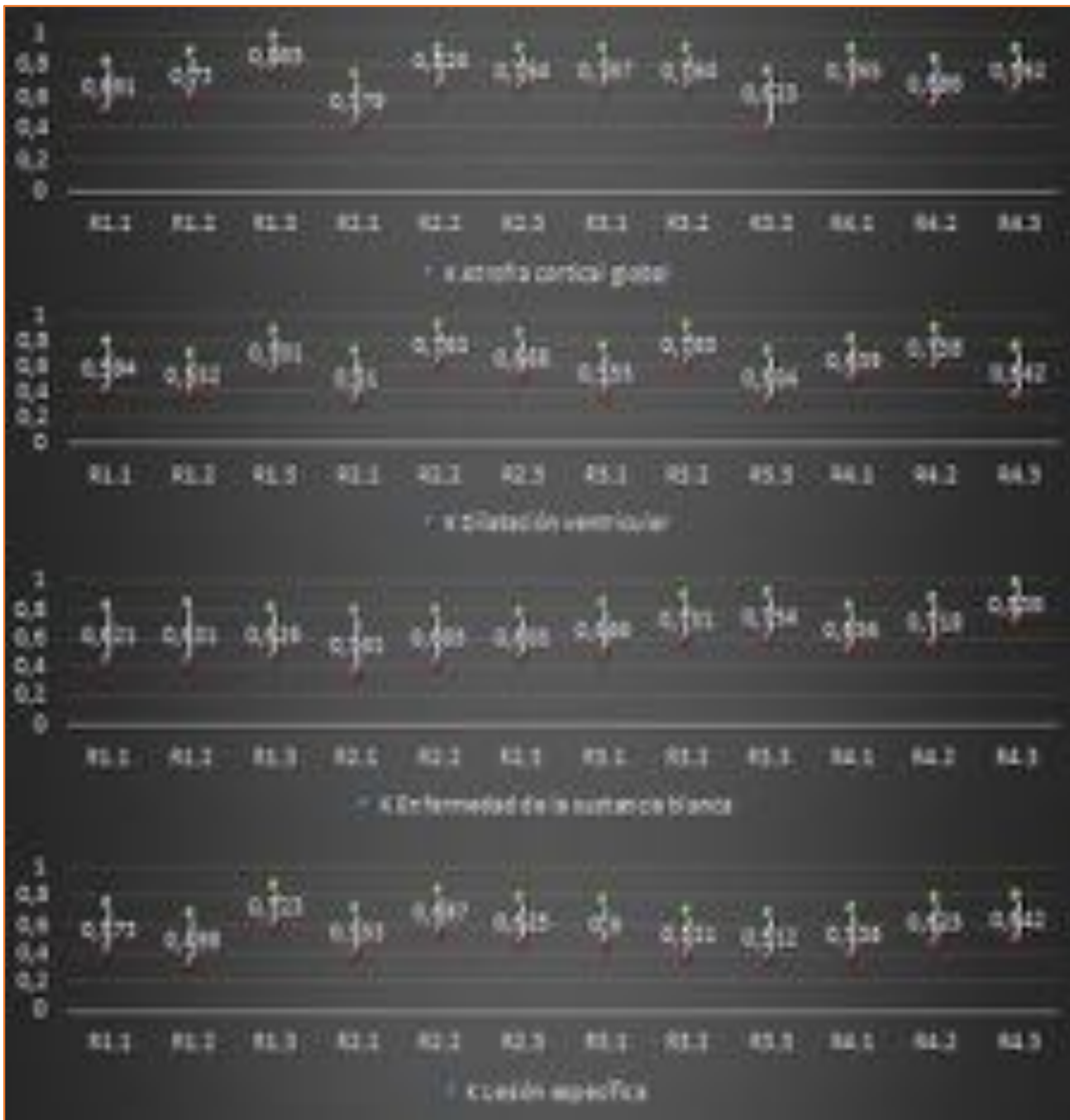


Figura 6

Resultados:

Concordancia:

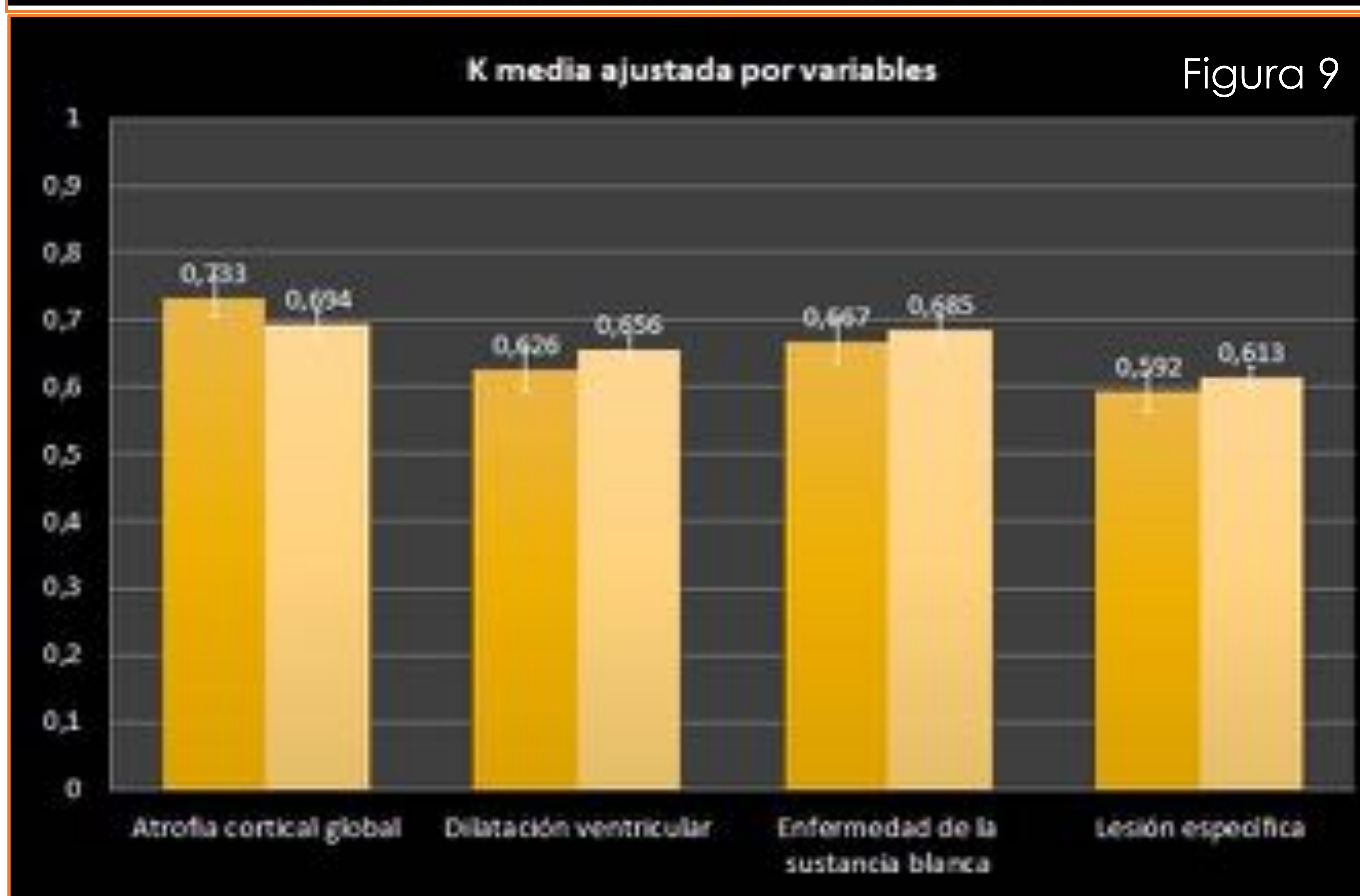
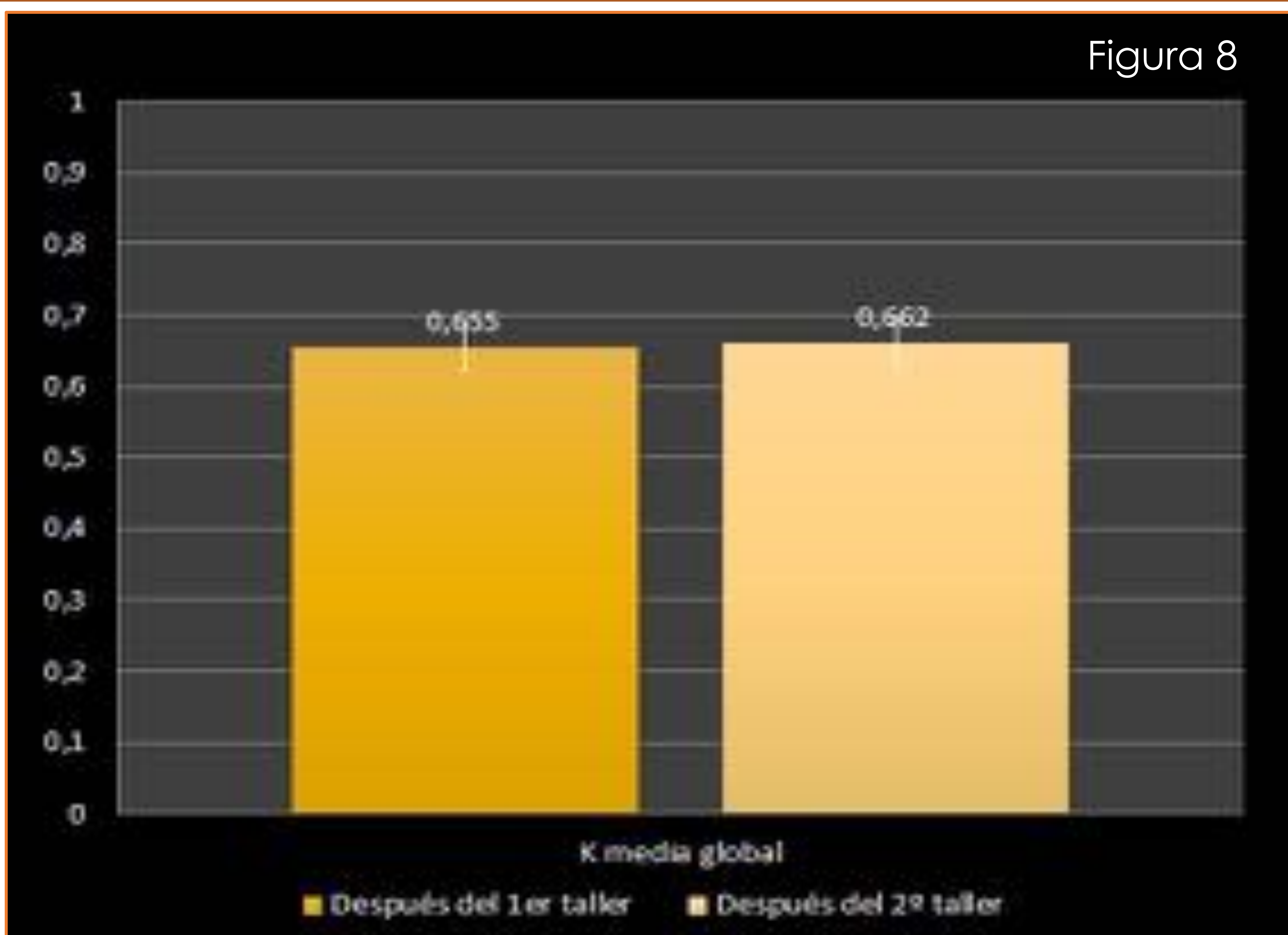
| | Después del 1 ^{er} taller | | | | Después del 2 ^o taller | | | |
|-------|------------------------------------|---------|-------|-------|-----------------------------------|---------|-------|-------|
| | Rango K | K media | Se | Es | Rango K | K media | Se | Es |
| ACG | 0,579-0,865 | 0,733 | 77,96 | 96,84 | 0,531-0,791 | 0,694 | 77,46 | 92,91 |
| DV | 0,504-0,763 | 0,626 | 73,58 | 89,36 | 0,411-0,836 | 0,656 | 68,21 | 94,48 |
| ESB | 0,561-0,839 | 0,667 | 82,59 | 89,68 | 0,497-0,796 | 0,685 | 82,32 | 90,61 |
| LE | 0,498-0,723 | 0,592 | 63,37 | 93,88 | 0,553-0,745 | 0,613 | 62,09 | 96,86 |
| Media | - | 0,655 | 74,38 | 92,44 | - | 0,662 | 72,52 | 93,72 |

Figura 7

Resultados:

Concordancia:

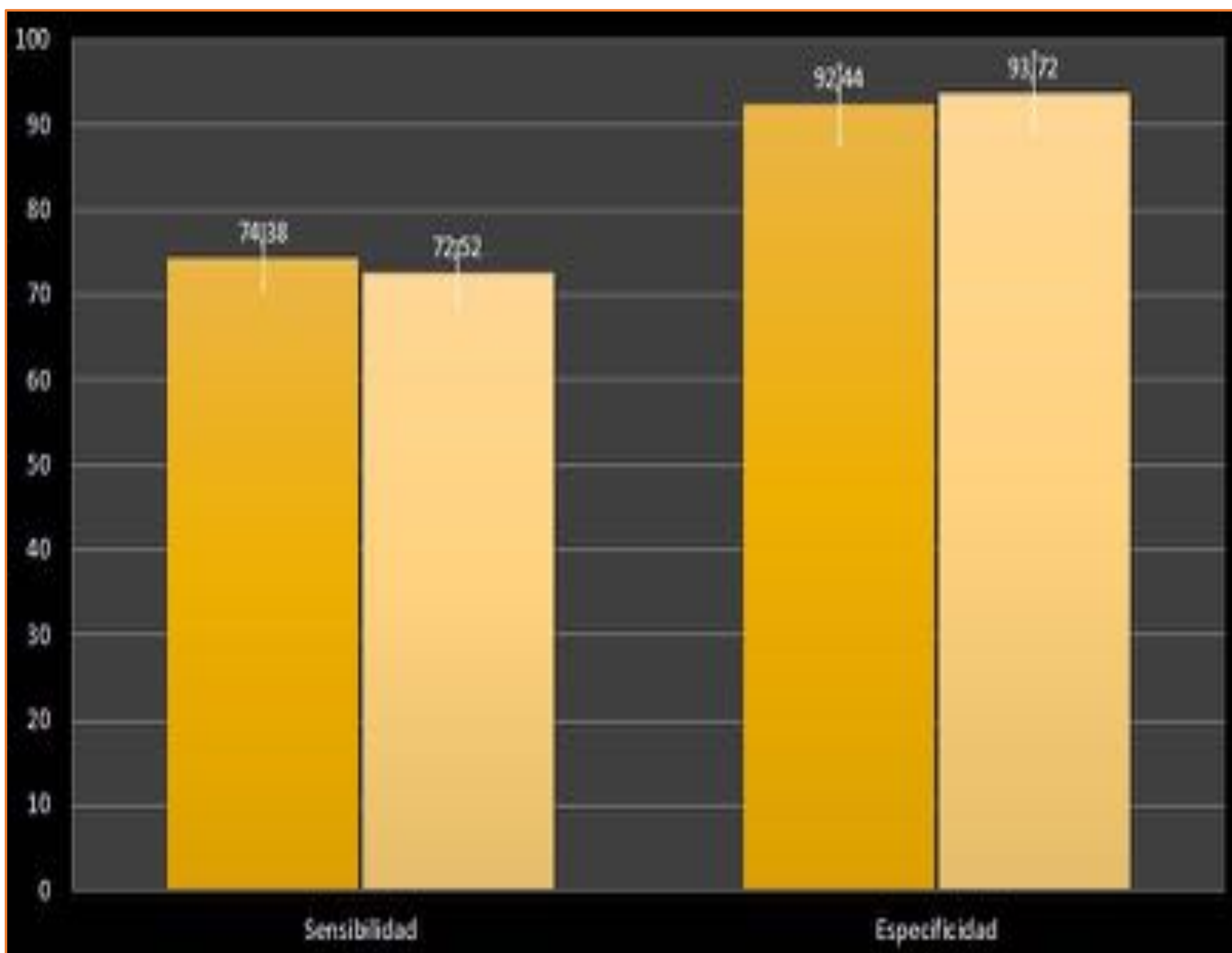
El K mejoró levemente tras el segundo taller en 3/9, 5/9, 6/9 y 6/9 residentes, con un aumento medio global del K de 0,007 (Figura 8), y un aumento medio del K menor que 0,03 para VA, ESB y LE (Figura 9).



Resultados:

Concordancia:

Figura 10



La sensibilidad global no mejoró después del segundo taller (72,52%) en comparación con el primer taller (74,38%).

La especificidad tras el segundo taller (93,72%) fue casi tan alta como después del primer taller (92,44%) (Figura 10).

El peor K (medias de 0,592 y 0,613 después del primer y segundo talleres) y los valores de sensibilidad (63,37% y 62,09% después del primer y segundo talleres) correspondieron a las lesiones específicas.

A diferencia de las otras variables, las diferentes posibilidades de esta variable podría explicar estos resultados.

Resultados:

Limitaciones:

No se tuvieron en cuenta las razones clínicas, sin embargo creemos que este tema tuvo poca influencia en el análisis de concordancia en todas las variables, excepto en la “lesión específica”, en la cual no podemos establecer ese potencial sesgo de variabilidad interobservador.

Tres lectores no pudieron reevaluar la muestra después del segundo taller, lo que puede haber afectado a la mejora del K tras el mismo.

Conclusión:

La lectura de la TC de la cabeza puede estandarizarse entre los residentes de radiología mediante un solo taller, independientemente de su año de residencia, con un acuerdo de moderado a bueno con un neurorradiólogo experimentado.

Proponemos un modelo de evaluación general para la TC craneal que se puede aplicar a otros estudios de concordancia.

Los talleres de capacitación deben implementarse rutinariamente como una herramienta de mejora en la evaluación e interpretación de los TC craneales urgentes entre los residentes de radiología.

Estas estrategias de actualización pueden implementarse para mejorar la atención médica del paciente.

Referencias:

1. National Centre for Health Statistics. U.S. Department of Health and Human Services web site; Health, United States, 2009 with special feature on medical technology.
<http://www.cdc.gov/nchs/data/hus/hus09.pdf>.
2. Pasquier F, Leys D, Weerts JG et-al. Inter- and intraobserver reproducibility of cerebral atrophy assessment on MRI scans with hemispheric infarcts. *Eur. Neurol.* 1997;36: 268-72.
3. Caraganis A, Mulder M, Kempainen RR, Brown RZ, Oswood M, Hoffman B, Prekker ME. Interobserver Variability in the Recognition of Hypoxic-Ischemic Brain Injury on Computed Tomography Soon After Out-of-Hospital Cardiac Arrest. *Neurocrit Care.* 2020. doi: 10.1007/s12028-019-00900-7.
4. im J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther.* 2005 Mar;85:257-68.
5. Fazekas F, Chawluk JB, Alavi A et-al. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *AJR Am J Roentgenol.* 1987;149: 351-6.