

# Indices of agreement between neurosurgeons and a radiologist in interpreting tomography scans in an emergency department

Índice de concordância entre neurocirurgiões e radiologistas na interpretação de tomografias de crânio na emergência

Jules Carlos Dourado<sup>1</sup>, Júlio Leonardo Barbosa Pereira<sup>1</sup>, Lucas Alverne Freitas de Albuquerque<sup>1</sup>, Gervásio Teles Cardos de Carvalho<sup>1</sup>, Patrícia Dias<sup>1</sup>, Laura Dias<sup>2</sup>, Marcos Bicalho<sup>3</sup>, Pollyana Magalhães<sup>1</sup>, Marcos Dellaretti<sup>1</sup>

## ABSTRACT

The power of interpretation in the analysis of cranial computed tomography (CCT) among neurosurgeons and radiologists has rarely been studied. This study aimed to assess the rate of agreement in the interpretation of CCTs between neurosurgeons and a radiologist in an emergency department. **Method:** 227 CCT were independently analyzed by two neurosurgeons (NS1 and NS2) and a radiologist (RAD). The level of agreement in interpreting the examination was studied. **Results:** The Kappa values obtained between NS1 and NS2 and RAD were considered nearly perfect and substantial agreement. The highest levels of agreement when evaluating abnormalities were observed in the identification of tumors, hydrocephalus and intracranial hematomas. The worst levels of agreement were observed for leukoaraiosis and reduced brain volume. **Conclusions:** For diseases in which the emergency room procedure must be determined, agreement in the interpretation of CCTs between the radiologist and neurosurgeons was satisfactory.

**Keyword:** cranial trauma, tomography, neurosurgeons, radiologists.

## RESUMO

O poder de interpretação na análise de tomografias de crânio (TCC) entre neurocirurgiões e radiologistas tem sido pouco estudada. O objetivo deste estudo é avaliar as taxas de concordância na interpretação de TCCs entre neurocirurgiões e radiologista em um departamento de emergência. **Método:** 227 TCCs foram independentemente analisadas por 2 neurocirurgiões (NC1 e NC2) e um radiologista (RAD). O índice de concordância nas análises foi estudada posteriormente. **Resultados:** O valor de Kappa obtido entre os NC1 e NC 2 e entre estes e RAD foram quase perfeitos e substancial respectivamente. O maiores índices de concordância quando avaliadas anormalidades foram observados na identificação de tumores, hidrocefalia e hematomas intracranianos. O piores índices foram observados com relação a leucaraiose e redução volumétrica. **Conclusão:** Para doenças apresentadas em um departamento emergência que demandam tratamento mais agressivo o índice de concordância na interpretação de TCCs entre RAD e NC foi satisfatório.

**Palavras-chave:** trauma craniano, tomografia, neurocirurgiões, radiologistas.

Imaging examinations are frequently performed during emergency care. Following the increased popularity of cranial computed tomography examinations (CCTs), the analysis and interpretation of cranial imaging has become common practice for the neurologist, neurosurgeon and radiologist<sup>1,2,3</sup>.

In numerous centers, full-time radiologists who evaluate CCTs are often unavailable, such that these examinations are interpreted by the neurologist or neurosurgeon on call<sup>4</sup>. Moreover, in many institutions, it is the emergency physicians

who initiate the treatment of patients based on their interpretation of radiographic images, including CCTs, which are then reviewed by a radiologist<sup>5</sup>.

It has been demonstrated that discrepancies can occur in the interpretation CCTs between radiologists and emergency physicians, in around 8 to 11% of cases<sup>6</sup>. A number of studies exist that compare the interpretation of examinations between radiologists and emergency physicians<sup>2,5,6,7</sup>; however, few studies discuss the accuracy

<sup>1</sup>Hospital Santa Casa de Belo Horizonte, Departamento de Neurocirurgia, Belo Horizonte MG, Brazil;

<sup>2</sup>Universidade Federal de Minas Gerais, Belo Horizonte MG, Brazil;

<sup>3</sup>Faculdade de Ciências Médicas de Minas Gerais, Belo Horizonte MG, Brazil.

**Correspondence:** Jules Carlos Dourado; Departamento de Neurocirurgia, Hospital Santa Casa de Belo Horizonte; Rua Padre Rolim, 492; 30130-090 Belo Horizonte MG, Brasil; E-mail: julesdourado@gmail.com

**Conflict of interest:** There is no conflict of interest to declare.

Received 12 January 2015; Received in final form 05 March 2015; Accepted 25 March 2015.

of neurosurgeons at interpreting imaging examinations compared with radiologists<sup>3</sup>.

The purpose of this study was to analyze the rate of agreement in the interpretation of CCTs between two neurosurgeons and a radiologist in an emergency department.

## METHOD

A total of 250 CCTs with a multislice cranial computer tomography scan were performed on patients admitted to the emergency department of São Lucas Hospital of the Santa Casa de Belo Horizonte between July and August 2011, which were randomly distributed for analysis by one of three professionals, two neurosurgeons and a radiologist specialized in neuroradiology. All the professionals independently interpreted all the CCTs and none of them knew the patients beforehand. Reports resulting from analysis of the CCTs were detailed in a standardized form, in which general epidemiological data obtained from medical records were analyzed. Twenty-three examinations were excluded because they were inadequately filled out, thus 227 CCTs were included in the statistical analysis.

Initially, each of the health professionals determined whether the CCT was normal or abnormal, followed by details of any abnormalities identified. The following groups were considered: ischemic and hemorrhagic stroke, intracranial hematomas (subdural hematoma, extradural hematoma or intraparenchymal hemorrhage), intracranial mass lesion, pneumocephalus, skull fractures, brain contusions, lesions with mass effect, midline deviation plus than 5 mm, hydrocephalus and other injuries not previously discriminated.

Statistical analysis was performed using the program SPSS 18.0. The following variables were determined: neurosurgeon 1 (NS1) x radiologist (RAD); neurosurgeon 2 (NS2) x radiologist; and neurosurgeon 1 (NS1) x neurosurgeon 2 (NS2). The level of agreement when identifying the CCT as either normal or abnormal was evaluated using the Kappa test. The same test was also used to identify the correlation between the radiologist and the two neurosurgeons regarding findings of ischemia, brain tumor, hydrocephalus, mass effect, midline deviation plus than 5 mm, pneumocephalus, and leukoaraiosis/reduced of brain volume to verify whether an tomographic alteration identified by the radiologist was also observed by the neurosurgeons. These values were expressed as percentage of agreement between each of the pairs. The Kappa test was also performed to evaluate the degree of agreement among these professionals. The Kappa values were divided for analysis: less than zero, no agreement; 0 to 0.19, poor agreement; 0.20 to 0.39, weak agreement; 0.40 to 0.59, moderate agreement; 0.60 to 0.79, substantial agreement; and 0.80 and 1.00, almost perfect agreement, as per Landis and Koch<sup>8</sup>.

## RESULTS

In the 227 CCTs analyzed, patient age ranged from one month to 90 years old. Sixty-three CCTs were described as normal by RAD, compared with 91 by NS1 and 110 by NS2. In several CCTs, more than one abnormality was described. In these cases, each abnormality was analyzed separately to determine whether each of them had been observed by all the examiners.

The initial evaluation to determine whether the CCT was normal or abnormal showed that the Kappa values were substantial between NS2 and RAD and between NS1 and NS2. Between NS1 and RAD, the Kappa value determined achieved a level which we consider to be almost perfect (Table 1).

Concerning the identification of abnormalities, the Kappa test showed a substantial level of agreement among the three professionals (Table 2).

Regarding the abnormalities reported, in comparisons between the neurosurgeons and the radiologist, the rate of agreement was substantial between NS1 and RAD for identifying findings of ischemia, intracranial hematoma (only subdural hematomas were found in this sample), brain tumor, midline deviation and hydrocephalus. Agreement for the finding of mass effect was moderate. For the findings of pneumocephalus agreement was weak, while for leukoaraiosis and reduced brain volume, it was poor. RAD always identified a greater number of changes than NS1 for all pathologies (Table 3).

When the findings of NS2 were compared with RAD, substantial agreement was shown for hematomas, hydrocephalus and intracranial tumors, moderate agreement was shown for ischemia, pneumocephalus, midline deviation and mass effect and poor agreement was shown for leukoaraiosis and reduced brain volume (Table 3). Again, RAD always identified a greater number of changes than NS2 for all pathologies. RAD also identified the greatest number of changes in all the groups analyzed.

**Table 1.** Kappa test values for agreement between examiners in the identification of normal CCTs.

	RAD	NS2
RAD	-	0.71
NS1	0.81	0.65

RAD: radiologist; NS1: neurosurgeon; NS2: neurosurgeon 2; CCTs: cranial computed tomography examinations.

**Table 2.** Kappa test values for agreement between examiners in the identification of abnormalities in CCTs.

	RAD	NS2
RAD	-	0.62
NS1	0.75	0.64

RAD: radiologist; NS1: neurosurgeon; NS2: neurosurgeon 2; CCTs: cranial computed tomography examinations.

**Table 3.** Principal abnormalities detected and the Kappa values for agreement among the neurosurgeons and radiologist.

Alteration	N° alt.	Kappa NS1	Kappa NS2	Mean Kappa
Ischemia	35	0.65	0.53	0.59
Intracranial hematomas*	15	0.77	0.72	0.75
Brain tumor	2	0.65	0.66	0.66
Pneumoencephalus	5	0.16	0.49	0.32
Mass effect	3	0.47	0.46	0.47
Midline deviation	4	0.79	0.48	0.64
Hydrocephalus	8	0.78	0.77	0.78
Leukoaraiosis and brain volume reduction	50	0.21	0.17	0.19

NS1: neurosurgeon; NS2: neurosurgeon 2. \*Only subdural hematomas were found.

## DISCUSSION

Research evaluating the quality of the analysis of health care professional who are not radiologists in their interpretations of imaging examinations, particularly those directly related to emergency care, where a full-time radiologist is not always available, has shown growth in recent years.

In this work, our group aimed to verify the rate of agreement between two neurosurgeons and a radiologist in the interpretation of CCTs performed in an emergency department.

Analysis of the results showed a substantial rate of agreement among the examiners in identifying normal and abnormal CCTs. This is fundamental when screening patients in an emergency room for subsequent treatment decisions. The rate of agreement of both NS1 (0.81, almost perfect agreement) and NS2 (0.71, substantial agreement) with RAD was high. Agreement between the two neurosurgeons was also substantial. Mukerji et al.<sup>3</sup> analyzed 192 CCTs to compare the rate of agreement among neurosurgeons and radiologists in identifying normal CCTs and determined Kappa values ranging between 0.51 and 0.85, similar to our results.

In our evaluation of the rate of agreement with RAD for diagnoses of abnormal tests, the rate for NS1 was 0.75 and for NS2 was 0.62, both of which indicate substantial agreement. Analysis of agreement regarding CT imaging evaluations according to pathology showed that the highest rates of agreement between the neurosurgeons and the radiologist were observed for intracranial hematomas, hydrocephalus and ischemic lesions. This is of great importance as these lesions often require emergency surgery procedures, where the wrong diagnosis could compromise the clinical outcome. Mukerji et al.<sup>3</sup> also demonstrated high levels of agreement for

diagnosing intracranial hemorrhage (subdural hematoma, intracerebral and subarachnoid hemorrhage) when they analyzed the interpretations of radiologists and neurosurgeons.

The worst rates of agreement in our series were observed when diagnosing changes interpreted by the radiologist as reduced brain volume and leukoaraiosis; the Kappa value for NS1 was 0.21 and for NS2 was 0.17, indicating poor agreement. Erly et al.<sup>9</sup> analyzed the interpretation of radiology residents compared with neuroradiologists in the evaluation of 1324 CCTs, demonstrating greater disagreement in skull fractures and cerebral ischemia. In the work of Mukerji et al.<sup>3</sup>, the main disagreements between radiologists and neurosurgeons were the loss of differentiation of gray and white matter and ischemia. In our series, the highest rate of disagreement concerned changes described as leukoaraiosis, where the radiologist identified far more cases than the neurosurgeons. This change is defined as a nonspecific radiological sign that has been observed in both healthy and sick individuals<sup>10</sup>. Leukoaraiosis in CCTs can be related to a risk of cerebrovascular disease, but the extent of its involvement remains uncertain and it has no influence on emergency procedures<sup>11</sup>.

This study has numerous limitations, particularly with respect to the limited number of examiners and the large number of CCTs diagnosed as normal, which reduced the universe of abnormalities that were evaluated.

Despite these, we can conclude that for pathologies that must be diagnosed quickly to determine procedures that should be performed while in the emergency department, the level of agreement between the three professionals analyzed was satisfactory but further investigations are necessary due to the studies limitations.

## References

- Berkseth TJ, Mathiason MA, Jafari ME, Cogbill TH, Patel NY. Consequences of increased use of computed tomography imaging for trauma patients in rural referring hospitals prior to transfer to a regional trauma centre. *Injury*. 2014;45(5):835-9. <http://dx.doi.org/10.1016/j.injury.2014.01.002>
- Gallagher FA, Tay KY, Vowler SL, Szutowicz H, Cross JJ, McAuley DJ et al. Comparing the accuracy of initial head CT reporting by radiologists, radiology trainees, neuroradiographers and emergency doctors. *Br J Radiol*. 2011;84(1007):1040-5. <http://dx.doi.org/10.1259/bjr/24581602>
- Mukerji N, Cahill J, Paluzzi A, Holliman D, Dambatta S, Kane RJ. Emergency head CT scans: can neurosurgical registrars be relied upon to interpret them? *Br J Neurosurg*. 2009;23(2):158-61. <http://dx.doi.org/10.1080/02688690902730723>

4. O'Leary MR, Smith M, Olmsted WW, Curtis DJ. Physician assessments of practice patterns in emergency department radiograph interpretation. *Ann Emerg Med.* 1988;17(10):1019-23. [http://dx.doi.org/10.1016/S0196-0644\(88\)80438-4](http://dx.doi.org/10.1016/S0196-0644(88)80438-4)
5. Alfaro D, Levitt MA, English DK, Williams V, Eisenberg R. Accuracy of interpretation of cranial computed tomography scans in an emergency medicine residency program. *Ann Emerg Med.* 1995;25(2):169-74. [http://dx.doi.org/10.1016/S0196-0644\(95\)70319-5](http://dx.doi.org/10.1016/S0196-0644(95)70319-5)
6. Mucci B, Brett C, Huntley LS, Greene MK. Cranial computed tomography in trauma: the accuracy of interpretation by staff in the emergency department. *Emerg Med J Emj.* 2005;22(8):538-40. <http://dx.doi.org/10.1136/emj.2003.013755>
7. Espinosa JA, Nolan TW. Reducing errors made by emergency physicians in interpreting radiographs: longitudinal study. *BMJ.* 2000;320(7237):737-40. <http://dx.doi.org/10.1136/bmj.320.7237.737>
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-74. <http://dx.doi.org/10.2307/2529310>
9. Erly WK, Berger WG, Krupinski E, Seeger JF, Guisto JA. Radiology resident evaluation of head CT scan orders in the emergency department. *AJNR Am J Neuroradiol.* 2002;23(1):103-7.
10. Verny M, Duyckaerts C, Pierot L, Hauw JJ. Leuko-araiosis. *Dev Neurosci.* 1991;13(4-5):245-50. <http://dx.doi.org/10.1159/000112168>
11. Helenius J, Tatlisumak T. Treatment of leukoaraiosis: a futuristic view. *Curr Drug Targets.* 2007 Jul;8(7):839-45. <http://dx.doi.org/10.2174/138945007781077436>