Inter-Observer Agreement in ASPECTS

Michael Städt[1]*, Michael Lell[1], Johannes Schwab[2] and Heinz Voit-Höhne[1]

[1]*Department of Radiology und Nuclear Medicine, Paracelsus Medical University, Nuremberg, Germany*

[2]*Department of Cardiology, Paracelsus Medical University, Nuremberg, Germany*

Abstract

Introduction: Stroke is the second most common cause of death worldwide. Its reliable assessment and effective subsequent treatment are therefore essential. Non-Enhanced Cranial Computed Tomography (NECT) and Perfusion-CT (PCT) are leading modalities in stroke imaging. The aim of our study was to evaluate the visual assessment of NECT and perfusion maps using ASPECTS with a focus on inter-observer agreement and inter-modality correlation.

Materials and methods: The database of our department was searched for patients with suspicion of ischemic stroke in the anterior circulation who underwent technically successful NECT and PCT. In total, 40 patients were randomly selected. ASPECTS evaluation for NECT and for PCT was conducted by four fellow (<5 years of experience) and four expert radiologists (>5 years of experience) in a high-volume stroke center. Bland-Altmann plots were created for modality correlation, Intraclass Correlation Coefficient (ICC) was used to analyse inter-observer agreement and dichotomisation was used for further evaluation.

Results: Significant disagreement regarding modality correlation for all fellow radiologists and for one expert radiologist was shown. Bland Altmann plots for the three remaining expert radiologists showed a correlation between increased experience and better agreement between imaging modalities.

Inter-observer agreement for NECT as well as for PCT was excellent for fellow radiologists. Inter-observer agreement for expert radiologists was only fair for NECT and moderate for PCT.

The inter-observer agreement for the dichotomized ASPECTS values showed only fair agreement for NECT and moderate agreement for PCT between fellow radiologists. Agreement between expert radiologists for NECT as well as for PCT was insufficient.

Conclusion: We conclude that PCT must be used with caution and as a support rather than an independent modality for clinical decisions.

Keywords: Aspects; CT; Inter-observer agreement; Perfusion CT; Stroke

Introduction

Stroke is the second most common cause of death worldwide [1]. Non-Enhanced Cranial Computed Tomography (NECT) remains the most common, and most important imaging modality used to rule out intracranial haemorrhage [2]. Ischemic stroke usually requires further imaging to establish diagnosis and to direct treatment options. In most cases CT- Angiography (CTA) and CT-Perfusion (PCT) are the modalities of choice [3,4].

Early signs of ischemia in NECT, especially in the Middle Cerebral Artery (MCA) territory [i.e. hypoattenuation, sulcal swelling or a hyperdense artery sign can be subtle and the extent of hypoperfusion

may be difficult to assess [5-7]. To standardize stroke evaluation, the Alberta stroke program early CT Score (ASPECTS) was established [8-10]. ASPECTS was originally used with NECT, but then also transferred to PCT [11-13].

Many studies regarding ASPECTS in NECT have shown a moderate to excellent inter-observer agreement in various constellations, although these groups consisted mainly of smaller groups of expert neuroradiologists and neurologists [8-12,14]. To our knowledge there is only one study from 2013 that includes an analysis of inter-observer agreement between fellow radiologists [15], a group that is more likely to be confronted with stroke imaging in non-office hours. This study showed moderate inter-observer agreement for NECT and a good to excellent agreement for PCT, but more recent studies have resulted in inter-observer disagreement of PCT and a general disagreement overall [16,17].

That is why further analysis of reproducibility and agreement of NECT and PCT in terms of ASPECTS in daily practice is required.

The aim of our study is to evaluate the visual assessment of NECT and PCT with a focus on inter-observer agreement as well as the modality correlation between two groups of experts and fellow radiologists.

Materials and Methods

Ethics statement

Internal review board approval was obtained for this evaluation. Due to the retrospective study design informed consent was waived.

Patients

The database of our department (PACS, picture archive and communication system) was searched for all cases of suspicion of stroke. Only patients who underwent NECT and PCT between

December 2015 and December 2019 were included in the study population.

Exclusion criteria were intracerebral haemorrhage, diagnosis other than ischemic strokes (e.g. stroke mimics such as epileptic seizures) and stroke in the posterior circulation. As the calculation of ASPECTS requires good imaging quality, NECT and PCT scans with poor quality were excluded.

A minimum National Institute of Health Stroke Scale (NIHSS) or minimum duration of neurological symptoms was not defined for this study, since our stroke protocol required a higher NIHSS of at least 8 to perform further imaging after the NECT in any case.

From a total of 222 patients, 35 cases were excluded. Out of the remaining, 40 cases were randomly selected and anonymised.

Imaging protocol

All imaging studies were performed on a 64-MDCT scanner (Somatom AS, Siemens Healthcare, Forchheim, Germany). The imaging protocol included NECT, PCT and CTA.

NECT was performed using 120 kV, 340 mAs, and slice thickness of 4.8 mm. Scan direction was caudo-cranial. To protect the lenses from direct radiation, the scan volume extended to the supraorbitomeatal line to the Vertex.

PCT volume had longitudinal coverage of 84 mm, including at least the level of basal ganglia and the lateral ventricles to ensure a properly adjusted imaging for ASPECTS grading. 50 ml of a non-ionic contrast agent (Solutrast 370, Bracco Imaging, Konstanz, Germany) was injected intravenously with a power injector and a flow of 7 ml/s being followed by 40 ml of saline with the same flow rate. After a delay of six seconds a stack of 30 images was acquired. Data acquisition was performed at 120 kV and 180 mAs, 10 mm contiguous axial slices were reconstructed for further PCT analysis.

PCT post processing

Post-processing was performed on a standard workstation with clinically available PCT software (Syngo VE32D, Siemens AG, Berlin, Germany). Cerebral Blood Volume (CBV) and Cerebral Blood Flow (CBF) maps were calculated with a deconvolution-based algorithm and saved for assessment by using the automatically generated threshold.

Image assessment

All NECT and PCT scans were assessed with widely and free available software (Osirix Version 5.8.5, Pixmeo, Bernex, Switzerland).

The observers consisted of a group of 10 people. 5 of them were experienced radiologists with more than 5 years of experience in stroke imaging, the other five were fellow radiologist with 1 to 4 years of experience. All observers had a training phase in ASPECTS of at least 6 weeks. Training in ASPECTS had been provided since the beginning of 2016. All observers independently documented their ASPECTS evaluation on a standardized template adapted from the article by the ASPECTS study group in 2000 [8].

The observers were blinded for all clinical information except the side of symptoms and the NIHSS. They assessed the NECT first, followed by the CBV and CBF-maps, having access to every acquired slice. CTA slices were not included in this study. As published in

previous studies [15] a very broad definition of early signs of ischemic changes was used, for example hypoattenuation and swelling of the brain. Standard window levels (80/35) for the NECT and PCT were used but could be altered by the observer.

Statistical Analysis

Modality correlation described the agreement between the NECT and PCT results from the same observer. One sampled t-tests were calculated for each observer and if significant disagreement ($p<0,05$) was shown they were excluded from further analysis. For the remaining cases a Bland-Altmann Plot was created and further linear regression analysis was performed to rule out proportional bias as listed in the original paper [18].

Inter-observer agreement was defined as the degree of agreement on either the NECT or the PCT values between the fellow radiologists or the experienced radiologists and was analysed using the Intraclass Correlation Coefficient (ICC).

In a further evaluation ASPECTS results from NECT and PCT were additionally divided into dichotomal categories for ASPECTS <6 and ≥ 6. Fleiss-Kappa was used for calculation. Level of agreement for ICC and Kappa were divided into the following categories: Poor for values under 0.20, fair for values between 0.21 to 0.4, moderate for values between 0.41 to 0.6 and good for values between 0.61 to 0.8. Values above 0.81 were considered excellent.

Results

Inter-modality correlation

One sampled t-tests showed significant disagreement for every fellow radiologist and one expert radiologist and was therefore excluded. Furthermore linear regression analysis showed a proportional bias for one of the expert radiologists and was excluded as well. The Bland-Altmann Plots for the remaining three expert radiologists are listed below (Figure 1).
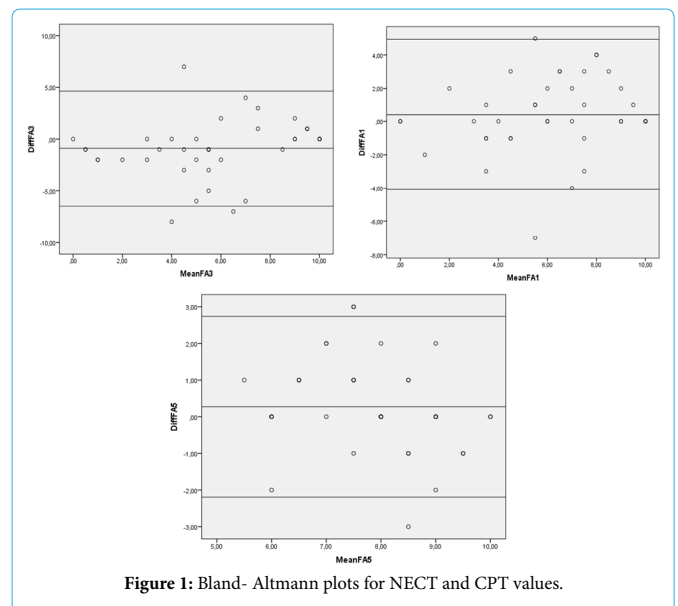

Figure 1: Bland- Altmann plots for NECT and CPT values.

Inter-observer agreement

ICC for the NECT among the fellow radiologist showed an excellent agreement (0, 88) as well as the ICC for the CPT (0, 89). Between the

expert radiologists ICC for NECT showed only fair agreement (0, 41), results for ICC in PCT were moderate (0, 63).

Kappa values for dichotomized values showed a fair agreement for the NECT values of fellow radiologists (K= 0, 36) and marginally fair values for expert radiologists (K=0, 23). Kappa values for PCT-values were moderate for fellow radiologists (K= 0, 56) and poor for expert radiologists (K= 0, 12) (Table 1).

	ICC-NECT	ICC-PCT	NECT-Kappa	PCT-Kappa
Fellow radiologists	0,88	0,89	0,36	0,56
Expert radiologists	0,41	0,63	0,23	0,12

Table 1: ICC and Kappa values for fellow and expert radiologist.

Discussion

Statistical analysis regarding inter-modality correlation showed significant disagreement for all fellow radiologists and for one expert radiologist, but the Bland-Altmann plots for the other three expert radiologists points towards a significant correlation between NECT and PCT. This indicates a link between increased experience and better agreement between imaging modalities.

To our best knowledge there is no similar study comparing ASPECTS evaluation for multiple fellow and expert radiologists. Numerous studies indicate that increasing experience results in better detection of ischemia and there have been many studies with different designs that show moderate to good modality correlation before. But these studies have either only included expert radiologists and physicians for the evaluation of the NECT or have compared different methods of grading the extent of ischemic changes [2,3,5,11,16,19,20].

Surprisingly the inter-observer agreement for NECT as well as for PCT was excellent for fellow radiologists. Contrary to this, inter-observer agreement for expert radiologists was only fair for NECT and moderate for PCT. Despite the fact that many studies show a higher inter-observer agreement [5,9-11,21] and recent studies dealing with perfusion come to the conclusion that PCT is a highly predictive method for ischemic changes in acute stroke [11,19,21,22], our findings point out that PCT can lead to insufficient agreement. Additionally, a recent study from 2016 by Khaw et al. compared PCT assessment by three neuroradiologists and showed similar results [17].

Dichotomisation was performed because at many institutions, including ours, ASPECTS < 6 is considered a contraindication for systemic lysis. The inter-observer agreement for the dichotomized ASPECTS values showed only fair agreement for NECT and moderate agreement for PCT between fellow radiologists. Surprisingly the agreement between expert radiologists for NECT as well as for PCT was insufficient. Data situation on this topic is not conclusive and difficult to compare due to mixed study designs. Older data showed better agreement [5,9,10,14] whereas newer studies show a manifold range of inter-observer variability [9,11,17,19,23-25]. In accordance to a recently published review from 2016 by Farzin et al. [16] our results point out that the extent of ischemia and consequent therapeutic decisions can be difficult to assess and thus unequally estimated even by specialised radiologists. Reasons for this insufficient agreement are unclear and should be the topic of further research. Thereby exact assessment of different cortical regions as well as the evaluation of technical methods such as thin slice reconstruction or increased temporal resolution in Perfusion should be considered [19,26,27].

Though being selected randomly we wanted to include every case from our database, even without MRI-follow up. Consequently we cannot compare the results with ischemic detection on MRI and therefore a precise comparison with many other studies is not possible, which can be seen as an essential limitation for this study. Other limitations result from the lack of MTT and TTP in our PCT-evaluation, especially since it has been pointed out that MTT leads to optimized detection [22].

Though statistical analysis, especially Kappa statistics and Bland-Altmann plots have been widely used in numerous studies dealing with ASPECTS, risks of miscalculation have been described and may have biased our results [23,28].

Conclusion

Assessment of ASPECTS is inconsistent between fellow and expert radiologists. Even evaluation regarding mechanical thrombolysis shows insufficient agreement, accounting for NECT as well as for PCT. Therefore we should conclude that PCT must be used with caution and as a support rather than a modality for clinical decisions.

Further studies evaluating the reasons are necessary to understand the causes of different assessment.

References

1. World Health Organization (2016) Global Health Estimates 2015: Deaths by cause, age, sex, by country and by region, 2000-2015. World Health Organization, Geneva, Switzerland.

2. Kaskar O, Goldstein LB (2017) Utility of urgent computed tomography angiography in the setting of intraparenchymal brain hemorrhage. J Stroke Cerebrovasc Dis 26: 417-419.

3. Schriger DL, Kalafut M, Starkman S, Krueger M, Saver JL (1998) Cranial computed tomography interpretation in acute stroke: Physician accuracy in determining eligibility for thrombolytic therapy. JAMA 279: 1293-1297.

4. Tan JC, Dillon WP, Liu S, Adler F, Smith WS, et al. (2007) Systematic comparison of perfusion-CT and CT-angiography in acute stroke patients. ANN Neurol 61: 533-543.

5. Kalafut MA, Schriger DL, Saver JL, Starkman S (2000) Detection of early CT signs of >1/3 middle cerebral artery infarctions: Interrater reliability and sensitivity of CT interpretation by physicians involved in acute stroke care. Stroke 31: 1667-1671.

6. Mair G, Boyd EV, Chappell FM, von Kummer R, Lindley RI, et al. (2015) Sensitivity and specificity of the hyperdense artery sign for arterial obstruction in acute ischemic stroke. Stroke 46: 102-107.

7. Wardlaw JM, Dorman PJ, Lewis SC, Sandercock PAG (1999) Can stroke physicians and neuroradiologists identify signs of early cerebral infarction on CT? J Neurol Neurosurg Psychiatry 67: 651-653.

8. Barber PA, Demchuk AM, Zhang J, Buchan AM (2000) Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. ASPECTS Study Group. Alberta Stroke Programme Early CT Score. Lancet 355: 1670-1674.

9. Coutts SB, Demchuk AM, Barber PA, Hu WY, Simon JE, et al. (2004) Interobserver variation of ASPECTS in real time. Stroke 35: 103-105.

10. Pexman JH, Barber PA, Hill MD, Sevick RJ, Demchuk AM, et al. (2001) Use of the Alberta stroke program early ct score (ASPECTS) for Assessing CT Scans in Patients with Acute Stroke. AJNR Am J Neuroradiol 22: 1534-1542.

11. Aviv RI, Mandelcorn J, Chakraborty S, Gladstone D, Malham S, et al. (2007) Alberta stroke program early ct scoring of ct perfusion in early stroke visualization and assessment. AJNR Am J Neuroradiol 28: 1975-1980.

12. Bisdas S, Donnerstag F, Ahl B, Bohrer I, Weissenborn K, et al. (2004) Comparison of perfusion computed tomography with diffusion-weighted magnetic resonance imaging in hyperacute ischemic stroke. J Comput Assist Tomogr 28: 747-755.

13. Xin Y, Han FG (2016) Diagnostic accuracy of computed tomography perfusion in patients with acute stroke: A meta-analysis. J Neurol Sci 360: 125-130.

14. Mak HK, Yau KK, Khong PL, Ching AS, Cheng PW, et al. (2003) Alberta stroke programme early ct score. Hypodensity of >1/3 middle cerebral artery territory versus Alberta stroke programme early CT Score (AS-PECTS): Comparison of two methods of quantitative evaluation of early CT changes in hyperacute ischemic stroke in the community setting. Stroke 34: 1194-1196.

15. van Seeters T, Biessels GJ, Niesten JM, van der Schaaf IC, Dankbaar JW, et al. (2013) Reliability of visual assessment of non-contrast ct, ct angiography source images and ct perfusion in patients with suspected ischemic stroke. PLoS One 8: e75615.

16. Farzin B, Fahed R, Guilbert F, Poppe AY, Daneault N, et al. (2016) Early CT changes in patients admitted for thrombectomy: Intrarater and inter-rater agreement. Neurology 87: 249-256.

17. Khaw AV, Angermaier A, Michel P, Kirsch M, Kessler C, et al. (2016) Inter-rater agreement in three perfusion-computed tomography evaluation methods before endovascular therapy for acute ischemic stroke. J Stroke Cerebrovasc Dis 25: 960-968.

18. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1: 307-310.

19. Finlayson O, John V, Yeung R, Dowlatshahi D, Howard P, et al. (2013) Interobserver agreement of ASPECT score distribution for noncontrast CT, CT angiography, and CT perfusion in acute stroke. Stroke 44: 234-236.

20. Wardlaw JM, Mielke O (2005) Early signs of brain infarction at CT: Observer reliability and outcome after thrombolytic treatment--systematic review. Radiology 235: 444-453.

21. Wintermark M, Flanders AE, VelthuisB, Meuli R, van Leeuwen M, et al. (2006) Perfusion-CT assessment of infarct core and penumbra: Receiver operating characteristic curve analysis in 130 patients suspected of acute hemispheric stroke. Stroke 37: 979-985.

22. Abels B, Villablanca JP, Tomandl BF, Uder M, Lell MM (2012) Acute stroke: A comparison of different CT perfusion algorithms and validation of ischaemic lesions by follow-up imaging. Eur Radiol 22: 2559-2567.

23. Lin K, Rapalino O, Law M, Babb JS, Siller KA, et al. (2008) Accuracy of the Alberta stroke program early ct score during the first 3 hours of middle cerebral artery stroke: Comparison of noncontrast CT, CT angiography source images, and ct perfusion. AJNR Am J Neuroradiol 29: 931-936.

24. Saur D, Kucinski T, Grzyska U, Eckert B, Eggers C, et al. (2003) Sensitivity and interrater agreement of CT and diffusion-weighted MR imaging in hyperacute stroke. AJNR Am J Neuroradiol 24: 878-85.

25. Schroeder J, Thomalla G (2016) A critical review of Alberta stroke program early ct score for evaluation of acute stroke imaging. Front Neurol 7: 245.

26. Abels B, Klotz E, Tomandl BF, Villablanca JP, Kloska SP, et al. (2011) CT perfusion in acute ischemic stroke: A comparison of 2-second and 1-second temporal resolution. AJNR Am J Neuroradiol 32: 1632-1639.

27. Riedel CH, Zoubie J, Ulmer S, Gierthmuehlen J, Jansen O (2012) Thin-slice reconstructions of nonenhanced CT images allow for detection of thrombus in acute stroke. Stroke 43: 2319-2323.

28. Feinstein AR, Cicchetti DV (1990) High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol 43: 543-549.

Henry Journal of Acupuncture & Traditional Medicine

Henry Journal of Anesthesia & Perioperative Management

Henry Journal of Aquaculture and Technical Development

Henry Journal of Cardiology & Cardiovascular Medicine

Henry Journal of Case Reports & Imaging

Henry Journal of Cell & Molecular Biology

Henry Journal of Tissue Biology & Cytology

Henry Journal of Clinical, Experimental and Cosmetic Dermatology

Henry Journal of Diabetes & Metabolic Syndrome

Henry Journal of Emergency Medicine, Trauma & Surgical Care

Henry Journal of Haematology & Hemotherapy

Henry Journal of Immunology & Immunotherapy

Henry Journal of Nanoscience, Nanomedicine & Nanobiology

Henry Journal of Nutrition & Food Science

Henry Journal of Obesity & Body Weight

Henry Journal of Cellular & Molecular Oncology

Henry Journal of Ophthalmology & Optometry

Henry Journal of Perinatology & Pediatrics