



Automated versus manual imaging assessment of early ischemic changes in acute stroke: comparison of two software packages and expert consensus

Friederike Austein¹ · Fritz Wodarg¹ · Nora Jürgensen¹ · Monika Huhndorf¹ · Johannes Meyne² · Thomas Lindner¹ · Olav Jansen¹ · Naomi Larsen¹ · Christian Riedel¹

Received: 11 October 2018 / Revised: 7 April 2019 / Accepted: 25 April 2019 / Published online: 10 May 2019
© European Society of Radiology 2019

Abstract

Aim The purpose of our study was to compare the agreement of both the total Alberta Stroke Program Early CT Score (ASPECTS) and region-based scores from two automated ASPECTS software packages and an expert consensus (EC) reading with final ASPECTS in a selected cohort of patients who had prompt reperfusion from endovascular thrombectomy (EVT).

Methods ASPECTS were retrospectively and blindly assessed by two software packages and EC on baseline non-contrast-enhanced computed tomography (NCCT) images. All patients had multimodal CT imaging including NCCT, CT angiography, and CT perfusion which demonstrated an acute anterior circulation ischemic stroke with a large vessel occlusion. Final ASPECTS on follow-up scans in patients who had EVT and achieved complete reperfusion within 100 min from NCCT served as ground truth and were compared to total and region-based scores.

Results Fifty-two patients met our study criteria. Moderate agreement was obtained between both software packages and EC for total ASPECTS and there was no significant difference in overall performance. However, the software packages differed with respect to regional contribution. In this cohort, the majority of infarcted regions were deep structures. Package A was more sensitive in cortical areas than the other methods, but at a cost of specificity. EC and software package B had greater sensitivity, but lower specificity for deep brain structures.

Conclusion In this cohort, using the final ASPECTS as ground truth, no clinically significant difference was observed for total ASPECT score between human or automated packages, but there were differences in the characteristics of the regions scored.

Key Points

- Some national stroke guidelines have incorporated ASPECTS in their recommendations for selecting patients for endovascular therapy.
- Computer-aided diagnosis is a promising tool to aid the evaluation of early ischemic changes identified on CT.
- Software packages for automated ASPECTS assessment differed significantly with respect to regional contribution without any significant difference in the overall ASPECT score.

Keywords Image interpretation · Computer assisted · Software validation · Stroke · Thrombectomy

Abbreviations

ASPECTS	Alberta Stroke Program Early CT Score
C	Caudate
CTP	CT perfusion
EC	Expert consensus
EIC	Early ischemic changes
EVT	Endovascular therapy
HU	Hounsfield units
I	Insula
IC	Internal capsule

✉ Friederike Austein
friederike.austein@uksh.de

¹ Department of Radiology and Neuroradiology, University Hospital, Schleswig-Holstein, Campus Kiel, Arnold-Heller-Str. 3, 24105 Kiel, Germany

² Department of Neurology, University Hospital, Schleswig-Holstein, Campus Kiel, Kiel, Germany

IQR	Interquartile range
L	Lentiform nucleus
NCCT	Non-contrast computed tomography
ROC	Receiver operating characteristic
TICI 3	Thrombolysis in cerebral infarction with complete reperfusion

Introduction

Recently, several of the randomized thrombectomy trials have used Alberta Stroke Program Early CT Score (ASPECTS) [1] for patient selection [2]. Some national stroke guidelines have incorporated ASPECTS in their recommendations for selecting patients for endovascular therapy (EVT) [3]. When an imaging-based selection approach is used clinically to select patients for EVT, ideally, it should demonstrate good inter-observer agreement. Unfortunately, the determination of early ischemic changes (EIC) and their translation into the ASPECTS have a considerable inter-rater variability [4, 5], which is influenced by rater experience. These studies tended to include a broad range of stroke severities and may not necessarily apply to the population eligible for EVT [4]. Computer-aided diagnosis is a promising tool to aid the evaluation of EIC identified on CT. The aim of this study was to evaluate the agreement of two CE-marked (Conformité Européenne) automated software solutions and the consensus of two neuroradiology experts (EC) based on a reference standard defined by final ASPECTS on follow-up imaging in those patients who had EVT and achieved prompt and complete reperfusion.

Methods

The study was approved by the local institutional review board and was prepared in accordance with the Guidelines for Reporting Reliability and Agreement Studies and the Standards for Reporting of Diagnostic Accuracy [6, 7].

Patient population

We retrospectively reviewed cases of patients having acute ischemic stroke who were treated with EVT between July 2013 and June 2017. Cases that met the following criteria were included: (1) occlusion of the internal carotid artery and/or middle cerebral artery; (2) complete recanalization corresponding to a TICI (thrombolysis in cerebral infarction) 3 score at the end of EVT; (3) time interval from CT to TICI 3 reperfusion < 100 min; and (4) availability of baseline NCCT head and a follow-up NCCT or MRI. Cases with significant motion artifacts were excluded from analysis. All patients had CT angiography and CT perfusion performed for confirmation of large vessel occlusion and perfusion deficit. Intravenous

thrombolysis was initiated in eligible patients. Clinical and imaging data were retrospectively analyzed.

ASPECTS readings by raters

All NCCT were retrospectively assessed by a consensus reading of two expert neuroradiologist (O.J. and F.W.), without access to other imaging studies or to clinical information and without time restraints. Modification of the window and level of the image contrast was allowed as needed. Contrary to the original ASPECT scoring system, which utilized only 2 brain sections [1], readers graded EIC in each of the 10 ASPECTS regions according to current methodology, which utilizes all images. EIC was defined as tissue hypoattenuation or loss of gray–white matter differentiation, as these changes have been associated with edema and irreversible injury. The 10 regions were divided in deep structures including caudate (C), lentiform nucleus (L), internal capsule (IC) and insula (I), and in superficial cortical areas including M1–M6.

Automated ASPECTS

e-ASPECTS software (Brainomix) is a decision support software that uses thin NCCT slices (1.0 mm) to perform the analysis. According to information provided by the company, the Brainomix software first resamples and standardizes the input DICOM data. Subsequently, voxel-wise early or non-acute signs of ischemia are identified using a machine learning classifier which has been trained on a large dataset (> 10,000 images) containing a wide range of clinical CTs from stroke patients and negative controls. After a patient-specific segmentation of the ASPECTS VOIs is generated, finally, the ASPECTS output score is calculated by classifying each VOI according to the results of the voxel-wise ischemia analysis. As manufacturers regularly release new software versions which claim improved performance, analysis was performed with both *e-ASPECTS* v6.1 and v7.1.

RAPID ASPECTS (iSchemaView) can analyze different slice thicknesses. Slice thicknesses of 2–3 mm are recommended as preferred dataset. Therefore, in each patient, a fully automated NCCT was processed on this software using thick slices (2.5 mm). To compare directly with *e-ASPECTS*, a further analysis using 1.0-mm slices was performed. According to information provided by the company iSchemaView, *RAPID ASPECTS* defines the 10 ASPECTS VOIs in both hemispheres, measures their mean Hounsfield units (HU), and calculates percentage HU differences between corresponding VOIs. Basing on these percentage HU differences, the VOIs are classified as ischemic or non-ischemic and the total ASPECT score is calculated. In order to avoid misinterpretation due to liquor or old infarctions, dark voxels are excluded from the mean HU measurements.

Both software tools are CE-marked.

Imaging protocol, CTP post processing, final ASPECTS

Imaging protocol

All CT scans were performed using a 64-slice CT scanner equipped with a 40-mm-wide detector (Philips Healthcare). NCCT scans of the head were performed in helical mode (0.65 mm thickness, kV 120, mA 250) and axial non-contrast CT images were reconstructed in 1.0-mm and 2.5-mm slice thickness applying the brain standard kernel with a fourth-generation iterative reconstruction algorithm (iDose level 2/filter UB). The imaging parameters for CTP were 80 kVp, 150 mAs, and 32×1.25 mm detector collimation and a scan duration of 60 s. After cerebral NCCT, CTP was conducted with the toggling table technique, allowing an extended coverage of the brain of 80 mm. A scan delay of 3 s was applied after injecting 60 mL (flow rate 5 mL/s) of iodinated contrast agent (350 mg I/mL Imeron 350, Bracco Imaging).

CTP image post processing and analysis

Perfusion analysis was performed with the RAPID™ software (iSchemaView).

Final ASPECTS

Final ASPECTS is the ASPECTS rated by consensus of two expert readers based at follow-up CT or MRI scans which were obtained between days 1 and 8. If both CT and MRI were performed, MRI data were used to assess the final ASPECTS.

Statistical analysis

Attribute agreement analysis (numerically equivalent to accuracy) for each single ASPECTS region was used to assess the agreement among the software packages, and EC with final ASPECTS as reference [8]. The agreement for total ASPECTS by multiple appraisers was measured by using the weighted Kappa statistic.

Additional assessments were made for each ASPECTS region to define the sensitivity, specificity, and accuracy of each software package and EC. The overall sensitivity, specificity, and accuracy with 95% confidence intervals in the cortical (M1–M6) and deep areas (IC, I, L, C) were evaluated for the software packages and EC. The sensitivity, specificity, and accuracy are generated from the default software outputs provided by the manufacturer.

For e-ASPECTS, the manufacturer reports the default operating point was set by selecting the most specific operating point that was non-inferior in both sensitivity and specificity

to an expert human scorer. For RAPID-ASPECTS, the default operating point in this analysis was the slider bar in the central position.

Score-based (total) receiver operating characteristic (ROC) curve analysis was performed using e-ASPECTS which allows generation of outputs using different operating points for sensitivity and specificity. RAPID does not output ROC curves; however, it allows adjusting its confidence level with a slider in the web interface. For this ROC analysis, the ground truth is positive for each ASPECTS point below 10, and negative for each ASPECTS point above 0, irrespective of which region is affected. As an example: if the physician scored 7, and the ground truth was 8, then this would result in 2 true positives, 1 false positive, and 7 true negatives. If the physician scored 8, and the ground truth was 7, this would result in 2 true positives, 1 false negative, and 7 true negatives. Weighted Kappa values were also calculated between each ASPECT score and the final ASPECTS. Continuous data with normal distribution were reported as mean \pm standard deviation; ordinal or non-normal data were reported as median and interquartile range (IQR). Categorical data were reported as proportions.

Statistical tests used to determine the significance of differences in variables are listed in the data tables and within the text where relevant statistical significance was set at $p < .05$. Statistical analysis was performed by using XLSTAT Version 2018.2.

Results

Demographic characteristics, clinical outcome, procedural characteristics and descriptive imaging findings, and ASPECTS assessment of the study patients

Retrospective analysis of our database identified 52 eligible patients. Patient demographics, clinical and imaging outcomes, procedural characteristics, and the ASPECTS assessments are summarized in Table 1.

Sixty-two percent of infarcted regions at follow-up were deep structures (Fig. 2) with the most common being the lentiform nucleus (19% of all infarcted regions, occurring in 60% of patients). Of the cortical regions involved, the most common region was the M2 region (10% of infarcted regions in 31% of patients).

Agreement of total and region-based ASPECTS

There was moderate agreement between EC, RAPID, and e-ASPECTS with the ground truth (Table 2). Attribute agreement analysis (accuracy) among operators and final

Table 1 Baseline demographic, clinical, and procedural characteristics. Imaging findings of patients with ICA or/ and MCA occlusion who achieved TIC1 3 recanalization

Number of patients	52
Male sex, <i>N</i> (%)	28 (54)
Age, years	76 (69–82)
NHSS ^(†)	14 (10–17)
Occlusion site, <i>N</i> (%)	
• ICA proximal	7 (13.5)
• ICA distal	0 (0)
• ICA terminus	6 (11.5)
• M1 segment, MCA	39 (75)
• M2 segment, MCA	11 (21.2)
Safe symptom onset, <i>N</i> (%)	30 (57.7)
Unknown symptom onset, <i>N</i> (%)	22 (42.3)
Symptom onset to NCCT, minutes ^(†)	66 (56–90)
NCCT to reperfusion TIC1 3, minutes ^(†)	79 (71–89)
Additional IV thrombolysis, <i>N</i> (%)	27 (51.9)
Patients receiving general anesthesia, <i>N</i> (%)	52 (100)
CT perfusion findings	
• Ischemic core volume, mL ^(†)	3 (0–13)
• Hypoperfusion Volume, mL ^(†)	101 (65–160)
• Mismatch Volume, mL ^(†)	92 (62–125)
Baseline ASPECTS ^(†)	
• Expert consensus reading	9 (8–10)
• RAPID ASPECTS	9 (7–10)
• e-ASPECTS	9 (8–10)
Final ASPECTS ^(†)	7 (9–6)
Follow-up CT, <i>N</i> (%)	39 (75.0)
Follow-up MRI, <i>N</i> (%)	13 (25.0)
Follow-up between days ^(†)	1 (1–3)

Data are displayed as median (interquartile range). Patients may have multiple occlusion sites so that total number rises up over 100%

mRS, modified Rankin scale; *NIHSS*, National Institutes of Health Stroke Scale; *CTP*, computed tomography perfusion; *ICA terminus*, internal carotid artery terminus; *MCA*, middle cerebral artery; *MRI*, magnetic resonance imaging; *TIC1 score*, thrombolysis in cerebral infarction

[†] Symbol means interquartile range

ASPECTS as reference showed no significant differences in percentage agreement for EC, RAPID-ASPECTS, and e-ASPECTS 7: 77% (CI 95% 73–80), 74% (70–78), and 72% (68–76) respectively (Table 2). The ASPECTS distribution within the cohort was skewed toward higher scores (smaller infarcts) predominantly in the deep regions, without significant differences in the median ASPECTS between automated and EC ratings (Figs. 1 and 2).

Superficial cortical areas were more often marked as positive with e-ASPECTS (69 with v6/312) than EC (26/312) or RAPID ASPECTS (31/312) and deep structures less often identified as positive with e-ASPECTS (20 with v6/208) than EC (52/208) or RAPID ASPECTS (60/208). The final ASPECTS detected 101 ischemic lesions in deep structures and 63 ischemic lesions in cortical areas, $p < 0.0001$ (chi-square statistic). Figure 2 shows total frequencies of EIC in deep and superficial cortical regions.

Sensitivity, specificity, and accuracy are shown for each region in Table 3. The overall sensitivity, specificity, and accuracy with 95% confidence intervals in the cortical (M1–M6) and deep areas (IC, I, L, C) can be seen in Table 4. Within the regional analyses, it can be seen that within the cortical regions, EC and e-ASPECTS tended to be more sensitive, but less specific than RAPID ASPECTS. Within deep regions, RAPID ASPECTS was more sensitive, but less specific.

ROC curve analysis

Both EC and RAPID ASPECTS are in close proximity or overlapping to the e-ASPECTS ROC curves with no significant difference between the software packages and the EC (Fig. 3). Areas under the ROC curves with confidence intervals and weighted Kappa values are presented in Table 2. There was no significant difference between any of the ROC curves. Using a bootstrapping technique to determine statistical significance, the p value between e-ASPECTS version 7 and RAPID 2.5 mm was 0.813.

Table 2 Agreement assessment with final ASPECTS as reference on follow-up imaging

	Attribute agreement	95% confidence intervals	Weighted Kappa	Area under the ROC curve	95% confidence intervals
EC ASPECTS	0.77	0.73, 0.80	0.26	–	–
RAPID ASPECTS 1.0 mm	0.72	0.68, 0.76	0.19	0.79	0.75, 0.84
RAPID ASPECTS 2.5 mm	0.74	0.70, 0.78	0.24	0.79	0.73, 0.82
e-ASPECTS v6	0.68	0.63, 0.72	0.19	0.81	0.77, 0.85
e-ASPECTS v7	0.72	0.68, 0.76	0.29	0.79	0.74, 0.83

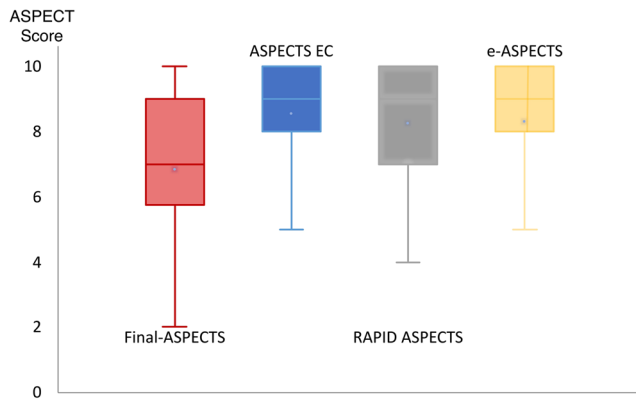


Fig. 1 Distribution of total ASPECTS results

Discussion

This is the first study that evaluated the total and region-based agreement of two different automated ASPECTS tools using final ASPECTS as the ground truth, based on follow-up imaging in patients who were promptly and successfully treated with EVT. A prior study that evaluated e-ASPECTS performance demonstrated that on average, this software was equivalent to expert neuroradiologists [9, 10]. Our findings support previous studies that demonstrated moderate agreement with the total ASPECTS for automated compared to human scorers [4].

A contributing factor for the substantial agreement between all methods for quantifying ASPECTS is the distribution of scores within this cohort which cluster around 8. This restricted distribution (i.e., smaller variance) results in less opportunity for potential discrepancies. Despite this, within the regional analysis of deep and cortical structures, we showed different sensitivities and specificities between humans and automated software packages. These differences are attributed to a different operating point used by the software packages and EC.

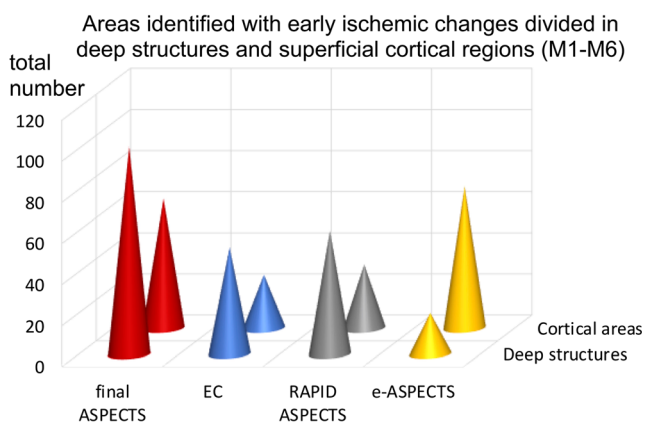


Fig. 2 Frequencies of early ischemic changes (EIC) on follow-up imaging and initial imaging by expert consensus (EC) and software packages of areas identified with EIC divided in deep structures (C, IC, L, I) and cortical regions (M1–M6)

It is notable that the ROC curve results are lower in this study than have been reported in previous studies of e-ASPECTS in different cohorts [10]. A likely explanation is the particular prevalence of damage in the ASPECTS regions within this study’s cohort. The most common finding in this cohort was infarction of the deep brain structures, in particular the lentiform nucleus, which was infarcted in 60% of patients at follow-up. RAPID ASPECTS and EC had higher sensitivity, but poorer specificity for detection of EIC in these deep structures. In contrast, e-ASPECTS showed lower sensitivity, but higher specificity. Areas of notable disagreement included the internal capsule and the insula. The internal capsule is known to be inconsistently scored among expert raters and has been the subject of low agreement in other studies [11, 12]. Further, the frequency with which the internal capsule was scored across methodologies varied from e-ASPECTS and EC which scored 0 and 4% involvement respectively to RAPID ASPECTS which scored involvement in 21% of cases.

The inverse pattern of sensitivity was observed in the cortical regions with e-ASPECTS scoring with greater sensitivity than EC and RAPID ASPECTS. Variability in cortical scoring may be due to the challenges of consistently defining anatomical borders for the superficial cortical regions (M1–M6). Despite this, identification of cortical infarction is important as these regions have been shown to have the greatest clinical eloquence, contribute to a greater proportion to the ASPECT score, and are therefore most likely to influence decision-making [13, 14]. It is possible that some discrepancy was observed as hypodensities that could be allocated to different cortical areas by different software programs or expert readers, resulting in discrepant region-based analysis [15]. This may underlay the increasing popularity of using volume measurements (such as a MRI- or CTP-defined ischemic core, or the e-ASPECTS CT infarct volume feature), although this approach may result in a loss of sensitivity to functional eloquence that can be provided by region-based scores.

Differences in slice thickness will result in different signal-to-noise ratios affecting the results, as was seen with the RAPID ASPECTS algorithm. In future studies, the influence of scan parameters should be explicitly investigated including the reconstruction algorithm as well as the slice thickness used for analysis. Previous work has demonstrated that machine algorithms tend to be less affected by reconstruction parameters than human raters [16]. The lower attribute agreement for the older version of e-ASPECTS demonstrates the improvements that are being made in machine learning, and the importance of using a consistent version of a software for analysis.

ASPECTS and e-ASPECTS have been shown to be a predictor of patient outcome following EVT [2, 17]. For automated software solutions to be clinically useful, not only is accuracy critical but also the presentation and traceability of the

Table 3 Specificity, sensitivity, and accuracy for each ASPECTS region

		Specificity	Sensitivity	Accuracy
M1	EC	0.95 (0.84, 0.99)	0.22 (0.03, 0.60)	0.83 (0.70, 0.92)
	RAPID ASPECTS 1.0 mm	0.86 (0.72, 0.95)	0.22 (0.03, 0.60)	0.75 (0.61, 0.86)
	RAPID ASPECTS 2.5 mm	0.93 (0.81, 0.99)	0.22 (0.03, 0.60)	0.81 (0.67, 0.90)
	e-ASPECTS v6	0.77 (0.61, 0.88)	0.33 (0.07, 0.70)	0.69 (0.55, 0.81)
	e-ASPECTS v7	0.98 (0.88, 1.00)	0.11 (0.00, 0.48)	0.83 (0.70, 0.92)
M2	EC	0.92 (0.78, 0.98)	0.38 (0.15, 0.65)	0.75 (0.61, 0.86)
	RAPID ASPECTS 1.0 mm	0.92 (0.78, 0.98)	0.19 (0.04, 0.46)	0.69 (0.55, 0.81)
	RAPID ASPECTS 2.5 mm	0.94 (0.81, 0.99)	0.25 (0.07, 0.52)	0.73 (0.59, 0.84)
	e-ASPECTS v6	0.67 (0.49, 0.81)	0.38 (0.15, 0.65)	0.58 (0.43, 0.71)
	e-ASPECTS v7	0.78 (0.61, 0.90)	0.31 (0.11, 0.59)	0.63 (0.49, 0.76)
M3	EC	0.95 (0.83, 0.99)	0.09 (0.00, 0.41)	0.77 (0.63, 0.87)
	RAPID ASPECTS 1.0 mm	0.88 (0.74, 0.96)	0.18 (0.02, 0.52)	0.73 (0.59, 0.84)
	RAPID ASPECTS 2.5 mm	0.90 (0.77, 0.97)	0.27 (0.06, 0.61)	0.77 (0.63, 0.87)
	e-ASPECTS v6	0.90 (0.77, 0.97)	0.27 (0.06, 0.61)	0.77 (0.63, 0.87)
	e-ASPECTS v7	0.98 (0.87, 1.00)	0.09 (0.00, 0.41)	0.79 (0.65, 0.89)
M4	EC	0.96 (0.85, 0.99)	0.14 (0.00, 0.58)	0.85 (0.72, 0.93)
	RAPID ASPECTS 1.0 mm	0.96 (0.85, 0.99)	0.14 (0.00, 0.58)	0.85 (0.72, 0.93)
	RAPID ASPECTS 2.5 mm	0.96 (0.85, 0.99)	0.00 (0.00, 0.41)	0.83 (0.70, 0.92)
	e-ASPECTS v6	0.91 (0.79, 0.98)	0.29 (0.04, 0.71)	0.83 (0.70, 0.92)
	e-ASPECTS v7	0.93 (0.82, 0.99)	0.57 (0.18, 0.90)	0.88 (0.77, 0.96)
M5	EC	0.95 (0.83, 0.99)	0.25 (0.05, 0.57)	0.79 (0.65, 0.89)
	RAPID ASPECTS 1.0 mm	0.95 (0.83, 0.99)	0.00 (0.00, 0.26)	0.73 (0.59, 0.84)
	RAPID ASPECTS 2.5 mm	0.92 (0.80, 0.98)	0.25 (0.05, 0.57)	0.77 (0.63, 0.87)
	e-ASPECTS v6	0.75 (0.59, 0.87)	0.58 (0.28, 0.85)	0.71 (0.57, 0.83)
	e-ASPECTS v7	0.88 (0.73, 0.96)	0.33 (0.10, 0.65)	0.75 (0.61, 0.86)
M6	EC	0.98 (0.88, 1.00)	0.12 (0.00, 0.53)	0.85 (0.72, 0.93)
	RAPID ASPECTS 1.0 mm	0.93 (0.81, 0.99)	0.12 (0.00, 0.53)	0.81 (0.67, 0.90)
	RAPID ASPECTS 2.5 mm	0.91 (0.78, 0.97)	0.12 (0.00, 0.53)	0.79 (0.65, 0.89)
	e-ASPECTS v6	0.86 (0.73, 0.95)	0.25 (0.03, 0.65)	0.77 (0.63, 0.87)
	e-ASPECTS v7	0.98 (0.88, 1.00)	0.38 (0.09, 0.76)	0.88 (0.77, 0.96)
Lentiform	EC	1.00 (0.84, 1.00)	0.58 (0.39, 0.75)	0.75 (0.61, 0.86)
	RAPID ASPECTS 1.0 mm	1.00 (0.84, 1.00)	0.35 (0.19, 0.55)	0.62 (0.47, 0.75)
	RAPID ASPECTS 2.5 mm	0.95 (0.76, 1.00)	0.32 (0.17, 0.51)	0.58 (0.43, 0.71)
	e-ASPECTS v6	1.00 (0.84, 1.00)	0.29 (0.14, 0.48)	0.58 (0.43, 0.71)
	e-ASPECTS v7	0.95 (0.76, 1.00)	0.35 (0.19, 0.55)	0.60 (0.45, 0.73)
Insular	EC	0.86 (0.67, 0.96)	0.54 (0.33, 0.74)	0.71 (0.57, 0.83)
	RAPID ASPECTS 1.0 mm	0.71 (0.51, 0.87)	0.50 (0.29, 0.71)	0.62 (0.47, 0.75)
	RAPID ASPECTS 2.5 mm	0.82 (0.63, 0.94)	0.62 (0.41, 0.81)	0.73 (0.59, 0.84)
	e-ASPECTS v6	0.96 (0.82, 1.00)	0.08 (0.01, 0.27)	0.56 (0.41, 0.70)
	e-ASPECTS v7	0.86 (0.67, 0.96)	0.29 (0.13, 0.51)	0.60 (0.45, 0.73)
Caudate	EC	0.96 (0.80, 1.00)	0.52 (0.32, 0.71)	0.73 (0.59, 0.84)
	RAPID ASPECTS 1.0 mm	0.92 (0.74, 0.99)	0.37 (0.19, 0.58)	0.63 (0.49, 0.76)
	RAPID ASPECTS 2.5 mm	0.84 (0.64, 0.95)	0.52 (0.32, 0.71)	0.67 (0.53, 0.80)
	e-ASPECTS v6	1.00 (0.86, 1.00)	0.30 (0.14, 0.50)	0.63 (0.49, 0.76)
	e-ASPECTS v7	0.96 (0.80, 1.00)	0.26 (0.11, 0.46)	0.60 (0.45, 0.73)
Internal capsule	EC	0.97 (0.84, 1.00)	0.05 (0.00, 0.26)	0.63 (0.49, 0.76)
	RAPID ASPECTS 1.0 mm	1.00 (0.89, 1.00)	0.37 (0.16, 0.62)	0.77 (0.63, 0.87)
	RAPID ASPECTS 2.5 mm	0.94 (0.80, 0.99)	0.47 (0.24, 0.71)	0.77 (0.63, 0.87)
	e-ASPECTS v6	1.00 (0.86, 1.00)	0.30 (0.14, 0.50)	0.63 (0.49, 0.76)
	e-ASPECTS v7	0.96 (0.80, 1.00)	0.26 (0.11, 0.46)	0.60 (0.45, 0.73)

results. Automated ASPECTS tools are designed as decision support rather than standalone diagnostics. Their utility relies not only on their absolute score, but the ability of the clinician user to see the results and adjust the default score according to their clinical interpretation. Both RAPID ASPECTS and e-ASPECTS visually present results in all slices and both software solutions analyze the whole MCA territory of the brain. e-ASPECTS generates a heatmap which marks that portion of

each region that shows EIC, allowing the clinician to visually inspect and review the results on a voxel basis. RAPID ASPECTS marks the whole area, even if only part of the region is affected. e-ASPECTS provides volumes of acute and non-acute ischemic areas, whereas RAPID ASPECTS provides Hounsfield units values for each region.

The study has some limitations. Only patients who were EVT candidates and had prompt and complete reperfusion

Table 4 Sensitivity, specificity, and accuracy across all cortical and all deep regions

	Specificity	Sensitivity	Accuracy
Cortical regions (N = 364)			
EC ASPECTS	0.96 (0.92, 0.98)	0.34 (0.25, 0.45)	0.80 (0.75, 0.84)
RAPID ASPECTS 1.0 mm	0.92 (0.88, 0.95)	0.21 (0.14, 0.31)	0.74 (0.69, 0.78)
RAPID ASPECTS 2.5 mm	0.93 (0.89, 0.96)	0.24 (0.16, 0.34)	0.75 (0.71, 0.80)
e-ASPECTS v6	0.83 (0.78, 0.87)	0.34 (0.25, 0.45)	0.70 (0.65, 0.75)
e-ASPECTS v7	0.93 (0.89, 0.95)	0.31 (0.22, 0.41)	0.77 (0.72, 0.81)
Deep regions (N = 156)			
EC ASPECTS	0.93 (0.85, 0.97)	0.40 (0.28, 0.52)	0.69 (0.61, 0.76)
RAPID ASPECTS 1.0 mm	0.88 (0.80, 0.94)	0.41 (0.30, 0.54)	0.67 (0.59, 0.75)
RAPID ASPECTS 2.5 mm	0.87 (0.78, 0.93)	0.54 (0.42, 0.66)	0.72 (0.65, 0.79)
e-ASPECTS v6	0.99 (0.94, 1.00)	0.14 (0.07, 0.25)	0.61 (0.53, 0.69)
e-ASPECTS v7	0.94 (0.87, 0.98)	0.20 (0.11, 0.31)	0.61 (0.53, 0.69)

were included, which limited the number of eligible patients and biased the cohort to patients with favorable ASPECTS ratings (patients with unfavorable scores are not treated with EVT). Therefore, our results may not apply to patients with lower ASPECT scores, which is critical when assessing the utility of the packages for patient selection for EVT. Furthermore, due to both the hyperacute nature of the imaging cohort and the time delay of up to 100 min between baseline imaging reperfusion, it is likely that some additional evolution of the infarct occurred, which accounts for the relative insensitivity of all methods tested.

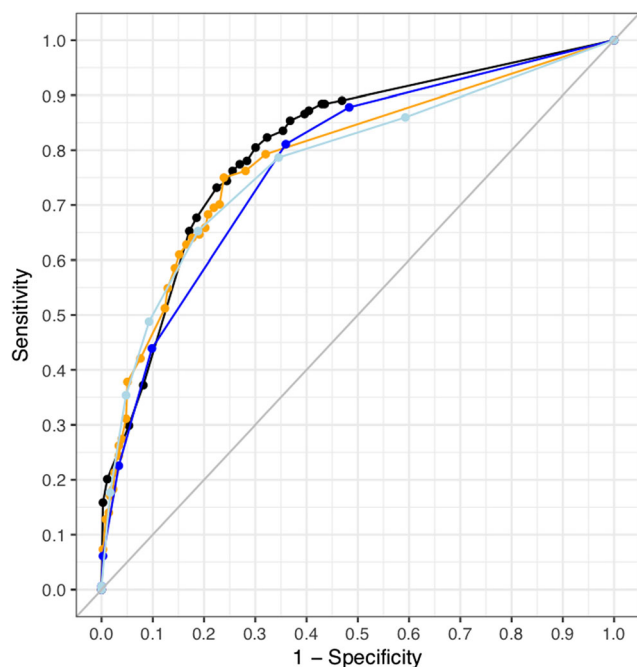


Fig. 3 Score-based (total) receiver operating characteristic (ROC) curve using multiple operating points (black: e-ASPECTS v6; orange: e-ASPECTS v7; dark blue: RAPID ASPECTS 1.0 mm; light blue: RAPID ASPECTS 2.5 mm)

Conclusions

Good agreement between both software packages and EC was obtained for total ASPECTS compared to ground truth. The packages differed with respect to regional contributions, without any significant difference in performance and without any implications in the clinical decision-making in this cohort. Fully automated ASPECT scoring is not designed to be used as a standalone tool clinically and both products in this study are intended for use as a decision support tool. Multidisciplinary neuroradiologic and neurologic expertise will always be required, but automated tools may facilitate decision-making. Automated ASPECTS can be used to assist decision-making, but other examination results, such as CTA, CTP, and clinical parameters, must also be considered. Automated ASPECTS tools have the potential to improve standardization and inter-rater agreement in both research and clinical practice, especially when images are being read by less-experienced readers.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Olav Jansen.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise.

No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology

- Retrospective
- Diagnostic study or prognostic study
- Performed at one institution

References

1. Barber PA, Demchuk AM, Zhang J, Buchan AM, Group AS (2000) Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. *Lancet* 355:1670–1674
2. Goyal M, Menon BK, Van Zwam WH et al (2016) Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet* 387:1723–1731
3. Powers WJ, Derdeyn CP, Biller J et al (2015) 2015 American Heart Association/American Stroke Association focused update of the 2013 guidelines for the early management of patients with acute ischemic stroke regarding endovascular treatment: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 46:3020–3035
4. Farzin B, Fahed R, Guilbert F et al (2016) Early CT changes in patients admitted for thrombectomy intrarater and interrater agreement. *Neurology* 87:249–256
5. Wilson AT, Dey S, Evans JW, Najm M, Qiu W, Menon BK (2018) Minds treating brains: understanding the interpretation of non-contrast CT ASPECTS in acute ischemic stroke. *Expert Rev Cardiovasc Ther* 16:143–153
6. Kottner J, Audigé L, Brorson S et al (2011) Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Int J Nurs Stud* 48:661–671
7. Bossuyt PM, Reitsma JB, Bruns DE et al (2003) The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 138:W1–W12
8. Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86:420
9. Herweh C, Ringleb PA, Rauch G et al (2016) Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients. *Int J Stroke* 11:438–445
10. Nagel S, Sinha D, Day D et al (2017) e-ASPECTS software is non-inferior to neuroradiologists in applying the ASPECT score to computed tomography scans of acute ischemic stroke patients. *Int J Stroke* 12:615–622
11. Gupta AC, Schaefer PW, Chaudhry ZA et al (2012) Interobserver reliability of baseline noncontrast CT Alberta Stroke Program Early CT Score for intra-arterial stroke treatment selection. *AJNR Am J Neuroradiol* 33:1046–1049
12. Pexman JH, Barber PA, Hill MD 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
13. Khan M, Baird GL, Goddeau RP Jr, Silver B, Henninger N (2017) Alberta Stroke Program Early CT Score infarct location predicts outcome following M2 occlusion. *Front Neurol* 8:98
14. Sheth SA, Malhotra K, Liebeskind DS et al (2018) Regional contributions to poststroke disability in endovascular therapy. *Interv Neurol* 7:533–543
15. van der Zwan A, Hillen B, Tulleken CA, Dujovny M, Dragovic L (1992) Variability of the territories of the major cerebral arteries. *J Neurosurg* 77:927–940
16. Seker F, Pfaff J, Nagel S et al (2018) CT reconstruction levels affect automated and reader-based ASPECTS ratings in acute ischemic stroke. *J Neuroimaging*. <https://doi.org/10.1111/jon.12562>
17. Pfaff J, Herweh C, Schieber S et al (2017) e-ASPECTS correlates with and is predictive of outcome after mechanical thrombectomy. *AJNR Am J Neuroradiol* 38:1594–1599

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.