Radiology

Automated Calculation of the Alberta Stroke Program Early CT Score: Feasibility and Reliability

Christian Maegerlein, MD • Johanna Fischer, MD • Sebastian Mönch, MD • Maria Berndt, MD • Silke Wunderlich, MD • Christian L. Seifert, MD • Manuel Lehm, MD • Tobias Boeckh-Behrens, MD • Claus Zimmer, MD • Benjamin Friedrich, MD

From the Department of Diagnostic and Interventional Neuroradiology (C.M., J.F., S.M., M.B., M.L., T.B.B., C.Z., B.F.) and Department of Neurology (S.W., C.L.S.), Klinikum rechts der Isar, Technical University Munich, Ismaninger Str 22, 81675 Munich, Germany. Received May 25, 2018; revision requested July 12; final revision received December 1; accepted December 12. **Address correspondence to** C.M. (e-mail: *christian.maegerlein@tum.de*).

Conflicts of interest are listed at the end of this article.

Radiology 2019; 291:141–148 • https://doi.org/10.1148/radiol.2019181228 • Content code: NR

Background: The Alberta Stroke Program Early CT Score (ASPECTS) evaluation is a qualitative method to evaluate focal hypoattenuation at brain CT in early acute stroke. However, interobserver agreement is only moderate.

Purpose: To compare ASPECTS calculated by using an automatic software tool to neuroradiologist evaluation in the setting of acute stroke.

Materials and Methods: For this retrospective study, consensus ASPECTS were defined by two neuroradiologists based on baseline noncontrast CTs collected from January 2017 to December 2017 from patients with an occlusion in the middle cerebral artery and from an additional cohort of patients suspected of having stroke and no large vessel occlusion. Imaging data from both baseline and follow-up CT was evaluated for the consensus reading. After 6 weeks, the same two neuroradiologists again determined ASPECTS by using only the baseline CT. For comparison, ASPECTS was also calculated from baseline CT images by using a commercially available software (RAPID ASPECTS). Both methods were compared by using weighted κ statistics.

Results: CT scans from 100 patients with middle cerebral artery occlusion (44 women [mean age ± standard deviation, 75 years ± 14] and 56 men [mean age, 71 years ± 14]) and 52 patients suspected of having stroke and no large vessel occlusion (19 women [mean age, 69 years ± 18] and 33 men [68 years ± 15]) were evaluated. Neuroradiologists showed moderate agreement with the consensus score (κ = 0.57 and κ = 0.56). Software analysis showed substantial agreement (κ = 0.9) with the consensus score. Software analysis showed a substantial agreement (κ = 0.78) after greater than 1 hour between symptom onset and imaging, which increased to high agreement (κ = 0.92) in the time window greater than 4 hours. The neuroradiologist raters did not achieve comparable results to the software until the time interval of greater than 4 hours (κ = 0.83 and κ = 0.76).

Conclusion: In acute stroke of the middle cerebral artery, the Alberta Stroke Program Early CT score calculated with automated software had better agreement than that of human readers with a predefined consensus score.

Online supplemental material is available for this article.

The intravenous administration of recombinant tissue plasminogen activator (or rtPA) was introduced into acute stroke therapy in the mid-1990s (1,2). The interpretation of early infarct signs in CT then became clinically relevant for the first time, as it was shown that the response to rtPA could be predicted based on the degree of initial infarct demarcation (2,3). Evidence also indicated that intravenous administration of rtPA can be harmful in patients with advanced early infarct signs due to a higher risk of intracranial hemorrhage (1,2,4). However, only rough estimations of the degree of early infarct signs were performed. One rule often applied for patient selection and subsequently used for mechanical thrombectomy was that early infarct signs should be confined to less than one-third of the middle cerebral artery territory (3,4).

Beginning in the 2000s, a more detailed analysis of early infarct signs was proposed as the Alberta Stroke Program Early CT score (ASPECTS) (5–9). According to this concept, segmental assessment of 10 defined middle cerebral artery vascular territories is performed to look for focal hypoattenuation of the cortex and in the basal ganglia, the reduction of gray and white matter differentiation, and the loss of the insular ribbon sign (10,11). The final ASPECTS evaluation is calculated by subtracting one point for each region with early infarct signs from the maximum score of 10.

The ASPECTS evaluation found increasing acceptance after mechanical thrombectomy was found to be effective for treatment of patients with an emergent large vessel occlusion even beyond time windows applicable for intravenous administration of rtPA (12,13). For some of the large randomized controlled trials that ultimately led to the establishment of thrombectomy as a standard procedure, an ASPECTS evaluation greater than or equal to 6 was used as an inclusion criterion. For this reason, a minimum ASPECTS evaluation is included in most national and international thrombectomy guidelines. For example, the guideline from the American Heart Association provides level IA evidence for thrombectomy in patients in the early window period

Abbreviation

ASPECTS = Alberta Stroke Program Early CT score

Summary

In candidates for thrombectomy, the Alberta Stroke Program Early CT score calculated with automated software showed better agreement with a predefined consensus score than with expert human readers, especially for scans obtained between 1–4 hours from symptom onset.

Key Points

- Two experienced neuroradiologists showed moderate agreement with a predefined Alberta Stroke Program Early CT score (ASPECTS) consensus score (κ = 0.57 and κ = 0.56).
- When automated software was used to calculate ASPECTS, there was substantial agreement (κ = 0.90) with the predefined consensus score ASPECTS evaluation.
- In acute stroke of the middle cerebral artery, ASPECTS by using automated software had better agreement with a predefined consensus score than did experienced neuroradiologists.

(<6 hours) with an ASPECTS evaluation greater than or equal to 6 (14).

Thus, the reliable quantification of the ASPECTS evaluation is of clinical relevance but suffers from the drawback of limited interobserver agreement (5,15). Two commercially available software programs are now available to automatically calculate the ASPECTS evaluation in a reasonable and clinically acceptable timeframe. Studies using the e-ASPECTS software (Brainomix, Oxford, United Kingdom) have reported encouraging results (16,17). The aim of our study was to evaluate another automated software–based analysis (RAPID ASPECTS) of the ASPECTS evaluation in comparison with expert neuroradiologists. This study was exploratory without a prespecified hypothesis.

Materials and Methods

For the main analysis, imaging data from consecutive patients who presented to our institution with an acute stroke between January 1, 2017 and December 31, 2017 due to emergent large vessel occlusion in the middle cerebral artery and were treated with thrombectomy were included in our study (cohort 1, *n* = 100). Additionally, a second cohort of consecutive patients with clinical suspicion of stroke and no emergent large vessel occlusion between January 1, 2017 and March 31, 2017 was analyzed (cohort 2, *n* = 52). Disease categories of cohort 2 are specified in Table E1 (online). To be included in the analysis, only patients with follow-up MRI were included for both cohorts. Based on the anonymous and retrospective study design, written consent was waived by the local ethics committee. This is an investigator-initiated study without any financial support.

Image Analysis by Neuroradiologists

To assign a reference standard ASPECTS evaluation, two board-certified neuroradiology attending physicians (C.M. and B.F., with 9 years and 7 years of experience, respectively) independently reviewed the complete set of available imaging at the acute stage, including noncontrast CT (5 mm, axial, incremental acquisition), perfusion CT (10 mm, axial), digital subtraction angiography imaging, and follow-up MRI (three-dimensional fluid-attenuated inversion recovery [repetition time, 4800 msec; echo time, 289 msec; inversion time, 1650 msec], diffusion-tensor imaging with 15 directions [repetition time, 9895 msec; echo time, 55 msec; *b* value, 1000 sec/mm^2]) obtained 3–5 days after treatment.

Neuroradiology readers were blinded to the results of the automated software evaluation (described next). Differences in assigned ASPECTS evaluation for the consensus reading were resolved through joint review of the imaging and discussion with a unanimous decision.

Six weeks later, the same two neuroradiologists independently assigned ASPECTS evaluations by using only the noncontrast baseline CT obtained at the time of each patient's presentation. For this analysis, only the hemisphere affected by the stroke was known to the readers. All other information regarding vessel occlusion, time metrics, and the follow-up imaging was withheld.

Image Analysis by Using Automated Software

For comparison, an automated software tool (RAPID ASPECTS, version 4.9; iSchemaView, Menlo Park, Calif) was used to calculate an ASPECTS score. RAPID ASPECTS software performs a series of operations to generate an automated ASPECTS evaluation. These operations include the following: *(a)* importing the Digital Imaging and Communications in Medicine data from the noncontrast CT; *(b)* tilt correcting the image along with removal of the skull base, calvarium, and cerebrospinal fluid spaces; *(c)* applying a standardized atlas to create an individualized grid that corresponds to the 10 ASPECTS regions on each hemisphere; *(d)* calculating the Hounsfield unit values and other relevant parameters for each of the 20 regions; *(e)* classifying each region as either normal or abnormal by using a machine learning–based algorithm and assessing which hemisphere is most likely to be the involved hemisphere; *(f)* applying plausibility checks and confidence thresholding; *(g)* generating an ASPECTS output map on a graphical user interface with the involved regions identified in red and reporting an ASPECTS evaluation; *(h)* sending a deidentified image of the output map to the picture archiving and communication system (Fig 1). More detailed information on how the software determines the ASPECTS evaluation is provided in Appendix E1 (online). No further machine learning or training of the software occurred during the study period.

Statistical Analysis

All statistical analyses were performed by using SPSS (version 23; IBM, Armonk, NY). Agreement between the different readers on the baseline ASPECTS evaluation was calculated by using a square-weighted κ. The weighted κ is calculated by using a predefined table of weights that measure the degree of disagreement between the two raters, so that the higher the disagreement, the higher the weight (ie, the farther apart are the judgments, the higher the weights assigned). A bootstrap procedure using 4000 bootstrap samples was performed to estimate 95% confidence intervals for the differences of weighted κ and to perform statistical tests on equality of weighted κ values (18). Agreement between the different readers on dichotomized ASPECTS evaluations were calculated by using Cohen κ. Correlations were calculated by using Spearman ρ. All data

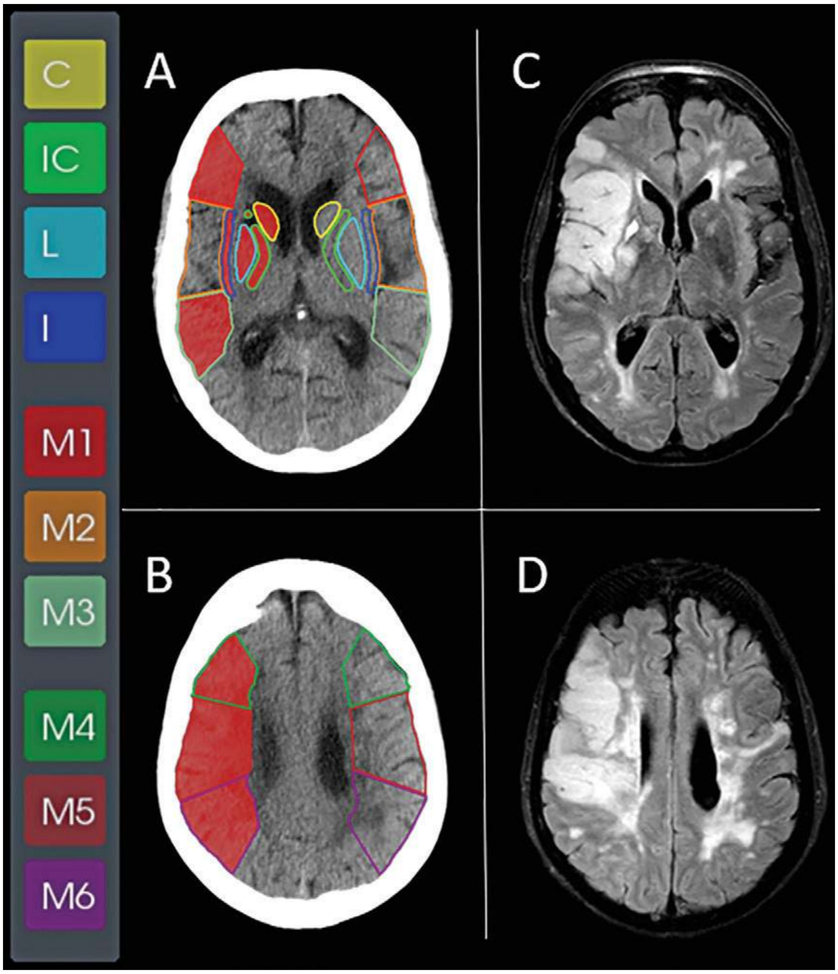

Figure 1: *A,B,* Axial CT images show territories with automatically calculated early infarct signs by using RAPID ASPECTS software (iSchemaView, Menlo Park, Calif) demonstrated in red. Automatically calculated ASPECTS evaluation was 1, while readers indicated scores of 7 and 8, respectively. Consensus ASPECTS evaluation was 2, differing in only one ASPECTS territory from software calculation. *C,D,* Axial MR images show infarct demarcation in fluid-attenuated inversion recovery 3 days after complete reperfusion during mechanical thrombectomy, which was achieved 150 minutes after symptom onset. Readers' scores were performed based on standard noncontrast CTs in picture archiving and communication system, unaware of both automated segmentation as well as automated scoring by using RAPID ASPECTS software. To compute ASPECTS, 1 point is subtracted from 10 for any evidence of early ischemic change for each of 10 defined middle cerebral artery (MCA) vascular territories. C = caudate head, I = insula, IC = internal capsule, L = lentiform nucleus, M1 = frontal operculum, M2 = anterior temporal lobe, M3 = posterior temporal lobe, M4 = anterior MCA, M5 = lateral MCA, M6 = posterior MCA.

are presented as medians with confidence intervals if not indicated otherwise. A P value $\leq$ .05 was considered to indicate statistical significance.

Results

Demographics of the Study Cohort

Baseline characteristics for cohort 1 and cohort 2 are shown in Table 1 and Table 2, respectively. After excluding patients for whom no adequate follow-up MRI was available (18 patients in cohort 1, 21 patients in cohort 2) or whose baseline images were technically inadequate to be analyzed by the human readers (three patients in cohort 1) or could not be processed with the software (22 patients in cohort 1, 10 patients in cohort 2), a total of 100 patients in cohort 1 and 52 patients in cohort 2 were included in our study (Fig 2).

The consensus score analysis showed a median ASPECTS evaluation of 9 (95% confidence interval: 7, 10; range, 1–10). Dichotomized using a threshold of greater than or equal to 6, there were 14 patients in cohort 1 with ASPECTS less than 6 and 86 patients with ASPECTS greater than or equal to 6. In cohort 2, there was 1 patient with ASPECTS less than 6 and 51 patients with ASPECTS greater than or equal to 6.

Comparison of Neuroradiologist versus Software Evaluation of ASPECTS

In the blinded analysis of cohort 1, reader 1, reader 2, and software analysis each showed a median ASPECTS evaluation of 9 with different interquartile ranges (reader 1, 9–10; reader 2, 8–10; software evaluation of ASPECTS, 6–10). Reader 1 and reader 2 showed only a fair agreement with the consensus score in the subsequent analysis of the ASPECTS evaluation (reader 1, κ = 0.57; reader 2, κ = 0.56; P = .90 for reader 1 vs reader 2).

Automated software analysis showed almost perfect agreement to the consensus score with a κ value of 0.9 (P = .05 for automated software vs reader 1; P = .006 for automated software vs reader 2) (Table 2, Fig 3). In cohort 2, automated software agreement with the consensus score was κ value of 0.6; reader 1 reached a κ value of 0.3 and reader 2 reached a κ value of 0.6. There was a correlation between the consensus score ASPECT and the scores of reader 1 and reader 2: lower ASPECT consensus score correlated with lower agreement between the neuroradiologists (both P < .001). In contrast, automated software performance was independent of agreement with consensus score (P = .16) (Fig 4). The clinically relevant threshold of an ASPECTS evaluation of 6 or greater versus less than 6 also showed substantial differences: automated software analysis showed moderate agreement (κ = 0.7) in comparison with the consensus score. Reader 1 and reader 2 showed fair and moderate agreement, respectively (κ = 0.4 and κ = 0.5), when compared with the consensus score.

Relationship between Symptom Onset to Imaging and ASPECTS

The average time from symptom onset to imaging in our cohort was 121 minutes ± 89 (standard deviation) for cohort 1. Both human readers and the automated software showed mini-

Table 1: Patient Characteristics of Cohort 1

Characteristic	Result
No. of patients	100
Age (y)	75 ± 14
Women	75 ± 14
Men	71 ± 14
Female sex (%)	44
National Institutes of Health Stroke Scale*†	13 (8, 17)
Time from symptom onset to imaging (min)	121 ± 89
90-day modified Rankin Scale*‡	3 (1, 6)

Note.—Unless otherwise indicated, data are means ± standard deviation. Cohort 1 included patients with acute stroke due to large vessel occlusion in the middle cerebral artery territory.

* Data are medians, with confidence intervals in parentheses.

† Score of 0 indicates no stroke symptoms; score of 1–4 indicates minor stroke; score of 5–15 indicates moderate stoke; score of 16–20 indicates moderate to severe stroke; score of 21–42 indicates severe stroke.

‡ Score of 0 indicates no symptoms; score of 1 indicates no significant disability; score of 2 indicates slight disability; score of 3 indicates moderate disability; score of 4 indicates moderately severe disability; score of 5 indicates severe disability; score of 6 indicates the patient is dead.

Table 2: Patient Characteristics of Cohort 2

Characteristic	Result
No. of patients	52
Age (y)	73 ± 16
Women	69 ± 18
Men	68 ± 15
Female sex (%)	36
National Institutes of Health Stroke Scale*†	8 (1, 12)
Time from symptom onset to imaging (min)	134 ± 79

Note.—Unless otherwise indicated, data are means ± standard deviation. Cohort 2 included patients with clinical suspicion of stroke but no large vessel occlusion in the middle cerebral artery territory.

* Data are medians, with confidence intervals in parentheses.

† Score of 0 indicates no stroke symptoms; score of 1–4 indicates minor stroke; score of 5–15 indicates moderate stoke; score of 16–20 indicates moderate to severe stroke; score of 21–42 indicates severe stroke.

mal agreement with the consensus score when less than 1 hour passed between symptom onset and imaging (reader 1, κ = 0.19; reader 2, κ = 0.12; software analysis, κ = 0.17).

With increasing time between symptom onset and imaging, better agreement between the two readers and the consensus score was observed. If the time between symptom onset and imaging was more than 4 hours, then both readers showed substantial agreement (κ = 0.83 and κ = 0.76, respectively) (Tables 3, 4). Automated software analysis, however, showed substantial agreement (κ = 0.78) after the first hour, which subsequently increased to an almost perfect agreement (κ = 0.92) at the time window of more than 4 hours (Fig 5). No time dependency could be detected in cohort 2.

Relationship between Treatment and ASPECTS

Based on treatment guidelines, we would not have treated seven of our 100 patients using software analysis as a basis for decision making. Among these seven patients (all of whom achieved successful endovascular reperfusion), one patient was functionally independent with a modified Rankin Scale (or mRS) score of 1 after 90 days, one patient had an mRS score of 3 after 90 days, one patient had an mRS score of 4 after 90 days, and four patients died. The patient with the functionally independent outcome is particularly interesting. The automated software assigned an ASPECTS evaluation of 1, the consensus score was an ASPECTS evaluation of 6 at diffusion-weighted imaging, and the two human readers each scored 8. This patient was treated with thrombectomy and had an ASPECTS evaluation of 6 in a patchy appearance and a favorable clinical outcome (Fig 6). Because the hypodensities in the acute stroke CT can be clearly visualized and were confirmed with Hounsfield unit measurement (see software attenuation values in Fig 4), it appears that some of these hypodense areas were indeed reversible after successful recanalization.

Discussion

In our study, we confirmed that experienced neuroradiologists may show only a fair agreement (κ = 0.56–0.57) in ASPECTS evaluation, either between each other or in comparison to a consensus score that incorporates the results of follow-up imaging. The interrater variability can have clinical consequences, as guideline-based endovascular treatment of the patient is often based on the ASPECTS evaluation. We used an automated software approach that determines the ASPECTS evaluations in a reasonable computing time (2–4 minutes) and therefore hypothesized to standardize and improve the accuracy of ASPECT scoring. Our results demonstrate that a fully automated software tool can determine the ASPECTS evaluation with greater agreement in comparison to a predefined consensus standard of reference (κ = 0.92) than an experienced neuroradiologist in patients with an acute occlusion of the middle cerebral artery.

Various studies have shown only modest to moderate agreement between two or more readers over the entire scale for determining ASPECTS evaluations (19,20). Therefore, the ASPECTS evaluation was frequently dichotomized in the literature (eg, more or less than an ASPECTS evaluation of 7) (7). Moderate to good interrater agreement could be reached with these dichotomized ASPECTS evaluation provisions.

In contrast to human readers, the software showed almost perfect concordance with the previously defined consensus score that allowed data from follow-up images to be considered. The time dependency was particularly striking. In patients imaged less than 1 hour since the onset of symptoms, both the human reader and the software showed poor agreement (κ = 0.12–0.19) with the predefined consensus score, likely because early infarct changes are typically extremely subtle or absent in the first hour after symptom onset. However, beyond 1 hour, the software-based analysis showed good detection of the infarcted area as demonstrated by a high agreement with the consensus score (κ = 0.78), while the human readers showed poor to satisfactory agreement (κ = 0.27–0.36). For patients in

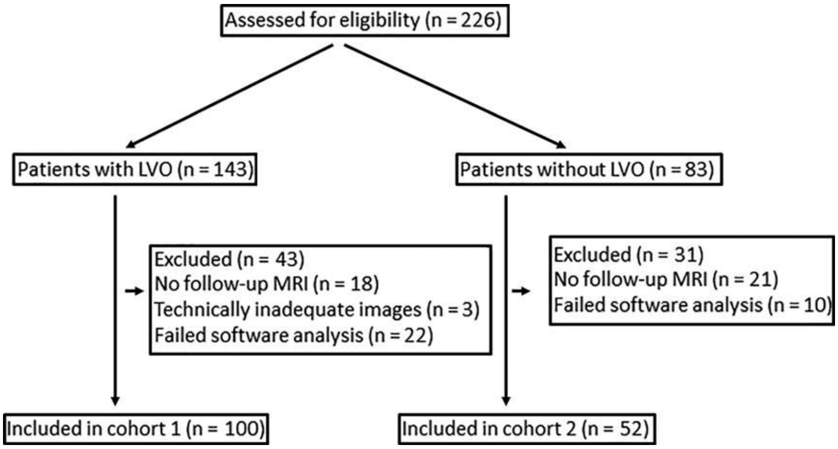

Figure 2: Flowchart shows patient selection and inclusion in analysis. LVO = large vessel occlusion.

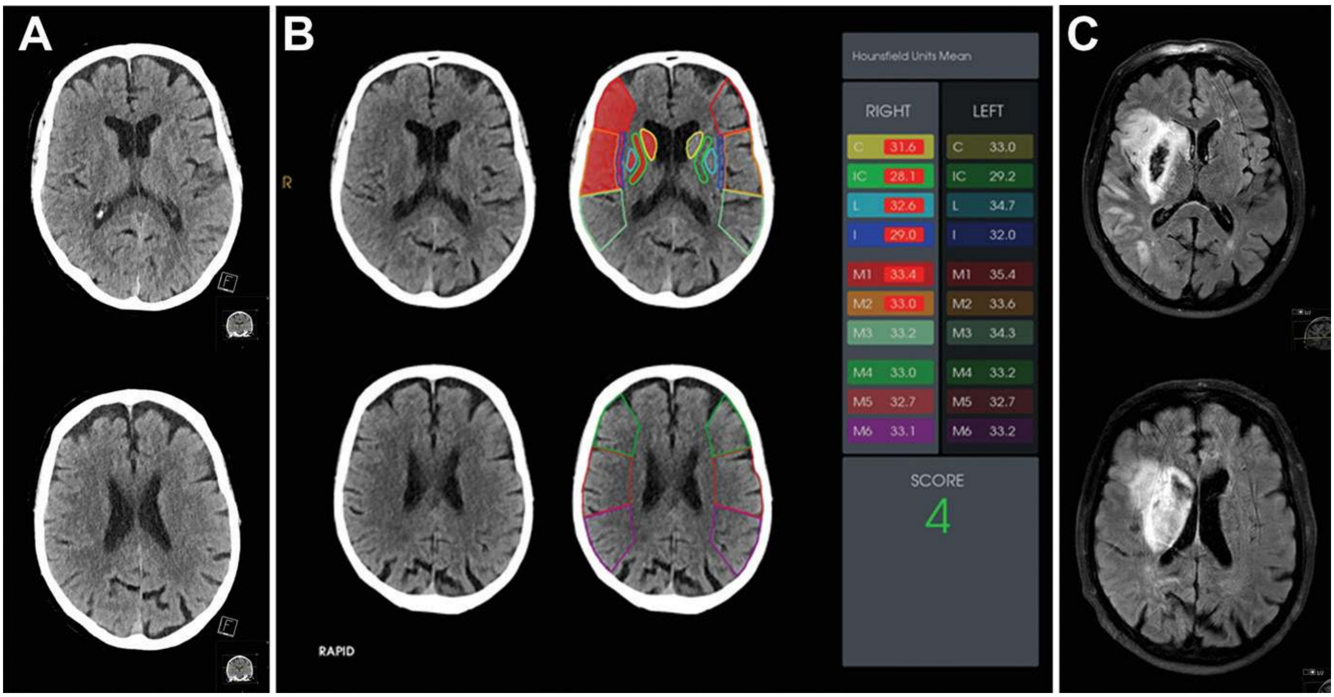

Figure 3: *A, B,* Axial CT images and, *C,* axial fluid-attenuated inversion recovery MR images in a 74-year-old woman with right-sided stroke due to occlusion of M1 segment. Automated software showed much greater agreement with consensus score than with human readers. *A,* Two human readers each scored 8. *B,* Software analysis assigned Alberta Stroke Program Early CT score (ASPECTS) evaluation of 4. Consensus score ASPECT was 4. *C,* This patient was promptly treated with thrombectomy (grade 3 on Thrombolysis in Cerebral Infarction scale) and had fluid-attenuated inversion recovery–ASPECTS evaluation of 3 on third day after intervention.

whom the onset of symptoms was more than 4 hours before imaging, human experts achieved comparable performance with the automated software (κ = 0.76–0.92). This is relevant because many patients with stroke do not present within the first 4 hours; however, in the early-window thrombectomy trials, the vast majority of patients were imaged before 4 hours. As mentioned, based on the software and with strict adherence to the guidelines, we would have excluded seven patients from thrombectomy. Five of those patients presented in the "late time window" beyond 4 hours. One patient presented between 1 and 2 hours after symptom onset and one patient presented in the time window between 2 and 3 hours.

The better performance of the software in early time windows is most likely explained by the subtle changes that could be detected by the software but were not easily detected by the human eye without the benefit of follow-up imaging to direct the readers' attention to these subtle findings. The human readers performed better at longer times after symptom onset because of clearer infarct demarcation. This is shown by the correlation between the time interval from symptom onset to imaging and the accuracy of the ASPECTS evaluation. These findings are corroborated by the literature (21,22). In cohort 2, software analysis was superior to one human reader and essentially identical to the other. Overall performance was slightly worse (κ = 0.6) in cohort 2 compared with

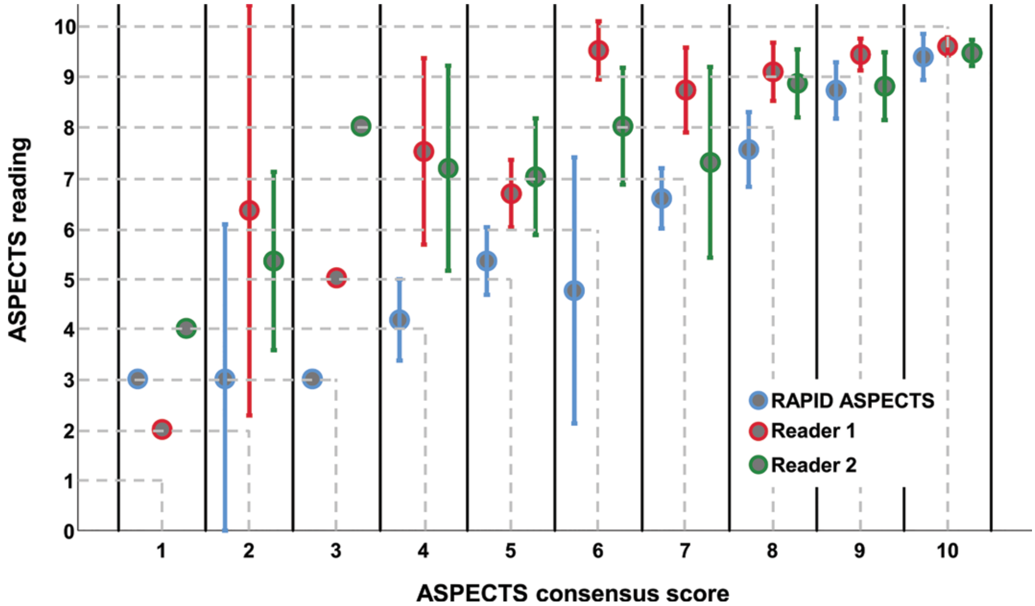

Figure 4: Plot shows correlation between magnitude of consensus score and accuracy of readers and software. Lower Alberta Stroke Program Early CT score (ASPECTS) consensus score correlated with lower accuracy of human readers (both $P < .001$). For software, there was no correlation between consensus score ASPECTS and difference between automated score and consensus score ($P = .16$). Dots indicate mean value and whiskers indicate 2x standard error of the mean.

Table 3: Square-weighted κ Values for Interobserver Agreement on ASPECTS Evaluation

Parameter	Reader 1	Reader 2	ASPECTS Calculated with Automated Software
Cohort 1			
Standard of reference consensus score	0.57	0.57	0.90
Reader 1	…	0.48	0.35
Reader 2	…	…	0.43
Cohort 2			
Standard of reference consensus score	0.30	0.60	0.60
Reader 1	…	0.40	0.35
Reader 2	…	…	0.50

Note.—ASPECTS = Alberta Stroke Program Early CT score.

Table 4: Relationship between CT Scan Time and Interobserver Agreement on ASPECTS Evaluation

Time from Symptom Onset to Imaging (min)	Reader 1 vs Standard of Reference Consensus Score	Reader 2 vs Standard of Reference Consensus Score	ASPECTS Calculated with Automated Software vs Standard of Reference Consensus Score
<60	0.19	0.12	0.17
>60 to <120	0.27	0.36	0.78
>120 to <240	0.64	0.52	0.77
>240	0.83	0.76	0.92

Note.— Data are κ values (calculated by using the square-weighted κ values in cohort 1). Both neuroradiology readers and automated software show greater agreement with the standard of reference Alberta Stroke Program Early CT score (ASPECTS) with greater imaging time following symptom onset.

the emergent large vessel occlusion cohort. This may be explained by the fact that the pretest probability of detecting an infarction lowers substantially when patients with only a clinical suspicion of ischemic stroke are examined. This is illustrated by the fact that the majority of these patients had normal or near-normal scans (Table E2 [online]). In cohort 2, seven patients showed an ASPECTS evaluation of 7 or less: three of them most likely induced by swelling due to a seizure, two showed substantial infarction at follow-up

MRI without large vessel occlusion at initial imaging (most likely an indication of spontaneous dissolution of the clot), and two of the patients had diffuse but subtle leukoencephalopathy. It should also be considered that the ASPECTS evaluation was not designed as a screening tool, but rather to describe the extent of infarction in patients with clinically definite ischemic stroke where reperfusion therapy (either with thrombolysis or endovascular) is being considered. Nevertheless, the ASPECTS evaluations are also relevant to help to determine a diagnosis of stroke.

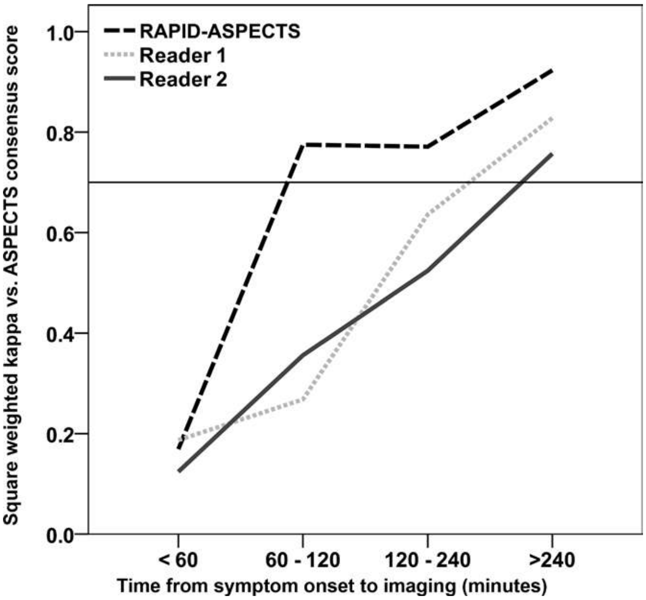

Figure 5: Graph visualizes Alberta Stroke Program Early CT score (ASPECTS) evaluation agreement with consensus score between human readers and fully automatic software based on time between symptom onset and imaging. Whereas software analysis tool had substantial agreement with consensus score after 1 hour, human readers did not reach similar level of agreement until greater than 4 hours.

There is controversy regarding the issue of whether patients with ASPECTS evaluations less than 6 benefit from thrombectomy and some centers routinely treat patients with lower scores. Ongoing randomized trials are addressing this issue (eg, Efficacy and Safety of Thrombectomy in Stroke with Extended Lesion and Extended Time Window [TENSION]). Regardless of the results of these trials, having a method to improve the accuracy and interrater reliability of ASPECT determinations is highly desirable. However, there are potential drawbacks to automated software. ASPECTS software is designed to be used in conjunction with an experienced reader to help validate the final score and to detect artifacts or technical issues that can lead to over- or underestimation of the true ASPECTS evaluation.

Our study had some limitations. Both the definition of the consensus score and the blinded ASPECTS evaluation analysis were performed by the same readers at different time points. However, this was a deliberate decision because a new set of readers for the consensus score read would introduce additional variation based on the well-established interreader variability inherit in ASPECT scoring. The interrater agreement between the human raters and the consensus score might have been less if different raters performed the reference standard read. To minimize a possible

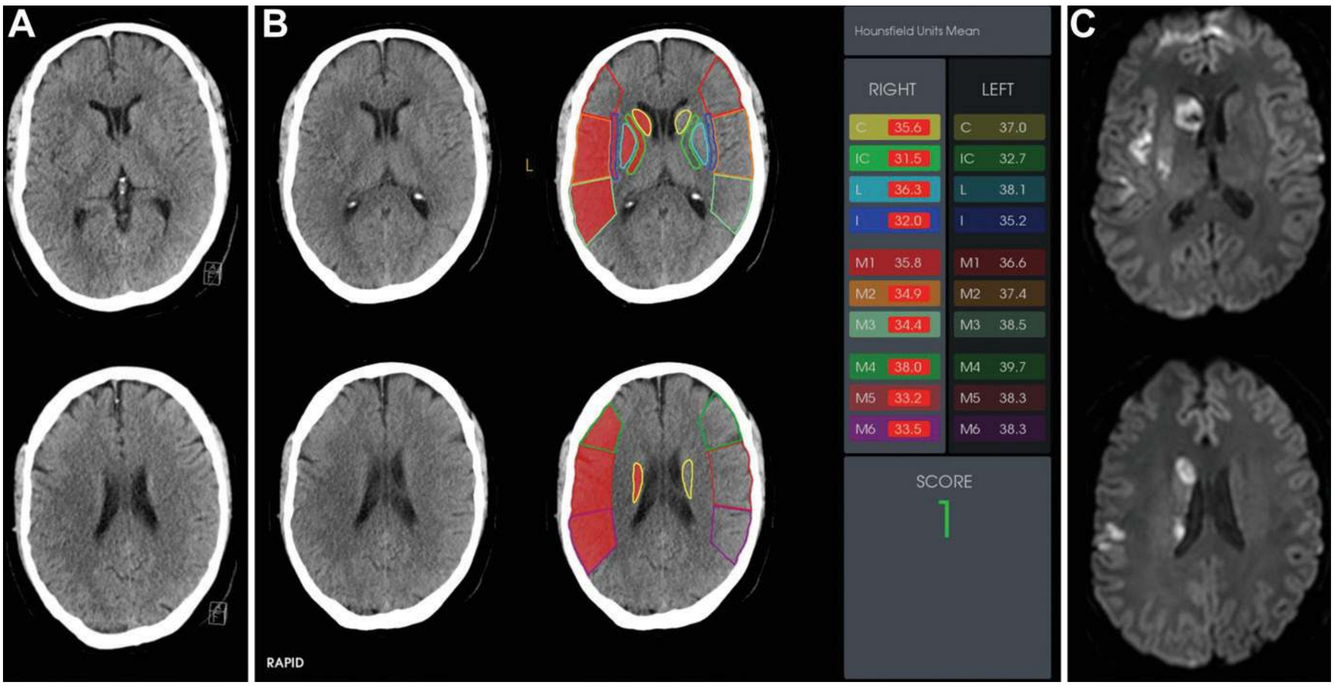

Figure 6: *A, B,* Axial CT images and, *C,* axial diffusion-weighted MR images in a 61-year-old man with right-sided stroke due to occlusion of M1 segment of middle cerebral artery (MCA). *A,* Two human readers each scored 8. *B,* Automated software assigned Alberta Stroke Program Early CT score (ASPECTS) evaluation of 1. Consensus score ASPECT was 6. *C,* This patient was treated with thrombectomy and had ASPECTS evaluation of 6 at diffusion-weighted imaging with multiple patchy lesions in MCA territory and favorable clinical outcome. Extensive areas of hypodensity at acute stroke CT can be clearly visually depicted and were verified to have low Hounsfield unit values (see Hounsfield unit values on analysis output screen in *B*); however, some of these hypodense regions appear normal at follow-up MRI, suggesting reversibility following successful recanalization.

influence of the two analysis runs on each other, a large time interval of 6 weeks was chosen so the readers would have no recollection of the original read. Another potential limitation was that no external selections of the image data were performed. Yet, we believe this limitation is minimal because we included all consecutive patients in our study period whose image data could be analyzed. Also, the knowledge of which hemisphere was affected in cohort 1, as well as the knowledge that the patients had a large vascular occlusion, may have given the human readers an advantage over the software. However, these data are typically available to human readers in routine clinical practice. From a statistical point of view, our results show a relatively narrow interquartile range. This truncation of score variability could alter the correlations that could be seen over a more diverse patient population. A perfectly accurate "ground-truth" ASPECTS evaluation is not achievable because there is no imaging study that provides definitive identification of regions of early ischemic injury. We tried to achieve the best possible consensus score by evaluating all information available, including multimodal imaging that was acquired at the acute stage. This included perfusion imaging, which is more sensitive than is noncontrast CT for identification of early ischemic core lesions, as well as CT angiography. Additionally, we analyzed the MRI after interventional stroke treatment, which is performed 3 days after treatment. Although successful endovascular treatment can have a tremendous impact on the evolution of ischemic lesions, we felt that the inclusion of the follow-up MRIs improved the accuracy of the consensus score reading because regions that progressed to infarction could be carefully scrutinized on the baseline scan to see if subtle changes were present. Another limitation was that about 20% (22 of 100 in cohort 1, 10 of 52 in cohort 2) of the CT data sets could not be analyzed by using automated software ASPECTS, often because the field of view was too small to meet the required 13 cm of coverage. In other CT data sets, the reason the scans could not be analyzed is unclear. In contrast, only three data sets could not be analyzed by the human readers. This finding demonstrates that human experts have more flexibility to interpret nonstandardized images. This further emphasizes that a thorough review of the software output by experienced neuroradiologists or neurologists is mandatory prior to treatment decisions.

In conclusion, in potential candidates for thrombectomy, the determination of the Alberta Stroke Program Early CT score evaluation with a fully automated software tool more closely matches a consensus reference standard, particularly in patients with a time interval between the onset of symptoms and imaging of 1 to 4 hours. Nevertheless, because of technical limitations and potential artifacts, a thorough review of the automated scores by experienced neuroradiologists or neurologists is extremely important.

Acknowledgments: Thr authors thank iSchemaView for providing further details about their proprietary algorithms used in the software described in this article. Thanks to Dr. B. Haller, IMISE, Technical University Munich, for the statistical advisory.

Author contributions: Guarantors of integrity of entire study, C.M., C.Z., B.F.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, C.M., J.F., S.M., M.B., C.Z., B.F.; clinical studies, C.M., S.M., S.W., C.L.S., M.L., T.B.B., C.Z., B.F.; statistical analysis, C.M., J.F., S.M., C.Z., B.F.; and manuscript editing, C.M., S.M., M.B., M.L., T.B.B., C.Z., B.F.

Disclosures of Conflicts of Interest: **C.M.** disclosed no relevant relationships. **J.F.** disclosed no relevant relationships. **S.M.** disclosed no relevant relationships. **M.B.** disclosed no relevant relationships. **S.W.** disclosed no relevant relationships. **C.L.S.** disclosed no relevant relationships. **M.L.** disclosed no relevant relationships. **T.B.B.** disclosed no relevant relationships. **C.Z.** disclosed no relevant relationships. **B.F.** disclosed no relevant relationships.

References

1. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. N Engl J Med 1995;333(24):1581–1587.
2. Hacke W, Kaste M, Fieschi C, et al. Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke: the European Cooperative Acute Stroke Study (ECASS). JAMA 1995;274(13):1017–1025.
3. von Kummer R, Allen KL, Holle R, et al. Acute stroke: usefulness of early CT findings before thrombolytic therapy. Radiology 1997;205(2):327–333.
4. Dávalos A, Toni D, Iweins F, Lesaffre E, Bastianello S, Castillo J. Neurological deterioration in acute ischemic stroke: potential predictors and associated factors in the European cooperative acute stroke study (ECASS) I. Stroke 1999;30(12):2631–2636.
5. Barber PA, Demchuk AM, Zhang J, Buchan AM. 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 2000;355(9216):1670–1674.
6. Pexman JH, Barber PA, Hill MD, et al. Use of the Alberta Stroke Program Early CT Score (ASPECTS) for assessing CT scans in patients with acute stroke. AJNR Am J Neuroradiol 2001;22(8):1534–1542.
7. Aviv RI, Mandelcorn J, Chakraborty S, et al. Alberta Stroke Program Early CT Scoring of CT perfusion in early stroke visualization and assessment. AJNR Am J Neuroradiol 2007;28(10):1975–1980.
8. Puetz V, Dzialowski I, Hill MD, Demchuk AM. The Alberta Stroke Program Early CT Score in clinical practice: what have we learned? Int J Stroke 2009;4(5):354–364.
9. Mak HK, Yau KK, Khong PL, et al. Hypodensity of >1/3 middle cerebral artery territory versus Alberta Stroke Programme Early CT Score (ASPECTS): comparison of two methods of quantitative evaluation of early CT changes in hyperacute ischemic stroke in the community setting. Stroke 2003;34(5):1194–1196.
10. Wardlaw JM, Mielke O. Early signs of brain infarction at CT: observer reliability and outcome after thrombolytic treatment--systematic review. Radiology 2005;235(2):444–453.
11. Truwit CL, Barkovich AJ, Gean-Marton A, Hibri N, Norman D. Loss of the insular ribbon: another early CT sign of acute middle cerebral artery infarction. Radiology 1990;176(3):801–806.
12. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. N Engl J Med 2018;378(1):11–21.
13. Albers GW, Marks MP, Kemp S, et al. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. N Engl J Med 2018;378(8):708–718.
14. Powers WJ, Rabinstein AA, Ackerson T, et al. 2018 guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2018;49(3):e46–e110 [Published corrections appear in Stroke 2018;49(3):e138 and Stroke 2018;49(6):e233–e234.].
15. Finlayson O, John V, Yeung R, et al. Interobserver agreement of ASPECT score distribution for noncontrast CT, CT angiography, and CT perfusion in acute stroke. Stroke 2013;44(1):234–236.
16. Nagel S, Sinha D, Day D, et al. 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 2017;12(6):615–622.
17. Herweh C, Ringleb PA, Rauch G, et al. Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients. Int J Stroke 2016;11(4):438–445.
18. Vanbelle S, Albert A. A bootstrap method for comparing correlated kappa coefficients. J Stat Comput Simul 2008;78(11):1009–1015.
19. Kobkitsuksakul C, Tritanon O, Suraratdecha V. Interobserver agreement between senior radiology resident, neuroradiology fellow, and experienced neuroradiologist in the rating of Alberta Stroke Program Early Computed Tomography Score (ASPECTS). Diagn Interv Radiol 2018;24(2):104–107.
20. McTaggart RA, Jovin TG, Lansberg MG, et al. Alberta stroke program early computed tomographic scoring performance in a series of patients undergoing computed tomography and MRI: reader agreement, modality agreement, and outcome prediction. Stroke 2015;46(2):407–412.
21. Naylor J, Churilov L, Chen Z, Koome M, Rane N, Campbell BCV. Reliability, reproducibility and prognostic accuracy of the Alberta Stroke Program Early CT Score on CT perfusion and non-contrast CT in hyperacute stroke. Cerebrovasc Dis 2017;44(3-4):195–202.
22. Naylor J, Churilov L, Rane N, Chen Z, Campbell BCV, Yan B. Reliability and utility of the Alberta Stroke Program Early Computed Tomography Score in hyperacute stroke. J Stroke Cerebrovasc Dis 2017;26(11):2547–2552. Note.— Data are κ values (calculated by using the square-weighted κ values in cohort 1). Both neuroradiology readers and automated software show greater agreement with the standard of reference Alberta Stroke Program Early CT score (ASPECTS) with greater imaging time following symptom onset.