Original research

Clinical evaluation of a deep-learning model for automatic scoring of the Alberta stroke program early CT score on non-contrast CT

Seong-Joon Lee,[1] Gyuha Park,[2] Dohyun Kim,[2] Sumin Jung,[2] Soohwa Song ![ORCID],[2] Ji Man Hong,[1] Dong Hoon Shin,[2,3] Jin Soo Lee ![ORCID] [1]

Correspondence to
Professor Jin Soo Lee, Department of Neurology, Ajou University School of Medicine Department of Neurology, Suwon, Gyeonggi-do 16499, Korea; jinsoo22@gmail.com and Professor Dong Hoon Shin, Department of Neurology, Gil Hospital, Gachon University1198, Guwol-Dong, Namdong-Gu, Incheon, 405-760, Republic of Korea; dr.donghoon.shin@gmail.com

![Check for updates]

ABSTRACT

Background Automated measurement of the Alberta Stroke Program Early Computed Tomography Score (ASPECTS) can support clinical decision making. Based on a deep learning algorithm, we developed an automated ASPECTS scoring system (Heuron ASPECTS) and validated its performance in a prespecified clinical trial.

Methods For model training, we used non-contrast computed tomography images of 487 patients with acute ischemic stroke (AIS). For the clinical trial, 326 patients (87 with AIS, 56 with other acute brain diseases, and 183 with no brain disease) were enrolled. The results of Heuron ASPECTS were compared with the consensus generated by two stroke experts using the Bland–Altman agreement. A mean difference of less than 0.35 and a maximum allowed difference of less than 3.8 were considered the primary outcome target. The sensitivity and specificity of the model for the 10 regions of interest and dichotomized ASPECTS were calculated.

Results The Bland–Altman agreement had a mean difference of 0.03 [95% confidence interval (CI): −0.08 to 0.14], and the upper and lower limits of agreement were 2.80 [95% CI: 2.62 to 2.99] and −2.74 [95% CI: −2.92 to −2.55], respectively. For ASPECTS calculation, sensitivity and specificity to detect the early ischemic change for 10 ASPECTS regions were 62.78% [95% CI: 58.50 to 67.07] and 96.63% [95% CI: 96.18 to 97.09], respectively. Furthermore, in a dichotomized analysis (ASPECTS >4 vs. ≤4), the sensitivity and specificity were 94.01% [95% CI: 91.26 to 96.77] and 61.90% [95% CI: 47.22 to 76.59], respectively.

Conclusions The current trial results show that Heuron ASPECTS reliably measures the ASPECTS for use in clinical practice.

INTRODUCTION

With the advent of mechanical thrombectomy (MT), various advanced neuroimaging techniques, such as multiphase computed tomography collaterals[1] or core-penumbra mismatch utilizing perfusion imaging[2] have been used for patient selection for reperfusion therapy.[3] However, the pretreatment ischemic core volume remains a strong independent predictor of clinical outcome in acute ischemic stroke (AIS) with occlusion of the proximal arteries.[4] The concept of the late window paradox further helped reveal that the ischemic core volume

WHAT IS ALREADY KNOWN ON THIS TOPIC
⇒ Automated methods of measuring ASPECTS may help in clinical decision making in the emergency department.

WHAT THIS STUDY ADDS
⇒ We have developed and validated a new deep learning-based automated ASPECTS program that could reliably measure ASPECTS among heterogeneous patients suspected of stroke in the emergency department.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY
⇒ Ongoing research regarding automated ASPECTS measurements will improve patient selection and outcomes of reperfusion therapy for ischemic stroke.

can also represent progression speed.[5] It is also well known that intravenous thrombolysis can be harmful in patients with large infarct cores.[6] Thus, the pretreatment ischemic core volume is arguably the most important clinical and imaging parameter that may predict response to reperfusion therapy.

The Alberta Stroke Program Early Computed Tomography Score (ASPECTS) can be obtained using non-contrast computed tomography (NCCT) and has good predictive power for treatment outcomes.[7] It is strongly predictive of clinical outcomes for thrombolysis[8] and standard time window MT.[9] It may also be utilized in late window MT.[10] Thus, a minimum ASPECTS score of ≥6 is suggested as an infarct volume criterion in international thrombectomy guidelines.[11] Recent trials have further extended the ASPECTS limit for thrombectomy efficacy to ≥3.[12]

However, because ASPECTS is based on NCCT, which has low sensitivity to early ischemic change (EIC), it is criticized for its low inter-rater reliability.[13] This may be a greater issue for less experienced practitioners with low-volume exposure to AIS. Accordingly, automated evaluation of the ASPECTS may guide clinicians in decision making. Two commercially available software programs exist for automated evaluation of ASPECTS: the e-ASPECTS software (Brainomix, Oxford, United Kingdom)[14]

BMJ

and RAPID ASPECTS program (iSchemaView, Menlo Park, Calif).[15] However, no study has validated the predictive ability of automated ASPECTS programs in a situation highly resembling ischemic stroke identification in the emergency department, in which there is a significant rate of other acute brain diseases, such as hemorrhage, or no acute brain disease.[16]

This study aimed to develop and validate an automated ASPECTS scoring system (Heuron ASPECTS version 1.0.0.0, Heuron Co., Ltd., Incheon, Republic of Korea) based on a deep-learning algorithm. Model learning was based on NCCT of patients with AIS with consecutive magnetic resonance imaging (MRI). A clinical validation trial was performed in patients presenting to the emergency department with suspected ischemic stroke. The diagnostic accuracy of Heuron ASPECTS was evaluated by analyzing the agreement between Heuron ASPECTS and the ASPECTS consensus of experts. Diagnostic accuracy was validated using a prespecified endpoint.

METHODS

The study protocols of the current clinical trial were approved by the Ministry of Food and Drug Safety of the Republic of Korea (KFDS-1233). Data collection and evaluation were approved by the Institutional Review Board of Gachon University Gil Medical Center (GDIRB2021-234) and Ajou University Medical Center (AJOUIRB-MDB-2020–189), respectively. Ethical standards of the 1964 Declaration of Helsinki and its later amendments were implemented. The need for written informed consent was waived owing to the retrospective data collection method of this study.

Designing the clinical trial

The primary endpoint of the clinical trial was generated based on the results of a previously reported study and the internal test results of Heuron ASPECTS.[17] In the reference study, the agreement between automated ASPECTS software, such as e-ASPECTS, Frontier ASPECTS (prototype v2, Siemens, Germany), and RAPID ASPECTS, was compared with the expert consensus as a reference standard, and it was concluded that reasonable performance was confirmed for all three programs. In the current study, the target mean difference of the Bland-Altman plot was determined to be 0.35, which was the average performance of the three reported ASPECTS software programs and the internal test performance of Heuron ASPECTS. The standard deviation of the difference was determined to be 1.54, which is the lowest value among the results from the reference study. In addition, based on the results of the reference study and internal test performance of Heuron ASPECTS, 95% confidence intervals (CIs) for the upper and lower limits of agreement were calculated, and the maximum allowed difference between the reference standard and Heuron ASPECTS was set at 3.8. A sample size (N) of 326 was calculated based on a significance level of 0.05, power of 80%, and dropout rate of 10%.[18]

Study population and data collection

The study population was generated according to the proportion of patients presenting to the emergency department, based on a previous study that reported 24% AIS, 19% other (acute) brain diseases, and 57% no brain disease, with an allowable difference range of ±10% (figure 1).[16]
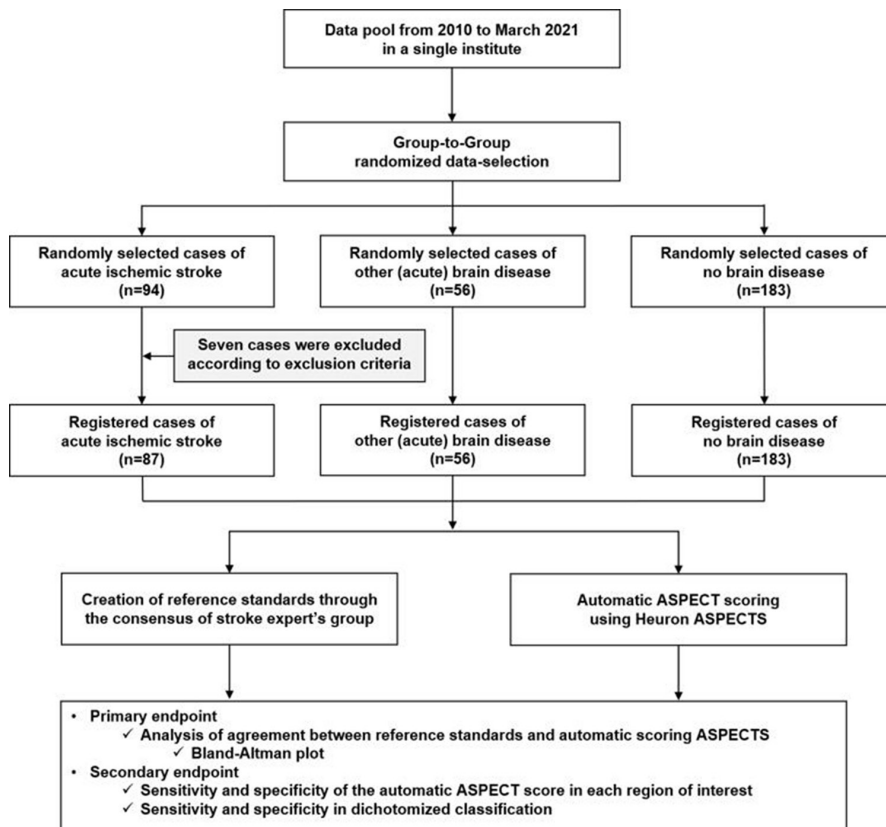

Figure 1 Workflow of the clinical test, including data collection. Flowchart shows patient selection with criteria and workflow to the clinical test for Heuron ASPECTS evaluation.

Table 1	Demographics of the included patients		
	Acute ischemic stroke group (n=87)	Other brain disease group (n=56)	No brain disease group (n=183)
Baseline characteristics			
Age, years (Mean±SD)	67±11	66±16	57±14
Male, n (%)	51 (58.6)	30 (53.6)	94 (51.4)
Smoker, n (%)	31 (36)	11 (19.6)	41 (22.4)
Hypertension, n (%)	53 (61.6)	32 (57.1)	80 (43.7)
Diabetes mellitus, n (%)	26 (30.2)	19 (33.9)	48 (26.2)
Hyperlipidemia, n (%)	12 (14)	2 (3.6)	37 (20.2)
Atrial fibrillation, n (%)	38 (44.2)	0 (0)	1 (0.5)
Cardiac disease, n (%)	33 (38.4)	4 (7.1)	25 (13.7)
Previous stroke, n (%)	13 (15.1)	6 (10.7)	10 (5.5)
NIHSS on admission (Mean±SD)	12±6	–	–
Time from onset to CT scan, min	84±50	–	–
Proportion of diseases in the other brain disease group (n=56)			
C79.30 (Secondary malignant neoplasm of brain)		1 (1.79)	
D18.01 (Hemangioma of intracranial structures)		1 (1.79)	
D43.2 (Brain, unspecified)		1 (1.79)	
G91.9 (Hydrocephalus, unspecified)		1 (1.79)	
I60.9 (Subarachnoid hemorrhage, unspecified)		2 (3.57)	
I61.0 (Intracerebral hemorrhage in hemisphere, subcortical)		31 (55.36)	
I61.1 (Intracerebral hemorrhage in hemisphere, cortical)		5 (8.93)	
I61.3 (Intracerebral hemorrhage in brain stem)		7 (12.50)	
I61.4 (Intracerebral hemorrhage in cerebellum)		2 (3.57)	
I61.5 (Intracerebral hemorrhage, intraventricular)		2 (3.57)	
I67.1 (Cerebral aneurysm, non-ruptured)		1 (1.79)	
Q04.6 (Congenital cerebral cysts)		1 (1.79)	
S06.5 (Traumatic subdural hemorrhage)		1 (1.79)	

CT, computed tomography; NIHSS, National Institute of Health Stroke Scale; SD, standard deviation.

Group-to-group randomized data selection was performed in patients aged ≥19 years with thrombolysis code activation due to suspected ischemic stroke who presented to the emergency department of Gachon University Medical Center between 2010 and March 2021. The final diagnosis of AIS was confirmed using diffusion-weighted MRI. Patients were excluded if cerebral infarction occurred in the posterior circulation, there was loss of information during anonymization, or CT (computed tomography) images had severe noise. Patients who presented with other acute brain diseases were classified as having other (acute) brain diseases. A detailed analysis of the final diagnoses is presented in table 1. The no brain disease group included those in whom a thrombolysis code was activated due to suspected ischemic stroke, but finally concluded that there was no organic brain disease. Cases were randomly collected from each group until the target number of cases was obtained for each group. All demographic information was anonymized and an independent ID was provided for this study. In the CT images of registered cases, slight differences were observed in the scanning model of the device, scanning protocol, or parameters depending on the scan date, but all were taken using devices from the same vendor (Siemens Healthineers, Erlangen, Germany), and the slice thickness varied from 3 to 5 mm.

Generation of the ASPECTS reference standard

Two experts with more than 10 years of clinical experience generated the ASPECTS reference standard. The two experts independently evaluated the presence of EIC at each region of interest (ROI) using NCCT images and then calculated the total score of the ASPECTS. If the results for each ROI evaluated by the two experts were different, the reference standard was determined by consensus between the two experts.

Automatic ASPECTS scoring by Heuron ASPECTS

The Heuron ASPECTS automated software is a solution developed to automatically calculate ASPECTS from NCCT images scanned for suspected acute ischemic stroke patients. As shown in online supplemental figure 1, the scanned NCCT images can be directly transmitted from the CT device to Heuron ASPECTS. Therefore, it is possible to receive and analyze CT images immediately after scanning, and the analyzed results are uploaded to PACS in the form of a report. It takes less than 5 minutes to complete the analysis after the NCCT image input, and the time required for data transmission varies depending on the environment of each institution.

Based on the deep learning technique of a convolutional neural network (CNN), it was trained using NCCT images of 557 patients with ischemic stroke collected from a single institution (Ajou University Medical Center). The ground truth for learning was derived by evaluating the presence of EIC and old infarction (OI) based on NCCT, MR-DWI, ADC, and FLAIR images taken within 1 hour of NCCT. The CNN model was trained and tested with fivefold cross-validation, and the EIC and OI probability values for each ROI for ASPECTS were assembled with the outputs of the five trained models. The EIC and OI classifications based on the ensemble probability values were determined based on the specific threshold values for each ROI. For classifying EIC or OI, the threshold was determined by finding the optimal correlation that satisfied the basic criteria of ≥90% specificity between inferenced ASPECTS and the reference standard.

As shown in figure 2A,B, NCCT images input into Heuron ASPECTS were initially pre-processed (such as noise and skull removal), and the ROI was automatically segmented for ASPECTS evaluation. The presence of EIC or OI in each ROI was independently classified. Basically, the ASPECTS is derived by deducting a point for each ROI evaluated as an EIC based on 10 points in each hemisphere. In Heuron ASPECTS, however, the ROI classified as OI was also deducted to derive the ASPECTS if there was at least one ROI classified as EIC in the same hemisphere.

For Heuron ASPECTS calculation for the clinical trial, an evaluator trained to use Heuron ASPECTS (a radiologist at the clinical trial institution blinded to the clinical data) imported the data into Heuron ASPECTS. After the generation of the reference standard and derivation of Heuron ASPECTS results for all cases, an independent statistical analyst collected and analyzed the results.

Statistical analysis

The primary outcome of this study was the performance of Heuron ASPECTS, which was the degree of agreement with the reference standard. It was evaluated by systematic
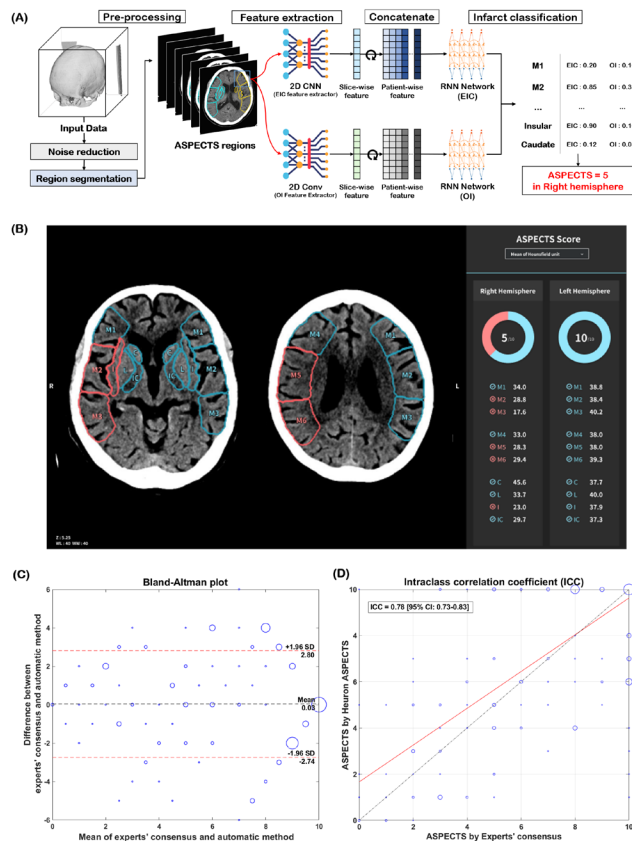
Figure 2 The automatic ASPECTS estimation based on the deep-learning model and its performance. (A) progress in estimating the aspects, and (B) output of Heuron ASPECTS. The study results are shown as a Bland–Altman plot (C) of experts' consensus and Heuron ASPECTS. The mean difference is 0.03, and the upper and lower limits of agreement are 2.80 and −2.74, respectively, satisfying prespecified primary outcomes. (D) The intraclass correlation coefficient (ICC) is 0.78 (95% CI: 0.73 to 0.83), showing good to excellent agreement.

differences between the experts' read and Heuron ASPECTS using the Bland–Altman plot, with prespecified study goals. Next, the performance of Heuron ASPECTS was assessed using the intraclass correlation coefficient (ICC), one-way random-effects, absolute agreement, and single-rater/measurement model, with the reference standard. ICC values less than 0.4 were deemed to have poor reliability, values between 0.4 to 0.59 to have fair reliability, values between 0.60 to 0.74 to have good reliability, and values greater than 0.75 to have excellent reliability.[19]

The secondary outcome of the current study was region-based analysis and analysis using dichotomized cut-off values. The sensitivity, specificity, and area under the receiver-operating characteristic curve (AUC) were analyzed to determine the performance of infarction discrimination at each ROI level. In addition, the ASPECTS-based dichotomized classification performance was analyzed, and both classification performances were dichotomized according to high (ASPECTS >6 vs. ≤ 6) and low (ASPECTS >4 vs. ≤ 4) cut-off values.

Bootstrapping with at least 2000 resamplings was used to calculate the CIs of sensitivity, specificity, and AUC for each variable. All statistical analyses were performed using MATLAB 2020b (MathWorks Inc., Natick, Massachusetts).

RESULTS
Demographics
In the clinical trial, 333 cases were collected through primary screening, and seven cases were excluded. Thus, 326 cases were enrolled in this study. All excluded cases presented with ischemic stroke in the posterior circulation. Of the 326 cases that were finally registered, 87, 56, and 183 were in the ischemic stroke, other brain disease, and no brain disease groups, respectively. The other brain disease group consisted of diseases such as subcortical intracerebral hemorrhage (ICH) (55.36%), ICH of the brainstem (12.5%), and cortical ICH (8.93%). There were no dropouts (figure 1).

The mean age of each group was 67±11, 66±16, and 57±14 years, and the proportion of males was 58.6%, 53.6%, and 51.4%, respectively. In the ischemic stroke group, baseline NCCT was performed within 84±50 min from the last normal time, and the mean National Institute of Health Stroke Scale (NIHSS) was 12±6. The detailed participant characteristics are shown in table 1.

Primary outcomes
Agreement was analyzed using the Bland–Altman plot and ICC between Heuron ASPECTS and experts' consensus for total ASPECTS. As shown in figure 2C,D, the mean difference between two results was 0.03 [95% CI: −0.08 to 0.14], and the upper and lower limits of agreement were 2.80 [95% CI: 2.62 to 2.99] and −2.74 [95% CI: −2.92 to −2.55], respectively, which satisfied the prespecified primary outcomes. The ICC was 0.78 [95% CI: 0.73 to 0.83] between the results of individual ASPECTS, showing good-to-excellent agreement.

Secondary outcomes
The lesion classification performance was calculated at each ROI. As shown in table 2, the specificity was over 90% for all ROIs, and the sensitivity varied depending on the ROI. For all regions, Heuron ASPECTS yielded a sensitivity of 62.78% [95% CI: 58.50 to 67.07], specificity of 96.63% [95% CI: 96.18 to 97.09], and an AUC of 0.88 [95% CI: 0.87 to 0.90]. For dichotomized ASPECTS, Heuron ASPECTS demonstrated a sensitivity of 94.01% [95% CI: 91.26 to 96.77], specificity of 61.90% [95% CI: 47.22 to 76.59], and an AUC of 0.89 [95% CI: 0.88 to 0.90] in the dichotomized criteria of >4 vs. ≤ 4, and sensitivity of 95.42% (95% CI: 92.89 to 97.95), specificity of 76.56% (95% CI: 66.18 to 86.94), and AUC of 0.80 (95% CI: 0.80 to 0.81) in the dichotomized criteria of >6 vs. ≤ 6.

DISCUSSION
The current clinical trial evaluated the reliability and consistency of the Heuron ASPECTS program in identifying EICs in NCCT images encountered in emergency department patients with suspected AIS. The Bland–Altman plot showed a comparable mean difference and maximum allowed difference with those of previously reported automated programs,[14] fulfilling the prespecified study goal. Furthermore, the reliability and consistency of the Heuron ASPECTS program was confirmed by ICC estimation. As a result of the reliability analysis, where the 95% CI of an ICC estimation was 0.73 to 0.83, the level of reliability was regarded as "good-to-excellent".

The ASPECTS score is quantitative and the agreement between total scores is important for evaluating its performance. Thus, the Bland–Altman plot has been utilized in previous studies regarding automated ASPECTS analysis.[14 17 20 21] As the current study was a clinical trial, we utilized pre-specified mean

Table 2 Sensitivity, specificity, and AUC of Heuron ASPECTS prediction in comparison with experts' consensus in each region of interest and dichotomized cut-off points

ROI for ASPECTS	Sensitivity, % (TP/TP+FN) (95% CI)	Specificity, % (TN/TN+FP) (95% CI)	AUC (95% CI)
M1	53.49 (23/43) (38.58 to 68.40)	98.85 (602/609) (98.00 to 99.70)	0.89 (0.83 to 0.94)
M2	56.25 (36/64) (44.10 to 68.40)	99.32 (584/588) (98.66 to 99.98)	0.89 (0.83 to 0.94)
M3	63.89 (23/36) (48.20 to 79.58)	96.10 (592/616) (94.58 to 97.63)	0.88 (0.83 to 0.93)
M4	45.16 (14/31) (27.64 to 62.68)	96.78 (601/621) (95.39 to 98.17)	0.88 (0.82 to 0.93)
M5	52.54 (31/59) (39.80 to 65.28)	98.82 (586/593) (97.95 to 99.69)	0.88 (0.83 to 0.94)
M6	58.06 (18/31) (40.69 to 75.44)	97.58 (606/621) (96.38 to 98.79)	0.88 (0.83 to 0.94)
Caudate	74.51 (38/51) (62.55 to 86.47)	94.84 (570/601) (93.07 to 96.61)	0.88 (0.83 to 0.94)
Lentiform	76.79 (43/56) (65.73 to 87.84)	94.46 (563/596) (92.63 to 96.30)	0.88 (0.83 to 0.93)
Insula	61.97 (44/71) (50.68 to 73.26)	97.25 (565/581) (95.92 to 98.58)	0.89 (0.83 to 0.94)
Int. Capsule	78.72 (37/47) (67.02 to 90.42)	92.40 (559/605) (90.28 to 94.51)	0.88 (0.83 to 0.94)
All ROIs	62.78 (307/489) (58.50 to 67.07)	96.63 (5828/6031) (96.18 to 97.09)	0.88 (0.87 to 0.90)
Dichotomized analysis	**Sensitivity, % (TP/TP+FN)(95% CI)**	**Specificity, % (TN/TN+FP)(95% CI)**	**AUC (95% CI)**
> 4 vs. ≤ 4	94.01 (267/284) (91.26 to 96.77)	61.90 (26/42) (47.22 to 76.59)	0.89 (0.884 to 0.90)
> 6 vs. ≤ 6	95.42 (250/262) (92.89 to 97.95)	76.56 (49/64) (66.18 to 86.94)	0.80 (0.80 to 0.81)

AUC, area under the curve; FN, false negative; FP, false positive; ROI, region of interest; TN, true negative; TP, true positive.

difference and maximum allowed difference values of Bland–Altman analysis as the primary outcome, which was benchmarked from previous studies.[22] In comparison, ICC values have been used for classification[19] and comparison[23] of reliability. A recent meta-analysis compared the performance of automated and manual ASPECTS predictions for early stroke changes. It reported good reliability between reference standards and both expert (ICC 0.62 [95% CI 0.52 to 0.71]) and automated predictions (ICC 0.72 [95% CI 0.61 to 0.80]), while concluding that automated prediction may be superior, based on higher ICC values.[23] The ICC in our study was 0.78 [95% CI: 0.73 to 0.82] showing good-to-excellent agreement, with results comparable to the meta-analysis.[23] In clinical situations, however, ASPECTS is often used in a dichotomized fashion, for selection of patients for reperfusion treatments. Our study results show high sensitivity and moderately high specificity, which is reasonable when we consider the role of automated ASPECTS to aid clinicians in selecting more patients to receive highly effective treatment.[24]

In segmental analysis, Heuron ASPECTS showed acceptable sensitivity, specificity, and high AUC (0.877–0.885) for all segments. A previous study reported a limited correlation between automated ASPECTS software (Brainomix e-ASPECTS, RAPID ASPECTS, and Frontier V2 [Siemens Healthcare GmbH, Forchheim, Germany]) and expert consensus, especially for the M3 segment (AUC, −0.027–0.693) and internal capsule (AUC, 0.000–0.691).[17] In a more recent study, the 3D-BHCA model, a deep-learning-based algorithm similar to Heuron ASPECTS, exhibited region-based ASPECTS analysis with better performance than that of human readers in the early time window.[25] Thus, automated ASPECTS utilizing deep-learning algorithms such as Heuron ASPECTS may be advantaged, especially in the identification of segmental-level EIC. To our knowledge, among the previously reported automated ASPECTS, the eASPECTS,[14] the RAPID ASPECTS,[15] and the method by Kuang et al[21] used a machine-learning algorithm,[17] while Frontier ASPECTS selects EIC based on brain densitometry.[20] In machine-learning, prior knowledge is essential; a researcher extracts suitable features from the input data, such as Houndsfield unit (HU), density, and HU difference compared with the contralateral side. However, in medical images with large domain sizes and information,

there are limits to the extractability of targeted features through machine-learning. It may also be limited when images have low signal-to-noise ratios and motion artifacts.[21] In contrast, the deep-learning algorithm can identify suitable features in large-sized domain data through the model itself with minimum information. Therefore, as the data used for training increases, the accuracy of deep-learning based models is believed to be higher than that of representative machine-learning algorithms.[26] To date, deep-learning methods of ASPECTS measurements have been reported,[25] but no comparative study between machine-learning and deep-learning has been performed.

Another strength of the current study is that it validated Heuron ASPECTS' ability to correctly measure the ASPECTS in a heterogeneous disease identical to emergency department presentations. Previous reports on automated ASPECTS programs have not addressed these issues. In diseases such as hemorrhage, perihematomal edema[27] results in changes in parenchymal densities or loss of gray–white matter discrimination. Such changes may be incorrectly recognized as EIC; hence, incorrect ASPECTS may occur. Our results show that such patients can be effectively screened using Heuron ASPECTS. In real-world practice, there is also a high number of patients with negative brain disease. In these patients, incidental OI lesions[28] may masquerade as EIC. We believe that the good predictive ability shown in the current study is partly due to the novel deep-learning algorithm design that discriminates between EIC and OI.

Apart from deep-learning analysis, the Heuron ASPECTS has some distinctive features. After preprocessing to remove the skull and artifacts, Heuron ASPECTS discriminated between OI and EIC. While some models exclude OI before the calculation of the ASPECTS,[29] Heuron ASPECTS detects signs of EIC and OI concomitantly. If EIC is present ipsilateral to the OI, the program subtracts the suspected OI along with the EIC from the ASPECTS, while OI by itself is not subtracted (online supplemental figure 2A,B). This pattern, in our opinion, most closely represents the human calculation of the ASPECTS. Heuron ASPECTS also internally discriminates hemorrhage; when hemorrhage is detected, Heuron ASPECTS does not undergo further calculation, resulting in an ASPECTS of 10 (online supplemental figure 2C,D). This is especially important in cases

of subtle hemorrhage, where ASPECTS suggestive of early ischemia can be clinically misleading. The hemorrhage detection ability of Heuron ASPECTS is expected to be confirmed in future studies.

Some limitations of this study should be noted. First, images of the study population were acquired at a single institution using a single vendor. There is a chance that differences in CT scanners and reconstruction methods may influence the predictive ability of automated ASPECTS software.[30] However, this effect is more pronounced in less experienced readers and less pronounced in automated software.[30] It should also be noted that the learning dataset was from a different institution, utilizing a different vendor and reconstruction method. Second, while this study has a strength in that it implemented a larger sample of other brain diseases or no brain disease cases compared with the stroke sample, there is a chance that the predictive ability of the Heuron ASPECTS may have been overestimated by the higher number of patients with negative EICs. Further analysis involving a large number of patients with ischemic stroke is needed in future studies.

CONCLUSION

In conclusion, Heuron ASPECTS reliably measured the ASPECTS on NCCT scans of patients suspected of having ischemic stroke. This was significant in both region-based analysis and clinically important cut-off values, and its ability was noninferior to that of the automated ASPECTS methods previously published. These findings suggest that deep-learning algorithm software may provide a useful aid to physicians caring for patients with stroke. Future studies are needed to evaluate this further.

ORCID iDs
Soohwa Song http://orcid.org/0000-0002-4723-5603
Jin Soo Lee http://orcid.org/0000-0002-7225-6166

REFERENCES

1 Menon BK, d'Esterre CD, Qazi EM, et al. Multiphase CT angiography: a new tool for the imaging triage of patients with acute ischemic stroke. *Radiology* 2015;275:510–20.
2 Xie Y, Oppenheim C, Guillemin F, et al. Pretreatment lesional volume impacts clinical outcome and thrombectomy efficacy. *Ann Neurol* 2018;83:178–85.
3 Lee JS, Demchuk AM. Choosing a hyperacute stroke imaging protocol for proper patient selection and time efficient endovascular treatment: lessons from recent trials. *J Stroke* 2015;17:221–8.
4 Han M, Choi JW, Rim N-J, et al. Cerebral infarct volume measurements to improve patient selection for endovascular treatment. *Medicine (Baltimore)* 2016;95:e4702.
5 Albers GW. Late window paradox. *Stroke* 2018;49:768–71.
6 National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med* 1995;333:1581–7.
7 Yoo AJ, Zaidat OO, Chaudhry ZA, et al. Impact of pretreatment noncontrast CT Alberta stroke program early CT score on clinical outcome after intra-arterial stroke therapy. *Stroke* 2014;45:746–51.
8 Barber PA, Demchuk AM, Zhang J, et al. 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;355:1670–4.
9 Goyal M, Demchuk AM, Menon BK, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. *N Engl J Med* 2015;372:1019–30.
10 Nguyen TN, Abdalkader M, Nagel S, et al. Noncontrast computed tomography vs computed tomography perfusion or magnetic resonance imaging selection in late presentation of stroke with large-vessel occlusion. *JAMA Neurol* 2022;79:22–31.
11 Powers WJ, Rabinstein AA, Ackerson T, et al. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: a guideline for healthcare professionals from the American heart association/american stroke association. *Stroke* 2019;50:e344–418.
12 Yoshimura S, Sakai N, Yamagami H, et al. Endovascular therapy for acute stroke with a large ischemic region. *N Engl J Med* 2022;386:1303–13.
13 Gupta AC, Schaefer PW, Chaudhry ZA, et al. Interobserver reliability of baseline noncontrast CT Alberta stroke program early CT score for intra-arterial stroke treatment selection. *AJNR Am J Neuroradiol* 2012;33:1046–9.
14 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;11:438–45.
15 Maegerlein C, Fischer J, Mönch S, et al. Automated calculation of the Alberta stroke program early CT score: feasibility and reliability. *Radiology* 2019;291:141–8.
16 Kim H.J. KDH, Park HI, Lee IY, et al. Prehospital delay factors after stroke and paramedic identification of stroke patients in a metropolitan city emergency medical service system. *J Korean Neurol Assoc* 2011;29:89–94.
17 Hoelter P, Muehlen I, Goelitz P, et al. Automated aspect scoring in acute ischemic stroke: comparison of three software tools. *Neuroradiology* 2020;62:1231–8.
18 Lu M-J, Zhong W-H, Liu Y-X, et al. Sample size for assessing agreement between two methods of measurement by bland-altman method. *Int J Biostat* 2016;12:/j/ijb.2016.12.issue-2/ijb-2015-0039/ijb-2015-0039.xml.
19 Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic* 1981;86:127–37.
20 Goebel J, Stenzel E, Guberina N, et al. Automated aspect rating: comparison between the frontier aspect score software and the brainomix software. *Neuroradiology* 2018;60:1267–72.
21 Kuang H, Najm M, Chakraborty D, et al. Automated aspects on noncontrast CT scans in patients with acute ischemic stroke using machine learning. *AJNR Am J Neuroradiol* 2019;40:33–8.
22 Giavarina D. Understanding Bland altman analysis. *Biochem Med (Zagreb)* 2015;25:141–51.
23 Adamou A, Beltsios ET, Bania A, et al. Artificial intelligence-driven aspects for the detection of early stroke changes in non-contrast CT: a systematic review and meta-analysis. *J Neurointerv Surg* 2023;15:e298–304.
24 Nogueira RG, Ribó M. Endovascular treatment of acute stroke. *Stroke* 2019;50:2612–8.
25 Naganuma M, Tachibana A, Fuchigami T, et al. Alberta stroke program early CT score calculation using the deep learning-based brain hemisphere comparison algorithm. *J Stroke Cerebrovasc Dis* 2021;30:105791.

26 Ng A. Chapter 4. scale drives machine learning progress. In: *Machine Learning Yearning*. 2018: 10–2.

27 Ironside N, Chen C-J, Ding D, *et al*. Perihematomal edema after spontaneous intracerebral hemorrhage. *Stroke* 2019;50:1626–33.

28 Saini M, Suministrado MSP, Hilal S, *et al*. Prevalence and risk factors of acute incidental infarcts. *Stroke* 2015;46:2722–7.

29 Wolff L, Berkhemer OA, van Es ACGM, *et al*. Validation of automated alberta stroke program early CT score (ASPECTS) software for detection of early ischemic changes on non-contrast brain CT scans. *Neuroradiology* 2021;63:491–8.

30 Seker F, Pfaff J, Nagel S, *et al*. CT reconstruction levels affect automated and reader-based ASPECTS ratings in acute ischemic stroke. *J Neuroimaging* 2019;29:62–4.