Machine learning volumetry of ischemic brain lesions on CT after thrombectomy—prospective diagnostic accuracy study in ischemic stroke patients

Jiri Kral [1,2] · Martin Cabal [1] · Linda Kasickova [1,3] · Jaroslav Havelka [4] · Tomas Jonszta [4] · Ondrej Volny [5,6] · Michal Bar [1,3]

Abstract

Purpose Ischemic lesion volume (ILV) is an important radiological predictor of functional outcome in patients with anterior circulation stroke. Our aim was to assess the agreement between automated ILV measurements on NCCT using the Brainomix software and manual ILV measurements on diffusion-weighted imaging (DWI).

Methods This was a prospective single-center observational study of patients with CT angiography (CTA) proven anterior circulation occlusion treated with endovascular thrombectomy (May 2018 to May 2019). NCCT ILV was measured automatically by the Brainomix software. DWI ILV was measured manually. The McNemar's test was used to test sensitivity and specificity. The Somer's delta was used to test the differences between concordant and discordant ASPECTS regions. The Bland-Altman plot was calculated to compare the differences between Brainomix and DWI ILVs.

Results Forty-five patients were included. Median Brainomix ILV was 23 ml (interquartile range [IQR], 15–39 ml), and median DWI ILV was 11.5 ml (IQR, 7–32 ml) in the TICI 2b-3 group. In the TICI 0-2a, the NCCT ILV was 39 ml (IQR, 18–62 ml) and DWI ILV was 30 (IQR, 11–105 ml). The DWI ILVs in patients with good clinical outcome (mRS 0–2) was significantly lower compared with patients with mRS ≥ 3 (10 mL vs 59 mL, $p = 0.002$). Similar trend was observed for Brainomix ILV measurements (21 mL vs 39 mL, $p = 0.012$). There was a high correlation and accuracy in the detection of follow-up ischemic changes in particular ASPECTS regions.

Conclusion NCCT ILV measured automatically by the Brainomix software might be considered a valuable radiological outcome measure.

Keywords Computed tomography · Software · Automatic · Final ischemia

✉ Michal Bar
michal.bar@fno.cz

1 Department of Neurology, Faculty Hospital Ostrava, 17. Listopadu 1790/5, 708 52 Ostrava-Poruba, Czech Republic

2 Faculty of Medicine, Masaryk University, Brno, Czech Republic

3 Faculty of Medicine, Ostrava University, Ostrava, Czech Republic

4 Department of Neuroradiology, Faculty Hospital Ostrava, Ostrava, Czech Republic

5 International Clinical Research Centre, Stroke Research Program, St. Anne's University Hospital, Brno, Czech Republic

6 Department of Neurology, St. Anne's University Hospital and Faculty of Medicine, Masaryk University, Brno, Czech Republic

Introduction

Baseline and follow-up diagnostic neuroimaging represent an inevitable part of stroke diagnosis, management, and outcome prediction. Follow-up non-contrast-computed tomography (NCCT) or magnetic resonance imaging (MRI) > 24 h post-recanalization treatment is clinically used for exclusion of hemorrhagic complications and visual estimation of the extent and location of acute ischemia [1–5].

Both the hypodensity (hypoattenuation) on follow NCCT and hyperintensity on diffusion-weighted imaging (DWI) MRI depict the same pathology, i.e., cellular swelling and increased water content. The DWI per se represents highly sensitive diagnostic modality in detecting acute ischemic lesions in stroke patients [6].

Previous studies showed that the ischemic lesion volume (ILV, i.e., volume of brain infarction/acute ischemia)

measured on follow-up scans (> 24 h post-treatment) correlated with the 3-month functional outcome defined as modified Rankin scale (mRS) [1, 3, 5]. Even though the early prediction of short-term or long-term clinical outcome based on ILV is considered valuable, the ILV is not measured routinely in the everyday clinical practice due to time demands of manual volumetric measurements and limited personnel resources. These limitations might be overcome by a reliable automatic software.

Our primary goal was to evaluate the level of agreement between the ILVs measured by an available automatic software (for NCCT) and manual outlining (for DWI-MRI) on follow-up stroke neuroimaging in patients who underwent mechanical thrombectomy for a large vessel occlusion (LVO) in the anterior cerebral circulation.

Methods

This was a prospective single-center observational study of patients with CT angiography (CTA) proven proximal occlusion in the anterior cerebral circulation (terminal internal carotid artery [ICA] and/or M1 segment of middle cerebral artery [MCA]) and premorbid modified Rankin stroke scale < 2, who were treated with endovascular thrombectomy (EVT) between May 2018 and May 2019 in the Comprehensive Stroke Center of Faculty Hospital Ostrava (Czech Republic). Ethical approval was obtained from the local Institutional Review Boards.

All patients with symptoms of acute ischemic stroke and no history of contrast allergy routinely undergo baseline NCCT and single-phase CTA from the aortic arch to the vertex at our center (CT perfusion was not performed during the study period). Follow-up neuroimaging included NCCT (a standard of care at our institution), and MRI performed 24 to 48 h after the EVT procedure (not a standard of care at our institution). Both follow-up neuroimaging methods were performed 30 min apart. This time period was selected in order to compare the reliability of ILV measurement on NCCT (automated analysis by the Brainomix software) and DWI-MRI manual outlining. The timing of follow-up imaging > 24 h was chosen since it represents the earliest time point for accurate delineation of acute ischemia volume [7].

Imaging protocol

NCCT was performed on a multi-detector spiral 64 series CT machine (Siemens Medical Systems, Erlangen, Germany). The NCCT examination was followed by CTA using 50–100 ml of iodine conjugate (Visipaque, GE Healthcare, Piscataway, NJ, USA), which was applied at a rate of 4 ml/s. The range of CTA was from the aortic arch to the distal intracranial artery. The width of the basic CT sections for further reconstruction was 0.75 mm.

MRI was performed on the 1.5 T Siemens Avanto (Siemens, Erlangen, Germany). Imaging protocol included the following: localizer; diffusion-weighted imaging (DWI) in 3 planes with b-factor values of 0 and 1000 s/mm2; reconstructed maps of the apparent diffusion coefficient (ADC) maps (ST 5 mm, gap 0.5 mm, duration: 20 s each + SHIM time); and fluid attenuated inverse recovery (FLAIR) sequences in axial and sagittal plane (ST 3 mm, gap 0 mm, TR 8000 ms, TE 105 ms, TI 2371 ms).

Imaging analysis

ILVs on follow-up DWI-MRI were manually outlined by consensus of 2 experienced neuroradiologists (> 10 years of experience) who were blinded to all clinical information apart from the affected hemisphere (2 volumetric measurements were done 2 weeks apart). The DWI scans with the b factor of 1000 s/mm^2 were used for the volumetric measurements of the cytotoxic ILV. The DWI hyperintense area was manually contoured in each slice. The total ILV was calculated as a sum of DWI areas in each slice multiplied by the effective slice thickness. The median of all four measurements was used for the statistical analyses.

ILVs on follow-up NCCT were measured automatically by the Brainomix software (Oxford, UK version 7.1; Fig. 1), a tool based on a combination of advanced image-processing and machine-learning algorithms. The image processing algorithm follows an AI approach, with a combination of traditional 3D graphics and statistical methods, and machine learning classification techniques. The input DICOM data is first resampled to correct any gantry tilt and to standardize the input resolution. Then, a fast proprietary registration approach is applied to re-align the data, removing any tilt and rotation. This ensures that the image is presented in a standard reference frame, which simplifies human interpretation of the scan. A scoring module uses a machine learning classifier to identify both early and non-acute signs of ischemia in order to derive a voxel-wise probability map. This classifier has been trained on a large dataset (> 10,000 images) containing a wide range of real-world CT scans from stroke patients and negative controls, with ground-truth data from additional imaging data such as MRI acquired within 1–2 h of the CT scan. This dataset contains examples of CT scans captured with scanners from all major manufacturers and from a wide range of countries worldwide. A patient-specific segmentation of the ASPECTS regions and ILV is computed, and finally, the output score or ILV is generated.

Fig. 1 Ischemic lesion volume measured automatically by the Brainomix software
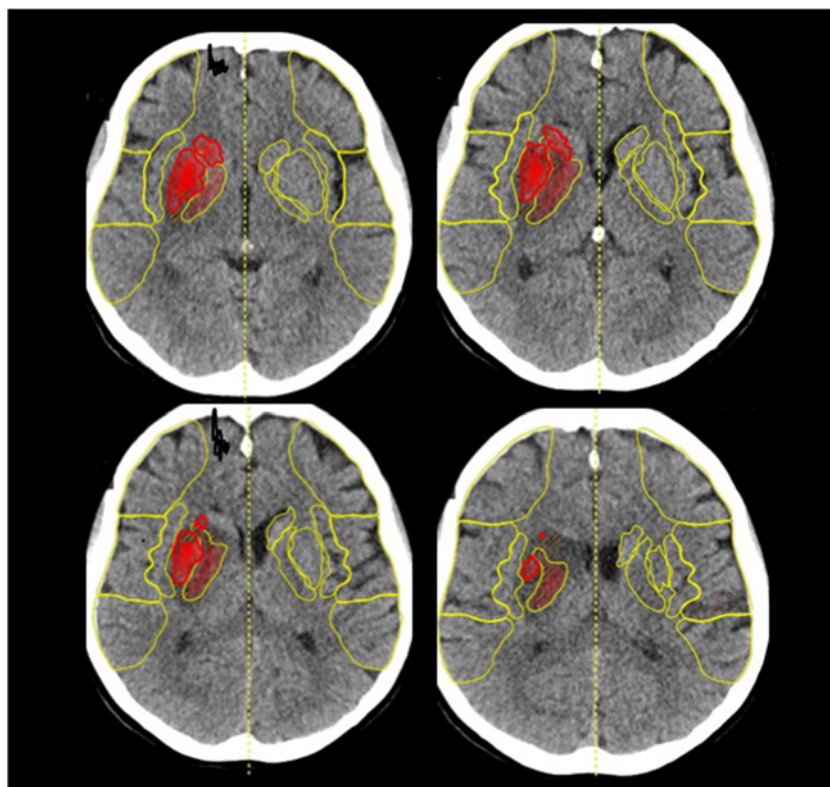
Table 1 Clinical, radiological, and outcome data of good and poor recanalization group

	Recanalization mTICI 0-2a (n = 13)	Recanalization mTICI 2b-3 (n = 32)	p values
Clinical baseline characteristics			
Age, mean (SD), years	66.5 (±15)	69.9 (±9)	0.69
Female sex, n (%)	6 (46.2)	19 (59.4)	-
Baseline NIHSS score, median (IQR)	17 (13–18)	16 (15–18)	0.92
Imaging characteristics			
Baseline NCCT, e-ASPECTS, median (IQR)	10 (8–10)	9.5 (8–10)	0.91
Baseline NCCT early ischemic changes volume (Brainomix), median (IQR), mL	18 (12–23)	15 (10–21)	0.83
Follow-up, e-ASPECTS, median (IQR)	7 (4–9)	8 (7–9)	0.17
Time from the end of EVT procedure to follow-up NCCT, median (IQR), hours	36 (29–39)	37 (27–41)	0.88
NCCT follow-up ILV (Brainomix), median (IQR), mL	39 (18–62)	23 (15–39)	0.26
Follow-up DWI ASPECTS, median (IQR), mL	7 (5–8)	8 (6–8)	0.38
DWI ILV, median (IQR), mL	30 (11–105)	16 (7–32)	0.06
Time from treatment to follow-up MRI, median (IQR), hours	36 (28–39)	36 (27–41)	0.9
Clinical outcomes			
24-h NIHSS score, median (IQR)	7 (4–18)	5.5 (3–9)	0.15
7-day or discharge NIHSS score, median (IQR)	4 (1–18)	2.5 (1–5)	0.25
7-day or discharge Barthel index, median (IQR)	60 (10–90)	92.5 (56–100)	0.07
90-day modified Rankin scale, median (IQR)	4 (2–5)	1 (0–3)	0.02

mTICI modified thrombolysis in cerebral infarction, *SD* standard deviation, *IQR* interquartile range, *NIHSS* National Institutes of Health Stroke Scale, *ASPECTS* Alberta Stroke Program Early CT Score, *DWI* diffusion-weighted imaging, *MRI* magnetic resonance imaging, *NCCT* non-contrast-computed tomography, *ILV* ischemic lesion volume

Table 2 Selected characteristics of patients with independent outcome (modified Rankin scale 0 to 2) versus dependent outcome (modified Rankin scale ≥ 3)

	mRS 0–2 (n = 28)	mRS ≥ 3 (n = 17)	p
Affected hemisphere, left, n (%)	13 (46%)	7 (41%)	
Age, mean (SD), years	68.3 (±11.3)	70 (±9.8)	0.58
Sex, females, %	42.9	47.1	
Baseline NIHSS, median (IQR)	16 (13–18)	17 (15–18)	0.31
Time from symptom onset to reperfusion, median (IQR), min	170 (144–189)	165 (110–185)	0.55
24-h NIHSS, median (IQR)	4 (2–7)	13 (8–18)	< 0.001
DWI ILV, median (IQR), mL	10 (7–22)	59 (17–126)	0.002
Brainomix NCCT ILV, median (IQR), mL	21 (13–35)	39 (20–64)	0.012
7-day or discharge NIHSS, median (IQR)	2 (0–2)	11 (6–18)	< 0.001

SD standard deviation, *IQR* interquartile range, *NIHSS* National Institutes of Health Stroke Scale, *ASPECTS* Alberta Stroke Program Early CT Score, *DWI* diffusion-weighted imaging, *NCCT* non-contrast-computed tomography, *ILV* ischemic lesion volume

Statistical analysis

Standard descriptive statistics were used to measure the central tendency and variability of baseline characteristics. Ordinal/continuous variables were compared by the Mann-Whitney *U* test or the *t* test based on their distribution. All tests were two-sided and the significance level was considered as 0.05. Categorical variables were compared using the Fisher exact test. The Somer delta was used to test the differences between concordant and discordant ASPECTS regions, and McNemar's test was used to test the sensitivity and specificity. The Bland-Altman plot was calculated to compare the differences between

Brainomix and DWI ILVs. Statistical analyses were performed using STATA version 14.2 (College Station, TX).

Results

Fifty patients were treated with EVT during the 12-month study period; five patients were excluded due to hemorrhagic transformation (HT) or parenchymal hematoma (PH) on follow-up imaging. Table 1 demonstrates the clinical, radiological, and outcome (short-term and long-term) characteristics based on the recanalization status. Table 2 demonstrates selected characteristics of patients with independent outcome versus dependent outcome. Baseline CT "early ischemic changes" volumes were similar between the groups with and without successful reperfusion, 18 mL (interquartile range [IQR], 12–23) and 15 mL (IQR, 10–21), respectively. Successful recanalization (TICI2b-3) was achieved in 32 (71%) patients. Median follow-up NCCT ILV was 23 ml (IQR, 15–39 ml), and median follow-up DWI ILV was 11.5 ml (IQR, 7–32 ml) in the TICI2b-3 group. In the TICI 0-2a, the median NCCT ILV was 39 ml (IQR, 18–62 mL) and the median DWI ILV was 30 (IQR, 11–105 mL). The DWI ILVs in patients with good clinical outcome (mRS 0–2) were significantly lower compared with the patients with mRS ≥ 3 (10 mL vs 59 mL, *p* = 0.002). The similar trend was observed for the Brainomix ILV measurements (21 mL vs 39 mL, *p* = 0.012). There was a high correlation and accuracy in detection of follow-up ischemic changes in particular ASPECTS regions and cortical versus subcortical ASPECTS regions (Table 3). Bland-Altman plot comparing the differences between the Brainomix software and DWI ILV measurements is demonstrated in Fig. 2. Dot plot for particular Brainomix and DWI ILV measurements is presented in Fig. 3.

Table 3 Accuracy of concurrent (Brainomix and DWI-MRI) detection of ischemic changes in cortical and subcortical ASPECTS regions

ASPECTS regions	N (%)	Somer's delta*	p value**	Accuracy
M1	40 (89%)	0.98	1.00	0.98
M2	33 (73%)	0.91	1.00	0.98
M3	35 (78%)	0.70	0.25	0.93
M4	41 (91%)	0.75	1.00	0.98
M5	38 (84%)	0.86	1.00	0.98
M6	38 (84%)	1.00	NA	1.00
Caudate	25 (56%)	0.91	0.48	0.96
Lentiform	11 (24%)	0.97	1.00	0.98
Insula	28 (62%)	1.00	NA	1.00
Internal capsule	0 (0%)	NA	NA	NA
Cortical regions	225 (83%)	0.86	0.13	0.97
Subcortical regions	109 (61%)	0.96	1.00	0.98

***Somer's delta, testing the difference between the number of concordant and discordant pairs. McNemar's test, comparing the sensitivity and specificity of two diagnostics methods in the same group of patients. Cortical regions = M1 to M6; subcortical regions = Caudate, Lentiform, Insula, Internal capsule

Fig. 2 Bland-Altman plot comparing the differences between the Brainomix software and manual DWI ILV measurements
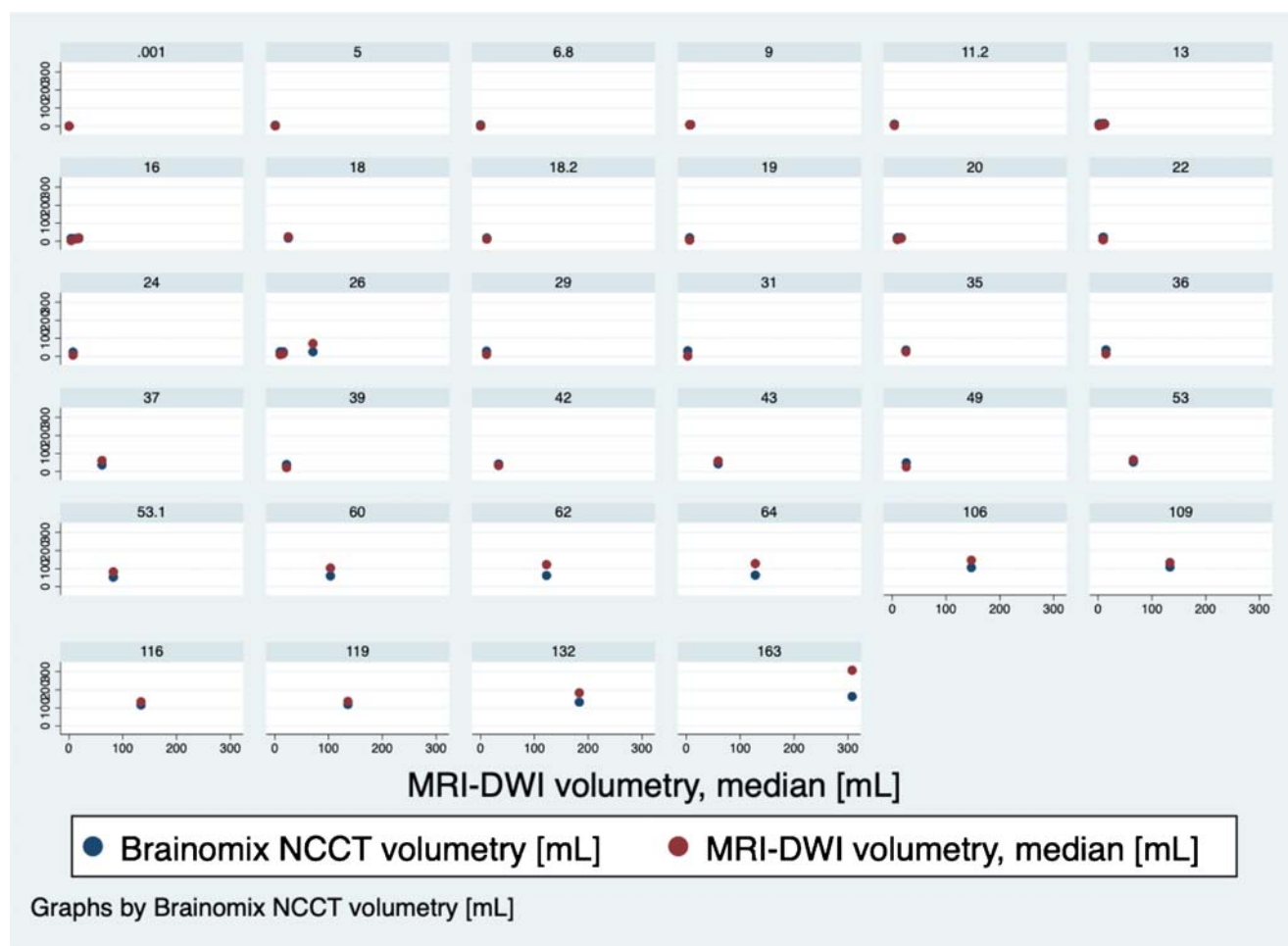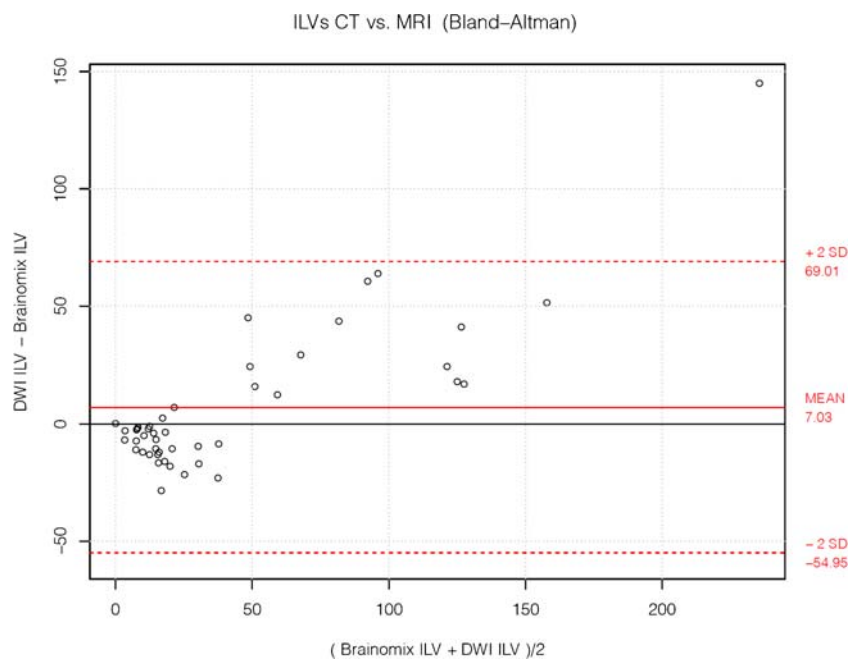
Fig. 3 Dot plot demonstrating particular ILV measurements. NCCT, non-contrast CT; MRI, magnetic resonance imaging; and DWI, diffusion-weighted imaging

Discussion

Our single-center observational study demonstrated a good correlation between the expert manual DWI ILV outlining and fully automated NCCT ILV (Brainomix) measurement in patients who underwent the EVT for anterior circulation stroke.

ILV on follow-up neuroimaging represents a biomarker of stroke treatment efficacy, and follow-up NCCT still represents the imaging modality of choice in most of the countries worldwide. Automated (software-based) measurement of ILV by a reliable software has a potential to overcome the shortcomings and time demands associated with manual infarct delineation. It was demonstrated previously that the Brainomix software was non-inferior compared with expert radiologists for baseline ASPECTS reading [7–9]. Its newer versions enable to measure the volume (extent) of early ischemic changes on baseline NCCT (in addition to the ASPECT score) and ILVs on follow NCCT in short processing time, which might be considered helpful and beneficial in everyday practice. We demonstrated that the Brainomix software and radiology experts detected the concurrent ischemic changes in particular ASPECTS regions on follow-up neuroimaging (NCCT vs DWI-MRI) with similar accuracy.

The relationship between ILV outlined manually on NCCT, DWI, or FLAIR at the median time of 42 h and favorable clinical outcome in patients receiving EVT was studied before and found no differences in outcome prediction regardless the used imaging modality—implying that providing a follow-up NCCT might be sufficient for ILV measurements [10]. Manual ILV delineation on 27-h NCCT or FLAIR was used in a study by Albers et al. [3]. Software-assisted manual outlining of 24 h-NCCT ILV by using the ITK-Snap software was used in a study by Bucker et al. [1]. In comparison to the studies mentioned above, our study was one of the first testing the clinical utility of fully-automated software in ILV measurement in a prospective EVT cohort. Since our study was limited by the number of enrolled patients, we did not test the association between ILVs and 3-month clinical outcome.

Our study has several limitations including small sample size and single-center experience. The ILVs were measured by using two different imaging techniques (NCCT vs DWI-MRI) and expert reading versus automated software, which have different infarct conspicuity. Additionally, the Brainomix software is not able to measure ILV in the presence of hemorrhagic transformation or parenchymal hemorrhage, which limits its generalizability and should be overcome in future versions of the Brainomix software.

In conclusion, NCCT ILV measured automatically by the Brainomix software might be considered an additional early radiological outcome measure in everyday clinical practice.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval All procedures were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all participants included in the study.

References

1. Bucker A, Boers AM, Bot JCJ, Berkhemer OA, Lingsma HF, Yoo AJ, van Zwam W, van Oostenbrugge R, van der Lugt A, Dippel DWJ, Roos YBWEM, Majoie CBLM, Marquering HA, MR CLEAN Trial Investigators (Multicenter Randomized Clinical Trial of Endovascular Treatment for Acute Ischemic Stroke in the Netherlands) (2017) Associations of ischemic lesion volume with functional outcome in patients with acute ischemic stroke: 24-hour versus 1-week imaging. Stroke 48:1233–1240. https://doi.org/10.1161/STROKEAHA.116.015156

2. Manning NW, Warne CD, Meyers PM (2018) Reperfusion and clinical outcomes in acute ischemic stroke: systematic review and meta-analysis of the stent-retriever-based, early window endovascular stroke trials. Front Neurol 9:301. https://doi.org/10.3389/fneur.2018.00301

3. Albers GW, Goyal M, Jahan R, Bonafe A, Diener HC, Levy EI, Pereira VM, Cognard C, Yavagal DR, Saver JL (2015) Relationships between imaging assessments and outcomes in solitaire with the intention for thrombectomy as primary endovascular treatment for acute ischemic stroke. Stroke 46:2786–2794. https://doi.org/10.1161/STROKEAHA.115.010710

4. Rangaraju S, Liggins JTP, Aghaebrahim A et al (2014) Pittsburgh outcomes after stroke thrombectomy score predicts outcomes after endovascular therapy for anterior circulation large vessel occlusions. Stroke 45:2298–2304. https://doi.org/10.1161/STROKEAHA.114.005595

5. Zaidi SF, Aghaebrahim A, Urra X, Jumaa MA, Jankowitz B, Hammer M, Nogueira R, Horowitz M, Reddy V, Jovin TG (2012) Final infarct volume is a stronger predictor of outcome than recanalization in patients with proximal middle cerebral artery occlusion treated with endovascular therapy. Stroke 43:3238–3244. https://doi.org/10.1161/STROKEAHA.112.671594

6. Van Everdingen KJ, Van Der Grond J, Kappelle LJ et al (1998) Diffusion-weighted magnetic resonance imaging in acute stroke. Stroke 29:1783–1790. https://doi.org/10.1161/01.STR.29.9.1783

7. Nagel S, Sinha D, Day D, Reith W, Chapot R, Papanagiotou P, Warburton EA, Guyler P, Tysoe S, Fassbender K, Walter S, Essig M, Heidenrich J, Konstas AA, Harrison M, Papadakis M, Greveson E, Joly O, Gerry S, Maguire H, Roffe C, Hampton-Till J, Buchan AM, Grunwald IQ (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(6):615–622. https://doi.org/10.1177/1747493016681020

8. Herweh C, Ringleb PA, Rauch G, Gerry S, Behrens L, Möhlenbruch M, Gottorf R, Richter D, Schieber S, Nagel S (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. https://doi.org/10.1177/1747493016632244

9. Pfaff J, Herweh C, Schieber S, Schönenberger S, Bösel J, Ringleb PA, Möhlenbruch M, Bendszus M, Nagel S (2017) E-ASPECTS correlates with and is predictive of outcome after mechanical thrombectomy. Am J Neuroradiol 38:1594–1599. https://doi.org/10.3174/ajnr.A5236

10. Yoo AJ, Chaudhry ZA, Nogueira RG, Lev MH, Schaefer PW, Schwamm LH, Hirsch JA, González RG (2012) Infarct volume is a pivotal biomarker after intra-arterial stroke therapy. Stroke 43:1323–1330. https://doi.org/10.1161/STROKEAHA.111.639401