



# Can machine learning of post-procedural cone-beam CT images in acute ischemic stroke improve the detection of 24-h hemorrhagic transformation? A preliminary study

Valerio Da Ros<sup>1</sup> · Andrea Duggento<sup>1</sup> · Armando Ugo Cavallo<sup>1</sup> · Luigi Bellini<sup>1</sup> · Francesca Pitocchi<sup>1</sup> · Nicola Toschi<sup>1</sup> · Alfredo Paolo Mascolo<sup>2</sup> · Fabrizio Sallustio<sup>2</sup> · Francesca Di Giuliano<sup>1</sup> · Marina Diomedì<sup>2</sup> · Roberto Floris<sup>1</sup> · Francesco Garaci<sup>1</sup> · Kamil Zelenák<sup>3</sup> · Iliaria Maestrini<sup>2</sup> 

Received: 5 August 2022 / Accepted: 12 October 2022 / Published online: 25 October 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

**Purpose** Hemorrhagic transformation (HT) is an independent predictor of unfavorable outcome in acute ischemic stroke (AIS) patients undergoing endovascular thrombectomy (EVT). Its early identification could help tailor AIS management. We hypothesize that machine learning (ML) applied to cone-beam computed tomography (CB-CT), immediately after EVT, improves performance in 24-h HT prediction.

**Methods** We prospectively enrolled AIS patients undergoing EVT, post-procedural CB-CT, and 24-h non-contrast CT (NCCT). Three raters independently analyzed imaging at four anatomic levels qualitatively and quantitatively selecting a region of interest (ROI) < 5 mm<sup>2</sup>. Each ROI was labeled as “hemorrhagic” or “non-hemorrhagic” depending on 24-h NCCT. For each level of CB-CT, Mean Hounsfield Unit (HU), minimum HU, maximum HU, and signal- and contrast-to-noise ratios were calculated, and the differential HU-ROI value was compared between both hemispheres. The number of anatomic levels affected was computed for lesion volume estimation. ML with the best validation performance for 24-h HT prediction was selected.

**Results** One hundred seventy-two ROIs from affected hemispheres of 43 patients were extracted. Ninety-two ROIs were classified as unremarkable, whereas 5 as parenchymal contrast staining, 29 as ischemia, 7 as subarachnoid hemorrhages, and 39 as HT. The Bernoulli Naïve Bayes was the best ML classifier with a good performance for 24-h HT prediction (sensitivity = 1.00; specificity = 0.75; accuracy = 0.82), though precision was 0.60.

**Conclusion** ML demonstrates high-sensitivity but low-accuracy 24-h HT prediction in AIS. The automated CB-CT imaging evaluation resizes sensitivity, specificity, and accuracy rates of visual interpretation reported in the literature so far. A standardized quantitative interpretation of CB-CT may be warranted to overcome the inter-operator variability.

**Keywords** Ischemic stroke · Thrombectomy · Cone-beam computed tomography · Machine learning · Cerebral hemorrhage

---

Valerio Da Ros and Andrea Duggento contributed equally to this work.

✉ Iliaria Maestrini  
ilamaestrini@hotmail.com

<sup>1</sup> Department of Biomedicine and Prevention, University Hospital of Rome “Tor Vergata”, Viale Oxford 81, Rome, Italy

<sup>2</sup> Stroke Center, Department of Systems Medicine, University Hospital of Rome “Tor Vergata”, Viale Oxford 81, 00133 Rome, Italy

<sup>3</sup> Clinic of Radiology, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, 03659 Martin, Slovakia

## Introduction

Hemorrhagic transformation (HT) represents the most severe and dreaded complication of acute ischemic stroke (AIS) and is an independent predictor of unfavorable outcome. The rate of HT is increased by revascularization procedure and is estimated to occur as symptomatic in 4.4% in the pooled analysis from 5 endovascular thrombectomy (EVT) trials [1] and in up to 6–7% of treated patients in both extended time window EVT trials [2, 3]. However, other factors might drive an increased risk of HT, such as sustained post-procedural hypertension and higher blood pressure variability within 24 h after EVT [4].

Follow-up brain imaging is usually performed within 24 h, or before in cases of neurological worsening, to detect HT, guide antithrombotic treatment regimen, even peri-procedural, and evaluate final infarct volume [5].

In the last years, the possibility of detecting peri-interventional intracranial hemorrhagic complications directly in the angiographic suite, using cone-beam computed tomography (CB-CT) scan, has gained increased attention [6]. However, most of the studies using CB-CT immediately after EVT have applied only a qualitative visual assessment [7, 8] to detect HT resulting in a variable predictive performance. The diagnostic challenge remains to differentiate abnormal post-EVT contrast staining (mimicking subarachnoid, intraventricular, or intracerebral hemorrhage) from actual intracranial hematomas [6–10]. Thus, the early detection of hemorrhagic outcomes using CB-CT immediately after EVT could improve the clinical care management of AIS patients. To date, there is only one study using an objective approach based on maximum-Hounsfield Unit (HU) measurements on CB-CT hyperdense lesions to predict 24-h HT [10]. This study showed promising results with an area under the curve (AUC), sensitivity, and specificity of 0.805, 96%, and 84%, respectively [10].

Machine learning (ML) application has been shown to improve the diagnostic accuracy of imaging in many fields, including AIS [11–13]. We hypothesize that the automated extraction of quantitative data from predefined regions of CB-CT immediately after EVT could help predict the 24-h HT risk in AIS patients and may implement visual neuroimaging assessment in the acute setting of AIS.

## Materials and methods

### Patient population

Between 15 January and 30 April 2022, all consecutive patients admitted at our Comprehensive Tertiary Stroke Center (“Tor Vergata” University Hospital of Rome, Italy), with a diagnosis of anterior circulation AIS due to large vessel occlusion (LVO) confirmed at brain non-contrast-computed tomography (NCCT) and CT-angiography undergoing EVT, immediate post-procedural CB-CT and 24-h post-treatment NCCT, were prospectively enrolled in this study and retrospectively reviewed.

We collected relevant demographic data including age, race/ethnicity, and sex, as well as clinical data, i.e., site of occlusion, stroke severity, pre-treatment modified Rankin Scale, time delays, cardiovascular risk factor, and stroke etiology, to define the population. Neuroradiological features were also collected including baseline Alberta Stroke Program Early CT Score (ASPECTS) on NCCT and Clot Burden Score (CBS) on CT-angiography. Final reperfusion

status after EVT was graded according to the modified TIC1 (mTIC1) score, in which mTIC1 2b, 2c, or 3 was accepted as successful recanalization [14]. The pre-written study protocol retraced our institutional clinical care routine and was conducted following the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. The local ethics committee approved prospective data collection on patients undergoing EVT (Registro Sperimentazioni, R.S. 25/18). Informed consent was obtained from all individual participants included in the study. Data is available upon reasonable request by the corresponding author.

### Image acquisition

Brain NCCT and CT-angiography were performed using a GE LightSpeed Plus 256 multislice CT scanner with conscious sedation when clinically indicated. CB-CT was acquired using a monoplane neuro-angiography X-ray system (AlluraClarity FD20; Philips Healthcare, Best, the Netherlands). The sensor area of the detector measures  $30 \times 40$  cm and consists of  $2586 \times 1904$  pixels. The acquisition protocol used was the commercially available “Xper-CT HD” protocol using the manufacturer’s default settings. The acquisition protocol consists of a rotational trajectory over  $200^\circ$  while acquiring 620 projection images at 30 frames per second for 20.8 s. The X-ray tube voltage was set to 120 kV, the focal spot to 0.7 mm, and a copper filter of 0.4 mm was used. The associated CT Dose Index measures 45 mGy and the effective dose measures 1.6 mSv [15, 16]. Pre-processing steps of the projection images include offset correction, gain correction, scatter correction, and water beam-hardening correction.

### Imaging evaluation

All images including immediate post-EVT CB-CT and 24-h NCCT were independently reviewed and graded by three raters (one radiology resident [F.P.], one senior neurointerventionalist [V.D.R.], and one senior neurologist [I.M.] with more than 5 years of experience) blinded to clinical data.

Images were analyzed on the same workstation with pre-set window levels. From both CB-CT and 24-h NCCT, four slices were selected at the following levels: (i) basal ganglia, (ii) insula, (iii) bodies of the lateral ventricles, and (iv) semioval centers height.

CB-CTs were read at a slice thickness of 3 mm, to match that of the 24-h NCCT. Each set of images was randomly and blindly analyzed in both a qualitative and quantitative fashion.

Image quality was assessed using a scoring system with a scale of 0 to 3, modeled on previously published techniques [17]: 0 — non-diagnostic or severely compromised, 1 — moderately compromised, 2 — mildly compromised, and 3 — excellent. The image quality of CB-CT was

further compared with 24-h NCCT and graded as worse or equal by the readers in consensus.

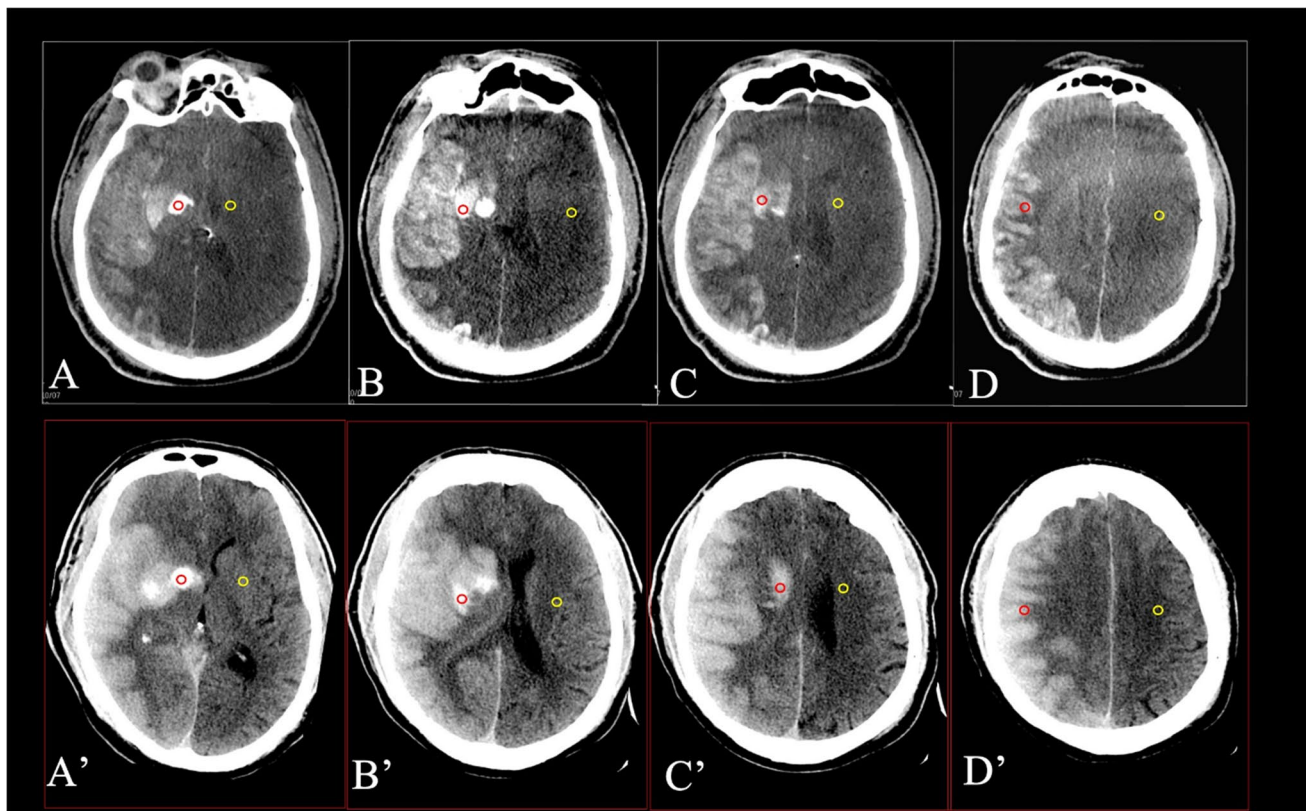
On CB-CT, four different patterns of pathology were identified: (i) hypodense pattern suggesting ischemic area; (ii) highly hyper-dense pattern in the subarachnoid space suggesting subarachnoid hemorrhage (SAH); (iii) smoothly hyper-dense pattern suggesting “parenchymal contrast staining” (PCS) due to the blood–brain barrier defect; (iv) highly hyper-dense pattern in the brain parenchyma suggesting intraparenchymal HT. All the three readers recognized one or more pathological patterns, when present, for each of the four slices abovementioned, in both CB-CT and 24-h NCCT. The density in HU of the pathological area was measured using a region of interest (ROI)  $< 5 \text{ mm}^2$ . In case more than one pattern was recognized in the same slice, ROIs were placed in the parenchymal areas where the density appeared more affected (in terms of hyperdensity). The same measurements were performed also in the analogous anatomical structures of the unaffected hemisphere. The differential HU-ROIs value was assessed between affected and unaffected brain hemispheres for each selected slice.

At 24-h NCCT, the HT was defined as any hyperdense lesion in the context of the infarcted area and then classified according to the radiological European Cooperative Acute Stroke Study (ECASS) I definition [18] in hemorrhagic infarction (HI1 or 2) and parenchymal hematoma (PH1 or 2). Pure SAH (or mixed to contrast extravasation) was defined as subarachnoid hyperdense lesions caused by vessel rupture and was similarly assessed on 24-h NCCT [19]. Whenever it was not possible to disentangle the pathological pattern at 24-h NCCT, a further follow-up neuroimaging performed within 7 days after symptoms’ onset was used to define the pattern. As a surrogate measure of lesion volume, we computed the number of anatomic levels affected for each patient, with a 4-point ordinal scale according to the possibility of having from 1 to 4 slices affected.

Figure 1 reports an example of the sampling modality of the ROIs in CB-CT and 24-h NCCT slices.

### Statistical analysis

We used the Statistical Package for the Social Sciences 22 for Windows to determine median values, interquartile



**Fig. 1** In the first line, we reported the region of interests (ROIs) placed on cone-beam computed tomography (CB-CT) slices at the level of the most hyperdense area at four predefined anatomical levels: (A) basal ganglia, (B) insula, (C) bodies of the lateral ventricles, and (D) semioval centers. At the bottom, we reported the correspond-

ing ROIs on 24-h non-contrast-enhanced computed tomography (NCCT) at the same four levels (A'–D'). CB-CT image quality was scored 2 (mildly compromised), and the comparison with the image quality of 24-h NCCT was graded as equal

ranges, and percentages. We compared groups for categorical variables with the  $\chi^2$  test with Yates’s correction or Fisher’s exact test, and for continuous variables with the Mann–Whitney  $U$  test.  $p$  Values lower than 0.05 were considered significant. Interrater reliability was assessed by using  $\kappa$  statistics.

In each ROI, we considered the average [Mean (HU)], standard deviation [STD (HU)], minimum [Min (HU)], and maximum values of HU [Max (HU)], in both ROIs of the pathological and unaffected hemisphere. We calculated the signal-to-noise ratio (SNR), and the contrast-to-noise ratio (CNR) as follows:

$$SNR = \frac{Mean(HU)_{path.}}{STD(HU)_{path.}}$$

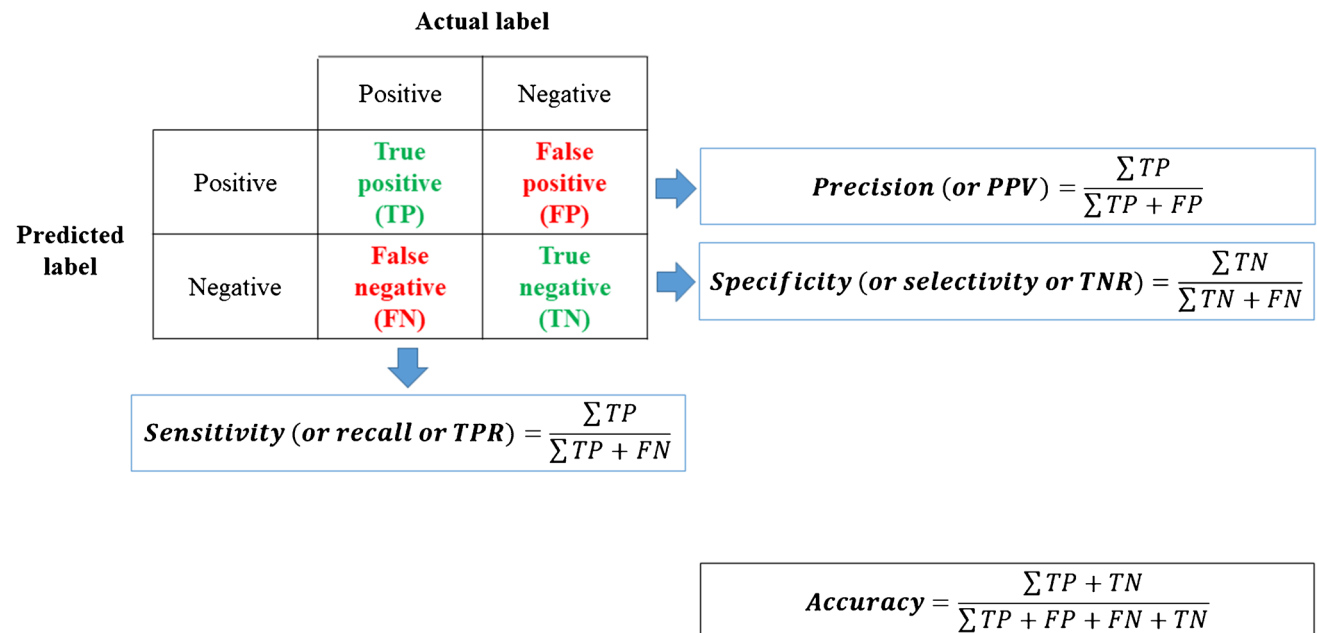
$$CNR = \frac{Mean(HU)_{path.} - Mean(HU)_{unaff.}}{\sqrt{STD(HU)_{path.}^2 + STD(HU)_{unaff.}^2}}$$

### Machine learning analysis

We tested whether mean (HU), SNR, CNR, max (HU), min (HU), and the number of anatomic levels affected could jointly be used to predict at 24-h follow-up two different outcomes: hemorrhagic (HC: HT and SAH) from non-hemorrhagic (non-HC: ischemic, PCS or “unremarkable” brain parenchyma) complications. Because the sample size was

too small (172 dyads of ROIs) to allow for a train-test split of the dataset, we adopted the leave-one-out validation technique, a special class of cross-validation technique, which ensures unbiased estimates of accuracy on unseen data. With leave-one-out cross-validation, we excluded one observation from ML training and used it for validation; this was repeated with all samples in the dataset. We tested several algorithms for ML classification: support vector machines, quadratic classifiers, decision trees, and artificial neural networks (with either one or two fully connected layers). All algorithms have been implemented in Python code with the scikit-learn package [20]. Leave-one-out validation was also needed to exclude classifiers that overperformed in the training set but underperformed in the validation set, indicating model overfitting resulting in a lack of generalization. The ML training was performed by a decision tree method (one decision tree per leave-one-out iteration). We selected the following figures of merit for model performances: (i) accuracy, (ii) specificity (also called selectivity), (iii) precision (also called positive predictive value (PPV)), and (iv) sensitivity (also called recall or true positive rate (TPR)) (Fig. 2).

All four figures of merit for performances are functions of the working threshold of the classifier, also called the “operating point.” The higher the threshold is set, the higher the specificity will be obtained but the lower the sensitivity will be. The working threshold was chosen based on the clinical needs, i.e., whether sensitivity and specificity are both equally important, or rather there is a criterion that suggests compromising sensitivity in favor of specificity or vice



**Fig. 2** Figures of merit for model performances: (i) accuracy (ACC); (ii) specificity, also called selectivity, or true negative rate (TNR); (iii) precision, also called positive predictive value (PPV), and (iv) sensitivity, also called recall or true positive rate (TPR). Adapted from Ma et al. [24]

versa. For example, the optimal working threshold is chosen maximizing the harmonic mean of the PPV and TPR, with additional weights if precision and recall have different clinical importance. For this purpose, the F $\beta$ -score is introduced:

$$F_{\beta} = (1 + \beta^2) \frac{\text{PPV} \times \text{TPR}}{(\beta^2)\text{PPV} + \text{TPR}}$$

where  $\beta$  is commonly chosen between three values:  $\beta = 1$  (unweighted harmonic mean PPV and TPR),  $\beta = .5$  (which weighs recall lower than precision), and  $\beta = 2$  (which weighs recall higher than precision).

Given our need to determine the risk of 24-h hemorrhage, the F2-score ( $\beta = 2$ ) was chosen as a figure of merit the machine learning algorithm was set to maximize.

## Results

### Population

Out of 50 patients, we enrolled 43 patients with a confirmed diagnosis of AIS due to anterior LVO undergoing immediate post-EVT CB-CT and 24-h NCCT. Seven patients were excluded since they did not undergo CB-CT for logistical issues related to the use of a secondary angio suite not equipped with CB-CT. Clinical, demographic, and neuroradiological characteristics and those related to treatment are reported in Table 1. Out of 43 patients, 23 presented HC (6 SAH and 17 HT). Out of 17 patients who developed HT, 5 patients presented a PH2 subtype, 9 patients a PH1 subtype, 2 patients HI2, and 1 patient HI1. For a comparison between CB-CT and 24-h NCCT scans, see Fig. 3.

According to clinical and demographic characteristics, patients who developed HC at 24 h differed from those non-HC for age (median age 64.5 years [IQR 56–75.5] versus 81 [IQR 73–82], respectively), and arterial hypertension (~61% vs 90%,  $p = 0.039$ ) (Table 1). Considering neuroradiological features, patients with HC compared to those with non-HC presented a lower ASPECTS (mean value 9 [IQR 8–10] vs 10 [IQR 9–10], respectively), and CBS (mean value 6 [IQR 3–6] vs 7 [IQR 6–8], respectively) on baseline NCCT and a higher number of anatomic levels affected on immediately post-procedural CB-CT (mean value 3 slices [IQR 1–3] vs 0 slices [IQR 0–2], respectively). There were no differences regarding treatment modalities, delays, mTICI scores, or functional outcome except for a higher 24-h NIHSS score in HC compared to non-HC patients (mean value 18 [IQR 13–21] vs 8 [IQR 6–18.5], respectively) (Table 1).

Out of 172 ROIs extracted from affected hemispheres, 92 ROIs were classified unremarkable, 5 were classified as PCS, 29 were classified as ischemia, 7 ROIs were classified as SAH, and 39 ROIs were classified as HT. The summary

statistics for the 172 dyads of ROI extracted from the 43 patients are reported in Table 2. The degree of agreement in image interpretation between the three raters was excellent ( $k = 0.9$ ). We needed to analyze further follow-up imaging for 7 patients (16.3%) to disentangle the neuroradiological pattern.

### Image quality

Mean image quality was significantly lower for CB-CT (1.88 [IQR 1–3]) than for 24-h NCCT (2.51 [IQR 2–3];  $p < 0.001$ ).

### Predictive value of CB-CT metrics for hemorrhagic complications

The median values of mean (HU), min (HU), max (HU), and CNR were all higher in the HC group (SAH or HT) compared to the non-HC group (unremarkable parenchyma, PCS, or ischemia). Median SNR showed no statistically significant association between “HC vs non-HC” groups.

For summary purposes, Fig. 4 shows the distribution of all samples concerning mean (HU) and CNR.

### Machine learning analysis

When the F2-score was chosen as a figure of merit (thus prioritizing sensitivity over specificity) and when all metrics were available, the ML method that best performed on the validation was a Bernoulli Naive Bayes Classifier with MinMaxScaler (with Additive Laplace/Lidstone) smoothing parameter  $\alpha = 1$  (one classifier per leave-p-out iteration). This latter was elected as the ML method classifier of choice.

In Fig. 5, results from the ML predictor are shown. Accuracy for classifying HC from non-HC cases was 82.0% with an AUC = 0.877, precision = 59.7%, sensitivity (recall) = 100%, and specificity = 75.4%.

## Discussion

The present study reported our preliminary experience of ML application on CB-CT performed immediately after EVT in patients with AIS, to predict the risk of 24-h HC. We proposed a new methodological approach for automatic early image analysis interpretation based on the Bernoulli Naive Bayes Classifier with MinMaxScaler with a 100% recall rate. We observed specificity of 75% and precision of ~60% by the trained model in HC prediction on unseen data.

Our methodology presents some strengths of innovation. First, the use of the highest HU values average hyperdense lesion on a single CB-CT slice compared with the contralateral unremarkable hemisphere resulted in a less time-consuming HC evaluation with an accuracy close to the previously

**Table 1** Bivariate comparison between patients with and without hemorrhagic complications (HC: subarachnoid hemorrhage and hemorrhagic transformation) according to baseline characteristics of the study population

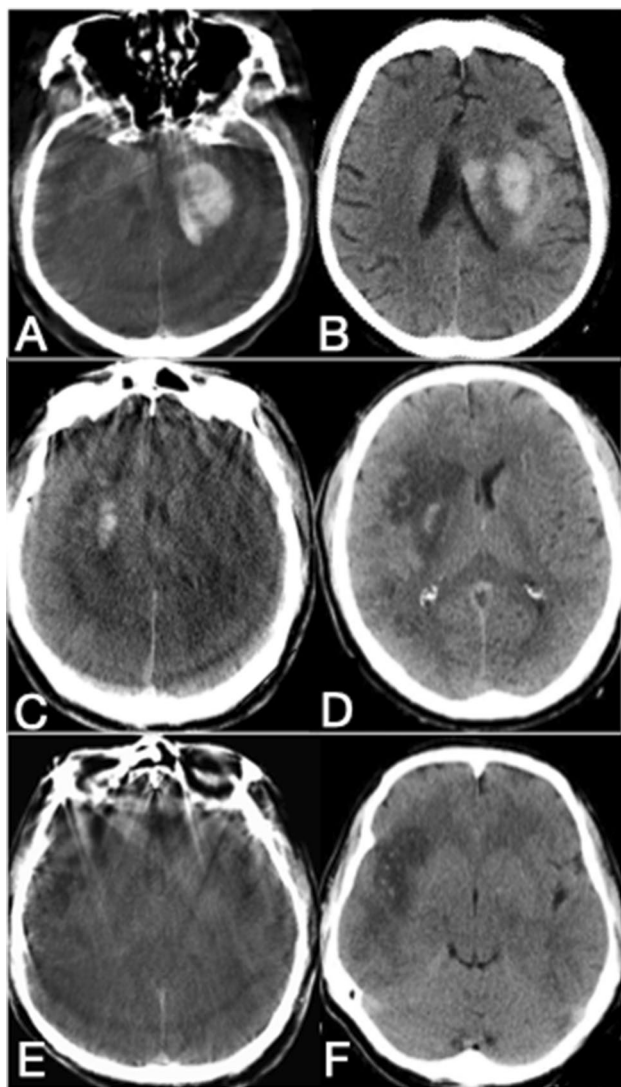
	All (n=43)	HC (n=23)	Non-HC (n=20)	p value
Demographic characteristics and vascular risk factors				
Age <sup>a</sup>	77 (69–83)	73 (64–80)	80 (74.5–84.5)	<b>0.031</b>
Male	26 (60.4)	15 (65.2)	11 (55.0)	0.545
Arterial hypertension	32 (74.4)	14 (60.9)	18 (90.0)	<b>0.039</b>
Hypercholesterolemia	15 (34.9)	9 (39.1)	6 (30.0)	0.749
Diabetes mellitus	10 (23.2)	4 (17.4)	6 (30.0)	0.473
Smoking habit	12 (27.9)	6 (26.1)	6 (30.0)	0.998
Atrial fibrillation	10 (23.2)	5 (21.7)	5 (25.0)	0.997
Ongoing antiplatelet therapy	22 (51.2)	9 (39.1)	13 (65.0)	0.129
Ongoing anticoagulant therapy	3 (7.0)	1 (4.3)	2 (10.0)	0.590
mRS before stroke	1 (0–1)	1 (0–1)	1 (1–1)	0.052
Clinical presentation				
Baseline NIHSS score <sup>a</sup>	18 (15–20)	18 (15–21)	17.5 (10–20)	0.212
Baseline SBP, mmHg <sup>a</sup>	150 (125–165)	150 (120–165)	147.5 (130–167.5)	0.599
Baseline DBP, mmHg <sup>a</sup>	80 (75–90)	90 (77–100)	80 (75–85)	0.901
Blood glucose concentration, mg/dl <sup>a</sup>	130 (105–153)	124 (103–140)	141.5 (107–168.5)	0.318
Wake-up stroke on unknown onset	17 (39.5)	8 (34.8)	9 (45.0)	0.545
TOAST classification				0.279
Undetermined origin	16 (37.2)	11 (47.8)	5 (25.0)	
Cardioembolic	24 (55.8)	11 (47.8)	13 (65.0)	
Large vessel occlusion	3 (7.0)	1 (4.3)	2 (10.0)	
Neuroradiological features				
Baseline ASPECTS <sup>a</sup>	9 (9–10)	9 (8–10)	10 (9–10)	<b>0.035</b>
Baseline CBS score <sup>a</sup>	6 (6–8)	6 (3–6)	7 (6–8)	<b>0.029</b>
Levels affected in CB-CT, nr <sup>a</sup>	1 (0–3)	3 (1–3)	0 (0–2)	<b>0.001</b>
Site of occlusion				0.351
Isolated MCA-M1	26 (60.5)	14 (60.9)	12 (60.0)	
Isolated MCA-M2	5 (11.6)	1 (4.3)	4 (20.0)	
Tandem	2 (4.6)	2 (8.7)	–	
Terminal ICA L-type	6 (13.9)	4 (17.4)	2 (10.0)	
Terminal ICA T-type	4 (9.3)	2 (8.7)	2 (10.0)	
Revascularization treatment				
Rt-PA	21 (48.8)	12 (52.2)	9 (45.0)	0.763
1st admission to telestroke spoke center	16 (37.2)	10 (43.5)	6 (30.0)	0.528
Onset-to-groin time, min <sup>a</sup>	298 (242–410)	288 (174–410)	304 (249–407)	0.733
Onset-to-recanalization time, min <sup>a</sup>	310 (260–418)	297 (240–421)	327 (271–418)	0.761
Onset-to-CB-CT time, min <sup>a</sup>	324 (285–430)	324 (289–430)	340.5 (277.5–435)	0.817
Device passages, nr <sup>a</sup>	2 (1–3)	2 (1–3)	1 (1–2)	0.051
Procedure duration, min <sup>a</sup>	20 (14–30)	22 (15–31)	18.5 (14–28)	0.770
Orotracheal intubation	5 (11.6)	1 (4.3)	4 (20.0)	0.167
Outcome				
TICI score $\geq$ 2b	36 (83.7)	20 (87.0)	16 (80.0)	0.687
Mortality rate	4 (9.3)	4 (17.4)	–	0.111
24-h NIHSS score <sup>a</sup>	14 (7–20)	18 (13–21)	8 (6–18.5)	<b>0.047</b>

Bold values denote statistical significance at the  $p < 0.05$  level

mRS modified Rankin Scale, NIHSS National Institute of Health Stroke Scale, SBP systolic blood pressure, DBP diastolic blood pressure, ASPECTS Alberta Stroke Program Early Computed Tomography Score, CBS clot burden score, MCA middle cerebral artery, ICA internal carotid artery, rt-PA recombinant tissue-plasminogen activator, CB-CT cone-beam computed tomography, TICI thrombolysis in cerebral infarction

Unless specified, values are the number of patients (%)

<sup>a</sup>Median (interquartile range)



**Fig. 3** (A) The machine learning (ML) evaluation of the cone-beam computed tomography (CB-CT) slice at the level of the bodies of the lateral ventricles performed immediately after successful endovascular thrombectomy (EVT) of the left MCA-M1 occlusion, and (B) 24-h non-contrast-enhanced computed tomography (NCCT) of the same patients detected a PH2 hemorrhagic transformation (HT); (C) the ML evaluation of the CB-CT slice at the level of the basal ganglia performed immediately after MT detected with good accuracy the PH1-HT but, less sharply the surrounding area of ischemic core hypodensity, and (D) 24-h NCCT of the same patients confirmed the PH1-HT with better delineation of the final ischemic core. (E) ML evaluation of the CB-CT slice at the level of the insula performed immediately after MT demonstrates with low-accuracy differences in hypodensity between the right affected insula and the left unaffected area, while the 24-h NCCT (F) of the same patient better delineates the final ischemic core at the level of the right insula secondary to unsuccessful right MCA-M2 occlusion

proposed whole lesion(s) average HU evaluation, proposed by Chen et al. [10]. Moreover, the differential HU-ROI measurements between affected (ROIs placed in the most hyperdense parenchymal area) and unaffected brain hemisphere performed at four different anatomic levels provide a panel

of data related to HU distribution in ROIs (Mean, Min, Max, SNR, and CNR). This last was used to train the ML algorithm and can potentially provide its further training.

Second, the ML analysis allows unbiased estimation of statistical metrics (accuracy, AUC, precision, recall) performances of the model trained on most of the data (i.e., 80%) in predicting an outcome on unseen data (i.e., 20%). In this way, the selection of classifiers overperformed in the training but underperformed in the validation set was avoided, testing the real validity and reproducibility of the model selected.

Third, by choosing the F2-score as a training learning goal, we explicitly searched for a ML model biased toward a higher recall supported by the need for a more conservative clinical approach, indulgent with type I errors (false positive rate), which minimizes type II errors (false-negative rate). In other words, we admitted the possibility to overestimate the rate of HC rather than underestimate it. Our results reflect this approach: in leave-*p*-out cross-validation, we obtained the highest possible sensitivity of 100%, corresponding to an ML algorithm that would report all HC, at the expense of a sub-optimal specificity (75%), which result in falsely reporting as “perspective hemorrhagic case” approximately one-fourth of patients with benignant outcomes. We considered this performance acceptable from a clinical perspective. Indeed, the objective interpretation of CB-CT with the use of ML in the angio suite immediately after or during EVT could be retranslated not only in a safer and tailored medical management in an acute setting (i.e., in the case of antiplatelet therapy administration during emergent stenting for AIS due to tandem occlusions) but also to eventually redefine a stricter neuroimaging follow-up (i.e., before the “standard” 24 h) in selected patients.

Despite these encouraging results, our study showed a roughly 60% precision for HC prediction. Although our model was trained by operator-defined ROI, this performance significantly differs from the sensitivity, specificity, and accuracy rates (often over 90%) of visual CB-CT assessment reported in the literature so far [7, 8, 10]. We may speculate that the low ML precision resizes the post-treatment CB-CT visual subjective performance, highlighting the need for a more standardized quantitative one, to overcome the operator variability interpretation.

Moreover, our ML method still has room for improvement as well. In fact, despite our good experience with the routine evaluation of CB-CT post-thrombectomy images, our ML approach is mainly based on the manually performed input we used, hampering the translation of complex disease patterns into a finite number of feature descriptors.

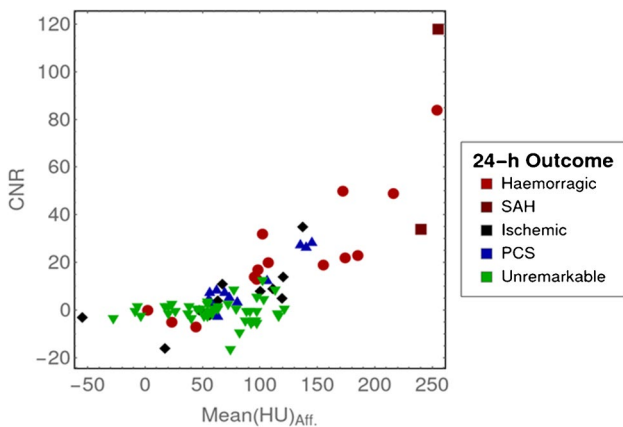
The use of a more sophisticated ML framework such as deep learning, based on deep convolutional neural networks, most used for pattern recognition tasks in images, automatically extracts relevant features from the training samples without manually designed features required as input. This artificial intelligence approach, when adequately trained, is

**Table 2** The median value within each pattern at the evaluation of the cone-beam computed tomography slice: mean HU [*Mean (HU)*], the minimum HU inside the ROI [*Min (HU)*], the maximum HU

inside the ROI [*Max (HU)*], SNR, and CNR on the affected side. Differences in median values ( $\Delta$ ) of *mean (HU)* within each pattern between the two sides (affected and unaffected)

Outcome		ROIs (n)	Affected side					$\Delta$ values between 2 sides Mean (HU)
			Mean (HU)	Min (HU)	Max (HU)	SNR	CNR	
Non-HC	Unremarkable	92	51.5	41.0	61.9	18.2	-0.25	-0.7
	PCS	5	111.8	95.4	126.6	12.6	9.2	28.5
	Ischemia	29	84.6	74.5	98.2	51.8	9.8	28.4
HC	SAH	7	167.7	153.6	168.6	155.1	47.1	110.4
	HT	39	111.7	97.1	124.9	43.5	22.5	70.0

HC hemorrhagic complications, ROI region of interest, PCS parenchymal contrast staining, HT hemorrhagic transformation, SAH subarachnoid hemorrhage, SNR signal-to-noise ratio, CNR contrast-to-noise ratio



**Fig. 4** Scatter plot of contrast-to-noise ratio (CNR) vs mean (HU) in five groups

expected to be more robust and accurate to the wide range of variations in features between different hemorrhagic/classes of hyper-densities on CB-CT.

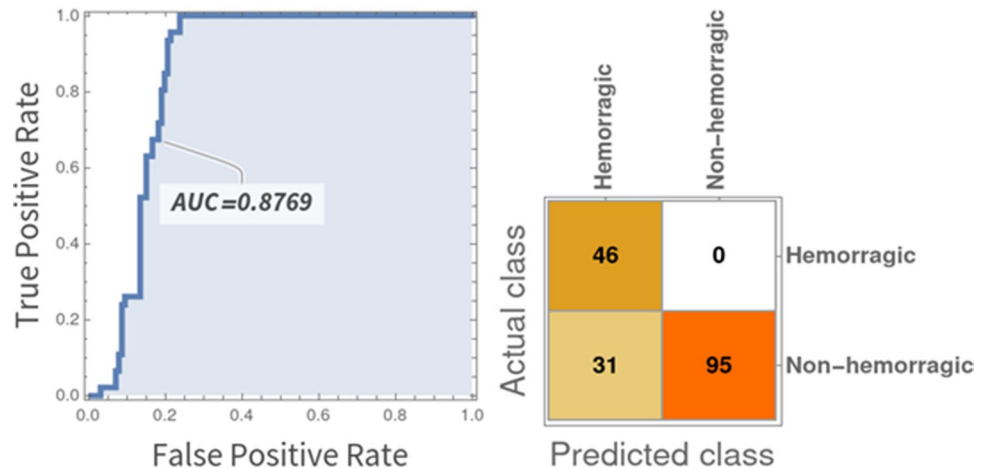
Importantly, since the deep learning process is automated, as long as the training set is much larger and more diverse,

it can easily analyze thousands of cases that even human experts may not be able to see and memorize in their lifetime [21]. If properly developed, validated, and implemented, this more sophisticated approach can be expected to be a more efficient data analytics artificial intelligence tool that might complement the experience of clinicians and improve patient care in terms of accuracy and workflow.

Our study presents three main limitations. First, the sample size was small due to the exploratory nature of the hypothesis. Nevertheless, the methodological procedure chosen, allowed us to extract data from 344 regions to perform an ML analysis.

Second, after the post-procedural CB-CT, HC was assessed only with 24-h NCCT with no systematically additional imaging follow-up, if not deemed necessary. The main reason is that our center represents a Comprehensive Stroke Center, so follow-up imaging was frequently performed in the Primary Stroke Centers where patients were further addressed. However, we specifically designed the study to assess the predictive value of CB-CT related to 24-h NCCT patterns, which is the most relevant question in clinical

**Fig. 5** Left: area under the curve (AUC) obtained by classification of hemorrhagic complications vs. non-hemorrhagic complications as obtained by a Bernoulli Naive Bayes Classifier trained over the whole dataset. Right: confusion matrix of the final classification





practice. Moreover, the persistence of hyper-densities at NCCT performed after 19 to 24 h has been demonstrated to strongly predict HT (specificity 100%, sensitivity 62.5%) and to be the only reliable method to differentiate PCS from HT [22]. Finally, in our population, symptomatic intracranial hemorrhage occurred in ~ 11% of cases, which is higher than classically reported rates of clinical trials and prospective stroke registries [1–3]. We could not exclude a selection bias due to the exclusion of a little share of patients for the already mentioned unavailability of the CB-CT in the secondary AIS angio suite. In fact, the percentage of symptomatic intracranial hemorrhage reported elsewhere in our population was 3.7%, which is close to the known clinical trial and stroke registries [23]. Nonetheless, this bias did not affect the study quality or the results' generalizability, since our research question concerns the discriminative capability of CB-CT in predicting HC.

## Conclusions

To the best of our knowledge, this is the first study applying ML to CB-CT performed immediately after EVT. The proposed approach seems to resize the few results available so far in the literature regarding the visual qualitative interpretation of the CB-CT images after EVT. Despite our preliminary data with ML being suboptimal, this tool could be adequately implemented, and this novel standardized quantitative CB-CT imaging analysis is warranted to overcome the inter-operator variability of interpretation, representing a further step towards a more precise and personalized medicine for our patients.

**Author contribution** Dr. Da Ros, Dr. Duggento, and Dr. Maestrini contributed substantially to the study's conception and design. Material preparation and data collection were performed by Dr. Da Ros, Dr. Bellini, Dr. Pitocchi, Dr. Mascolo, and Dr. Maestrini. Neuroimaging analysis was performed by Dr. Da Ros, Dr. Pitocchi, and Dr. Maestrini. Statistical analyses were performed by Dr. Duggento, Dr. Cavallo, and Dr. Maestrini. All authors made a substantial contribution to data interpretation. The first draft of the manuscript was written by Dr. Da Ros, Dr. Duggento, Dr. Cavallo, and Dr. Maestrini. All authors revised critically previous versions of the manuscript. All authors read and approved the final manuscript.

**Data Availability** Data is available upon reasonable request from the corresponding author.

## Declarations

**Ethics approval** All procedures performed in the study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments. The local ethics committee approved prospectively data collection on patients undergoing

EVT (Registro Sperimentazioni, R.S. 25/18). All the procedures performed were part of the routine care.

**Informed consent** Informed consent was obtained from all individual participants included in the study. The authors affirm that human research participants provided informed consent for the publication of the images in Fig. 3A–F.

**Competing interests** The authors have no relevant financial or non-financial interests to disclose.

## References

- Goyal M, Menon BK, van Zwam WH et al (2016) Endovascular thrombectomy after large-vessel ischaemic stroke: a metaanalysis of individual patient data from five randomised trials. *Lancet* 387(10029):1723–1731. [https://doi.org/10.1016/S0140-6736\(16\)00163-X](https://doi.org/10.1016/S0140-6736(16)00163-X)
- Albers GW, Marks MP, Kemp S et al (2018) Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N Engl J Med* 378(8):708–718. <https://doi.org/10.1056/NEJMoal713973>
- Nogueira RG, Jadhav AP, Haussen DC et al (2017) Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N Engl J Med* 378(1):11–21. <https://doi.org/10.1056/NEJMoal706442>
- Mistry EA, Mistry AM, Nakawah MO et al (2017) Systolic blood pressure within 24 hours after thrombectomy for acute ischemic stroke correlates with outcome. *J Am Heart Assoc* 6(5):e006167. <https://doi.org/10.1161/JAHA.117.006167>
- Puntonet J, Richard ME, Edjlali M et al (2019) Imaging findings after mechanical thrombectomy in acute ischemic stroke. *Stroke* 50(6):1618–1625. <https://doi.org/10.1161/STROKEAHA.118.024754>
- Nicholson P, Cancelliere NM, Bracken J et al (2021) Novel flat-panel cone-beam CT compared to multi-detector CT for assessment of acute ischemic stroke: a prospective study. *Eur J Radiol* 138:109645. <https://doi.org/10.1016/j.ejrad.2021.109645>
- Kau T, Hauser M, Obmann SM et al (2014) Flat detector angio-CT following intra-arterial therapy of acute ischemic stroke: identification of hemorrhage and distinction from contrast accumulation due to blood-brain barrier disruption. *Am J Neuroradiol* 35(9):1759–1764. <https://doi.org/10.3174/ajnr.A4021>
- Payabvash S, Khan AA, Qureshi MH et al (2016) Detection of intraparenchymal hemorrhage after endovascular therapy in patients with acute ischemic stroke using immediate postprocedural flat-panel computed tomography scan. *J Neuroimaging* 26(2):213–218. <https://doi.org/10.1111/jon.12277>
- Leyhe JR, Tsogkas I, Hesse AC et al (2017) Latest generation of flat detector CT as a peri-interventional diagnostic tool: a comparative study with multidetector CT. *J Neurointerv Surg* 9(12):1253–1257. <https://doi.org/10.1136/neurintsurg-2016-012866>
- Chen L, Xu Y, Shen R et al (2021) Flat panel CT scanning is helpful in predicting hemorrhagic transformation in acute ischemic stroke patients undergoing endovascular thrombectomy. *Biomed Res Int* 13(2021):5527101. <https://doi.org/10.1155/2021/5527101>
- Obermeyer Z, Emanuel EJ (2016) Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med* 375(13):1216–1219. <https://doi.org/10.1056/NEJMp1606181>
- Soun JE, Chow DS, Nagamine M et al (2021) Artificial intelligence and acute stroke imaging. *AJNR Am J Neuroradiol* 42(1):2–11. <https://doi.org/10.3174/ajnr.A6883>

13. Zeleňák K, Krajina A, Meyer L et al (2021) How to improve the management of acute ischemic stroke by modern technologies, artificial intelligence, and new treatment methods. *Life (Basel)* 11(6):488. <https://doi.org/10.3390/life11060488>
14. Zaidat OO, Yoo AJ, Khatri P et al (2013) Recommendations on angiographic revascularization grading standards for acute ischemic stroke: a consensus statement. *Stroke* 44(9):2650–2663. <https://doi.org/10.1161/STROKEAHA.113.001972>
15. Jiang X (2013) Patient radiation dose when using XperCT and rotational X-ray. Philips Healthcare. Murarrie, QLD, Australia. <https://doi.org/10.13140/RG.2.2.14118.73281>
16. Wang C, Nguyen G, Toncheva G et al (2014) Evaluation of patient effective dose of neurovascular imaging protocols for C-arm cone-beam CT. *AJR Am J Roentgenol* 202(5):1072–1077. <https://doi.org/10.2214/AJR.13.11001>
17. Maier IL, Leyhe JR, Tsogkas I et al (2018) Diagnosing early ischemic changes with the latest-generation flat detector CT: a comparative study with multidetector CT. *AJNR Am J Neuroradiol* 39(5):881–886. <https://doi.org/10.3174/ajnr.A5595>
18. Fiorelli M, Bastianello S, von Kummer R et al (1999) Hemorrhagic transformation within 36 hours of a cerebral infarct: relationships with early clinical deterioration and 3-month outcome in the European Cooperative Acute Stroke Study I (ECASS I) cohort. *Stroke* 30(11):2280–2284. <https://doi.org/10.1161/01.str.30.11.2280>
19. Yoon W, Jung MY, Jung SH et al (2013) Subarachnoid hemorrhage in a multimodal approach heavily weighted toward mechanical thrombectomy with solitaire stent in acute stroke. *Stroke* 44(2):414–419. <https://doi.org/10.1161/STROKEAHA.112.675546>
20. Pedregosa F, Varoquaux G, Gramfort A et al (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830
21. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444. <https://doi.org/10.1038/nature14539>
22. Dekeyser S, Nikoubashman O, Lutin B et al (2017) Distinction between contrast staining and hemorrhage after endovascular stroke treatment: one CT is not enough. *J Neurointerv Surg* 9(4):394–398. <https://doi.org/10.1136/neurintsurg-2016-012290>
23. Sallustio F, Pampana E, Davoli A et al (2018) Mechanical thrombectomy of acute ischemic stroke with a new intermediate aspiration catheter: preliminary results. *J Neurointerv Surg* 10(10):975–977. <https://doi.org/10.1136/neurintsurg-2017-013679>
24. Ma J, Ding Y, Cheng JCP, Tan Y, Gan VJL, Zhang J (2019) Analyzing the leading causes of traffic fatalities using XGBoost and grid-based analysis: a city management perspective. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2019.2946401>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.