


# Computed tomography angiography-based deep learning method for treatment selection and infarct volume prediction in anterior cerebral circulation large vessel occlusion

Acta Radiologica Open  
10(11) 1–10  
© The Author(s) 2021  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/20584601211060347  
[journals.sagepub.com/home/arr](https://journals.sagepub.com/home/arr)  


Lasse Hokkinen<sup>1</sup> , Teemu Mäkelä<sup>1,2</sup>, Sauli Savolainen<sup>1,2</sup> and Marko Kangasniemi<sup>1</sup>

## Abstract

**Background:** Computed tomography perfusion (CTP) is the mainstay to determine possible eligibility for endovascular thrombectomy (EVT), but there is still a need for alternative methods in patient triage.

**Purpose:** To study the ability of a computed tomography angiography (CTA)-based convolutional neural network (CNN) method in predicting final infarct volume in patients with large vessel occlusion successfully treated with endovascular therapy.

**Materials and Methods:** The accuracy of the CTA source image-based CNN in final infarct volume prediction was evaluated against follow-up CT or MR imaging in 89 patients with anterior circulation ischemic stroke successfully treated with EVT as defined by Thrombolysis in Cerebral Infarction category 2b or 3 using Pearson correlation coefficients and intraclass correlation coefficients. Convolutional neural network performance was also compared to a commercially available CTP-based software (RAPID, iSchemaView).

**Results:** A correlation with final infarct volumes was found for both CNN and CTP-RAPID in patients presenting 6–24 h from symptom onset or last known well, with  $r = 0.67$  ( $p < 0.001$ ) and  $r = 0.82$  ( $p < 0.001$ ), respectively. Correlations with final infarct volumes in the early time window (0–6 h) were  $r = 0.43$  ( $p = 0.002$ ) for the CNN and  $r = 0.58$  ( $p < 0.001$ ) for CTP-RAPID. Compared to CTP-RAPID predictions, CNN estimated eligibility for thrombectomy according to ischemic core size in the late time window with a sensitivity of 0.38 and specificity of 0.89.

**Conclusion:** A CTA-based CNN method had moderate correlation with final infarct volumes in the late time window in patients successfully treated with EVT.

## Keywords

CT angiography, stroke, deep learning, machine learning, convolutional neural network

Received 4 March 2021; Accepted 28 October 2021

## Introduction

Endovascular thrombectomy (EVT) has been the standard of care for some years now for patients with ischemic stroke and large vessel occlusion (LVO) presenting within 6 h of

<sup>1</sup>HUS Medical Imaging Center, Radiology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

<sup>2</sup>Department of Physics, University of Helsinki, Helsinki, Finland

### Corresponding author:

Lasse Hokkinen, HUS Medical Imaging Center, Radiology, University of Helsinki and Helsinki University Hospital, P.O. Box 340 (Haartmaninkatu 4), Helsinki 00290, Finland.

Email: [lasse.hokkinen@hus.fi](mailto:lasse.hokkinen@hus.fi)



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

symptom onset.<sup>1</sup> With large trials showing the efficacy and safety of endovascular therapy up to 24 h after time from last known well,<sup>2</sup> the need for advanced neuroimaging and interpretation of these studies has surged.

The main questions to answer for determining eligibility for thrombectomy are: (1) is there salvageable brain tissue, that is, is there a mismatch between the ischemic core and penumbra and (2) how large is the ischemic core. Currently, both computed tomography perfusion (CTP)-based methods and magnetic resonance imaging (MRI)-based perfusion-weighted imaging with diffusion-weighted imaging (DWI) are the mainstay for determining the core and penumbra in stroke diagnosis and treatment selection. In patients with stroke symptoms, computed tomography angiography (CTA) is routinely acquired for LVO detection and detecting significant carotid artery stenoses.<sup>3</sup>

With the growing number of patients possibly eligible for thrombectomy, the need for patient triage has surged, and there is still a need for alternative triage methods, as current imaging methods do not always provide the information needed or they may not be available in hospitals outside of comprehensive stroke centers.

To this end, several studies have looked at CTA-based deep learning methods in ischemic stroke detection with promising results<sup>4–7</sup> and our previous study suggested that a convolutional neural network (CNN) model could be useful in determining eligibility for thrombolytic therapy.<sup>5</sup> However, CNN performance in final infarct estimation might vary depending on treatment selection and could be different in patients receiving thrombolytic therapy versus patients treated with EVT.

In this study, we set out to investigate whether our CTA-based CNN model can predict the final infarct volume in patients with acute ischemic stroke (AIS) treated successfully with mechanical thrombectomy. The premise was that with successful recanalization, the baseline infarct core determined by the CNN should roughly match the final infarct volume on follow-up imaging. CNN performance was analyzed for patients presenting either in the early (0–6 h) or late (6–24 h) time window and compared to a commercially available CTP-based software (RAPID, iSchemaView) to investigate the ability of this method in predicting final infarct volume.

## Materials and methods

The data that support the findings of this study are available from the corresponding author upon reasonable request. Helsinki University Hospital ethics committee approved this retrospective study and patients' informed consent was waived.

## Study population

We retrospectively studied the clinical and imaging findings of consecutive stroke suspected cases that presented to Helsinki University Hospital between February 2018 and March 2020. The total number of thrombectomies performed in our institution during this time period was 505. Inclusion criteria for this study were: (1) stroke code activated, (2) admission stroke protocol imaging performed using fast CTA acquisition protocol and CTP, (3) anterior circulation LVO at CTA (CCA, ICA, MCA M1, MCA M2), (4) successful EVT, defined as Thrombolysis in cerebral infarction score 2b or 3 and (5) a follow-up NCCT or an MRI study with DWI performed no later than 5 days after the onset of symptoms. Patients with hemorrhagic transformation of infarct resulting in parenchymal hematoma with mass effect were excluded. This exclusion was done so that the hematomas and related edema would not produce errors in infarct volume measurements. Patient characteristics are presented in [Table 1](#).

## Image acquisition and preprocessing

A majority of patients ( $n = 82$ ) were imaged in the acute setting using a Siemens Somatom Definition Edge (Siemens Healthineers, Erlangen, Germany) 128-slice CT scanner. The CTA imaging parameters were tube voltage 120 kVp, reference current time 150 mAs, pitch 1.3, reconstruction kernel I30f, and slice thickness/increment 0.75/0.5 mm. The iodine concentration of the contrast agent was 350 mg/mL with an amount of 50 mL and injection rate of 5 mL/s. The timing of the scan was 12 s after time to peak of the test bolus in the ascending aorta. Five patients were imaged with a 128-slice Siemens Somatom Definition Flash (Siemens Healthineers, Erlangen, Germany) and two patients with a 128-slice GE Revolution EVO (GE Healthcare, Milwaukee, WI). CT was used for follow-up in 78 patients and 11 patients had a follow-up MRI. All follow-up studies were performed in our institution and a majority ( $n = 64$ ) were performed with the same scanner as the CTA (Siemens Somatom Definition Edge). Six follow-up CT studies were performed with a Siemens Somatom Definition Flash and eight with a GE Revolution EVO. Follow-up MRI studies were performed with a Siemens Magnetom Verio 3T ( $n = 5$ ), a Siemens Magnetom Skyra 3T ( $n = 3$ ), and a Siemens Magnetom Avanto 1.5 T ( $n = 3$ ). Almost all follow-up studies ( $n = 87$ ) were performed 24 h after admission. In two patients the follow-up study was performed 2 or 3 days from admission. Images were anonymized and stored on a server running the Extensible Neuroimaging Archive Toolkit, or XNAT, version 1.1.6.<sup>8</sup>

All follow-up studies were evaluated for final infarct volume. A senior neuroradiologist (MK) and a radiologist in training (LH), with over 20 and over 5 years of experience,

**Table 1.** Patient characteristics.

No. of patients	89
Age, mean (SD, range)	67 (13.3, 28-92)
Female, <i>n</i> (%)	46 (52)
NIHSS, median (IQR) <sup>a</sup>	12 (7–17)
Time from symptom onset to CT imaging (min), median (IQR) <sup>b</sup>	121 (71–228)
Time from symptom onset to recanalization (min), median (IQR) <sup>2</sup>	215 (169–348)
Most proximal target occlusion location, <i>n</i> (%)	
CCA	1 (<1)
Proximal ICA	3 (3)
Distal ICA	10 (11)
MCA M1	54 (61)
MCA M2	21 (24)
Intravenous thrombolysis, <i>n</i> (%)	40 (45)

<sup>a</sup>NIHSS was reported for 86 patients.

<sup>b</sup>Exact time from symptom onset was unknown in 32 patients.

SD: standard deviation; IQR: interquartile range; NIHSS: national institutes of health stroke scale; CT: computed tomography; CCA: common carotid artery; ICA: internal carotid artery; MCA: middle cerebral artery; CNN: convolutional neural network; CTP-RAPID: computed tomography perfusion RAPID; CBF: cerebral blood flow.

respectively, segmented the infarcted regions on follow-up CT and diffusion weighted MRI scans in consensus using 3D Slicer image processing and visualization platform.<sup>9</sup> No blinding was used regarding image assessment. Image data preprocessing and 3D convolutional neural network implementation was conducted by a physicist (TM).

The CNN had been previously trained and validated and has been used in a recent publication.<sup>5</sup> A total of 150 patients with a suspected AIS of the middle cerebral artery territory were retrospectively selected for CNN development. From this population, 75 were diagnosed with stroke and 75 were stroke mimics based on acute neurological symptoms and imaging findings. The CNN was trained on 20 non-stroke and 20 stroke patient CTA-SI volumes with manually delineated lesion targets. This was equal to 1400 axial images of which 20% included an ischemic lesion. Additional five non-stroke and five stroke volumes were used as validation data. Two test data sets were used, both consisting of 25 stroke and 25 stroke mimic cases. The volumes in both training and inference were resampled to isotropic 0.5×0.5×0.5 mm<sup>3</sup> resolution. Cases used in CNN training were not part of the study population for this study. The CNN consisted of a two-channel input, 40 3D convolutional layers with 3×3×3 kernel size, 16 filters each and valid padding, followed by a fully connected layer with 50 neurons and two output neurons (lesion/background) and softmax activation, producing voxel-by-voxel lesion presence confidences. Skip connections passing single layers were used to encourage gradient propagation. The network was fed with 147×147×147-voxel sub-volumes with equal number of stroke lesion positive and negative sub-volumes in each batch. The second input channel was the corresponding (left-right-mirrored) sub-volume from the contra-

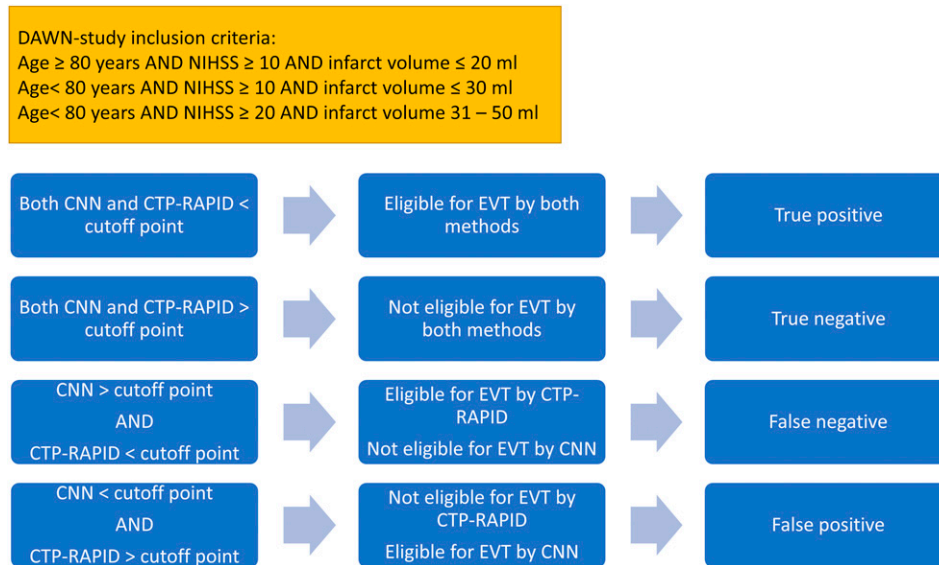
lateral hemisphere. The model was trained using batch-size eight and Adam optimizer for 30 epochs after which the validation loss, calculated on the separate set of 10 CTA-SI volumes, stopped improving. The network was implemented using Keras library version 2.2.4<sup>10</sup> and Tensorflow version 1.12.0.<sup>11</sup>

### Study design

Lesion volumes from CNN outputs and manual segmentations of final infarcts were calculated from all lesions in the affected cerebral hemisphere. Computed tomography perfusion-RAPID ischemic core estimations were reported as calculated by the software.

Only lesions in the affected cerebral hemisphere detected by the CNN were selected for the analysis with a volume threshold of >0.1 mL and a probability threshold of 0.5 for lesion inclusion. False positive lesions in the contralateral hemisphere or cerebellum were discarded from the analysis. This approach was chosen because in LVO, the site of arterial occlusion, and thus the affected hemisphere is readily identifiable from CTA.

Convolutional neural network performance was compared against a commercial software (RAPID, iSchema-View) in determining the infarct core volume derived from CTP as this is a validated and widely used method for treatment selection. The effect of two clinically relevant time windows (0–6 h and 6–24 h from symptom onset to start of CT protocol) on CNN and CTP-RAPID output accuracy in final infarct volume prediction was also tested. Patients whose time of symptom onset could not be accurately inferred from patient history (*n* = 32), were assigned to the 6–24 h time window. These were patients with



**Figure 1.** Accuracy of the convolutional neural network (CNN) in triaging patients for endovascular thrombectomy (EVT) was assessed by defining CNN results as true positives, true negatives, false negatives, or false positives using criteria from the DAWN-study. Sensitivity, specificity, negative, and positive predictive value for the CNN prediction were then derived.

either a wake-up stroke or patients who were found with stroke symptoms and who were last known to be well  $\leq$ 24 h from presentation to the emergency department.

ASPECTS anatomical regions were visually evaluated to determine CNN performance regarding anatomical accuracy against expert segmentation, that is, did the CNN predicted lesions' locations match the final infarct locations within the ASPECTS regions in the middle cerebral artery territory. Individual regions were labeled "positive" or "negative" for ischemic changes by a radiologist (LH), as determined by the CNN from acute phase CTA and by manual segmentations from follow-up CT. Manual segmentations were considered as ground truths. Accuracy, sensitivity, specificity, and Sørensen-Dice similarity coefficient were calculated from the regions' true or false labeling.

Convolutional neural network was compared to CTP-RAPID in determining patient eligibility for EVT according to criteria from the DAWN trial (DWI or CTP Assessment with Clinical Mismatch in the Triage of Wake Up and Late Presenting Strokes Undergoing Neurointervention) in patients that presented in the 6–24 h time window.<sup>2</sup> The cutoff points were ischemic core volumes  $\leq$ 20 mL,  $\leq$ 30 mL, or  $\leq$ 50 mL depending on patient age and National Institutes of Health Stroke Scale score. Volume outputs from the CNN and CTP-RAPID were compared according to Figure 1. The number of true positives, true negatives, false positives, and false negatives was then calculated and sensitivity, specificity, negative, and positive predictive value for the CNN prediction were derived.

### Statistical analysis

A linear model was fitted between the CNN derived volume outputs, manually segmented final infarct volumes, and CTP-RAPID ischemic core volumes (defined by cerebral blood flow (CBF)  $<$  30%). Pearson correlation coefficients ( $r$ ) were calculated to evaluate the correlation of CNN and CTP-RAPID derived volumes against final infarct volumes and CNN derived volumes to CTP-RAPID core volumes. Bland-Altman plots of agreement between infarct volume estimates and final infarct volumes and between CNN and CTP-RAPID derived estimates were also calculated. Calculations were performed using MATLAB version 2018b (MathWorks, Natick, MA, USA). Intraclass correlation coefficient (ICC) estimates and their 95% confidence intervals were calculated using SPSS Statistics for Windows, version 27.0 (IBM Corp., Armonk, NY) based on a single rater/measurement, absolute agreement, two-way mixed-effects model.

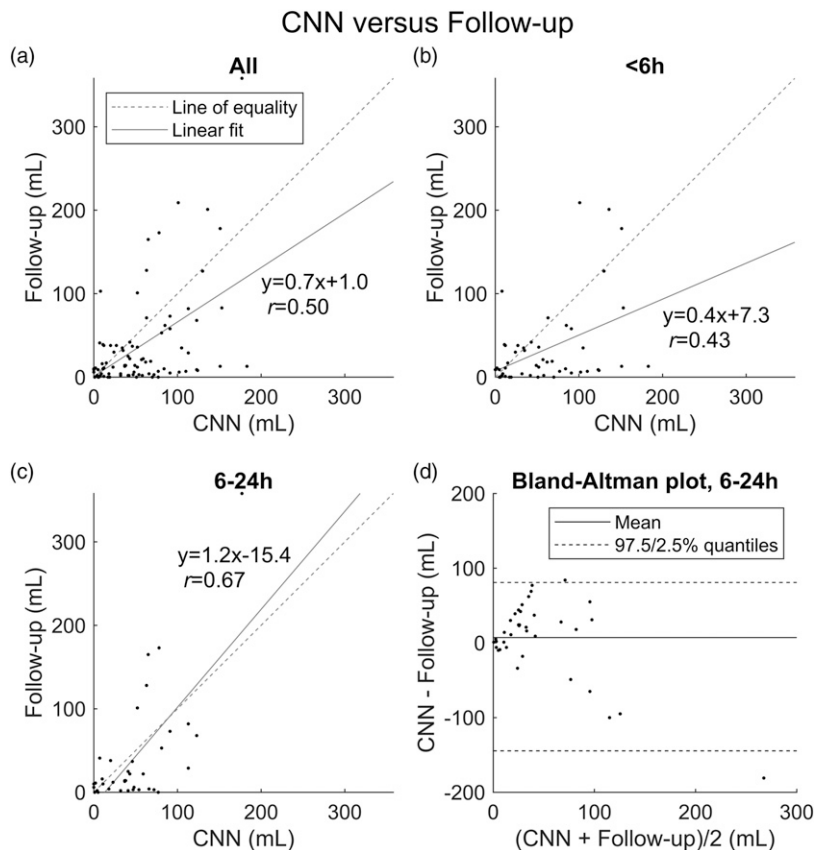
### Results

A total of 117 patients met the inclusion criteria. Patients with hemorrhagic transformation of infarct resulting in parenchymal hematoma with mass effect were excluded, leaving 89 patients for the analysis. Mean infarct volume estimates for CNN and CTP-RAPID outputs as well as final infarct volumes are presented in Table 2. No discernible infarct was found in 10 patients on follow-up imaging.

**Table 2.** Infarct lesion volumes provided by the CNN, CTP-RAPID, and measurements from follow-up imaging in mL, mean (SD, range).

	All cases ( <i>n</i> = 89)	0–6 h time window ( <i>n</i> = 51)	6–24 h time window ( <i>n</i> = 38)
CNN output (mL)	54 (45, 0–183)	58 (49, 0–183)	48 (39, 0–177)
CTP-RAPID infarct core (mL)	28 (36, 0–207)	31 (33, 0–106)	24 (39, 0–207)
Final infarct volume (mL)	36 (58, 0–358)	32 (50, 0–209)	41 (69, 0–358)

CNN: convolutional neural network, CTP-RAPID: Computed tomography perfusion RAPID.

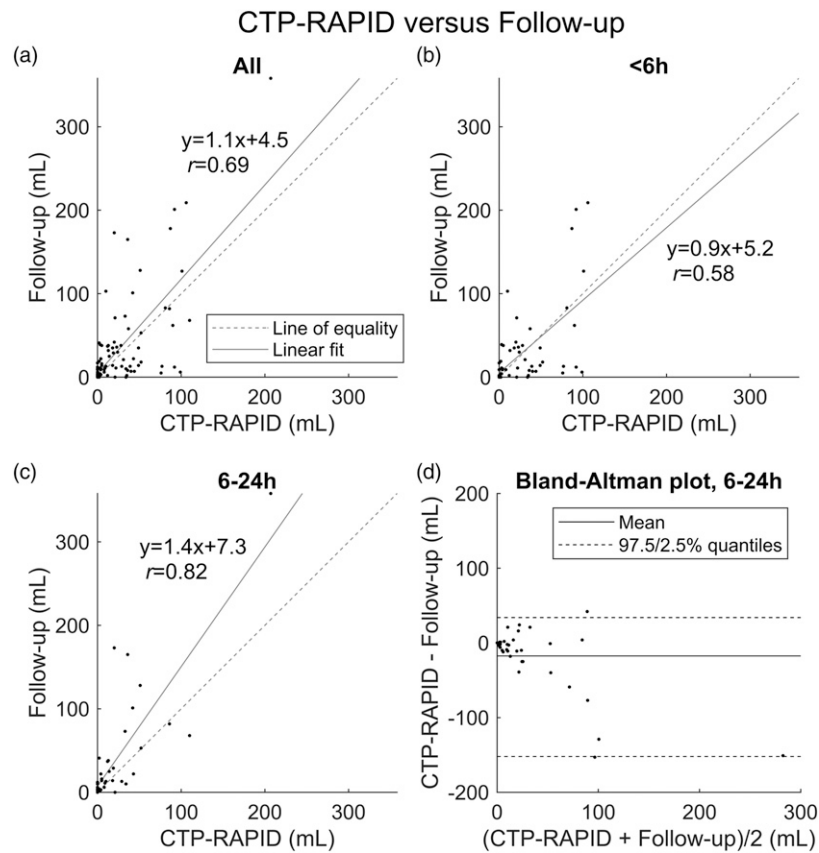


**Figure 2.** Lesion volume (mL) correlation between convolutional neural network (CNN) output and manual segmentation from follow-up imaging. (a) All cases (*n* = 89). (b) Patients imaged <6 h from symptom onset (*n* = 51). (c) Patients imaged 6–24 h from symptom onset (*n* = 38). (d) Bland-Altman plot of agreement between lesion volume estimates of the CNN based on acute phase CT angiography (CTA) and final infarct volumes in the 6–24 h time window.

Among all cases, irrespective of time from symptom onset, CTP-RAPID showed a tendency for more accurate estimates of final infarct volumes ( $r = 0.69$ , slope 1.1,  $p < 0.001$ ), while the CNN had a tendency for overestimation ( $r = 0.50$ , and per linear regression, final infarct volume =  $0.7 \times \text{CNN volume} + 1.0 \text{ mL}$ ,  $p < 0.001$ ) as shown (Figures 2 and 3). The CNN underestimated infarct volumes compared to final infarct volumes in 28% of cases with a mean volume difference of 38.9 mL (SD 44.7 mL). Compared to CTP-RAPID, the CNN underestimated infarct volumes in 11% of cases with a mean volume difference of 13.9 mL (SD 12.5 mL).

Correlations with final infarct volumes in the early 0–6 h time window were  $r = 0.43$  ( $p = 0.002$ ) for the CNN and  $r = 0.58$  ( $p < 0.001$ ) for CTP-RAPID. Computed tomography perfusion-RAPID showed a trend for more accurate volume estimates (slope 0.9 vs. 0.4 for the CNN) as shown (Figures 2 and 3).

In the late 6–24 h time window, a better correlation with final infarct volumes was found for both the CNN ( $r = 0.67$ , slope 1.2,  $p < 0.001$ ) and CTP-RAPID ( $r = 0.82$ , slope 1.4,  $p < 0.001$ ) (Figures 2 and 3). Both methods showed a trend for underestimating infarct volumes. The mean volume difference between CNN output and final infarct volume



**Figure 3.** Lesion volume (mL) correlation between CT perfusion RAPID (CTP-RAPID) ischemic core (CBF <30%) and manual segmentation from follow-up imaging. (a) All cases ( $n = 89$ ). (b) Patients imaged <6 h from symptom onset ( $n = 51$ ). (c) Patients imaged 6–24 h from symptom onset ( $n = 38$ ). (d) Bland-Altman plot of agreement between lesion volume estimates of CTP-RAPID and final infarct volumes in the 6–24 h time window.

**Table 3.** Reliability of the convolutional neural network (CNN) and CT perfusion RAPID (CTP-RAPID) in predicting final infarct volume, intraclass correlation coefficients and their 95% confidence intervals.

	Intraclass correlation	95% confidence interval	<i>p</i> -value
All cases			
CNN Output versus final infarct volume	0.46	0.28–0.61	<0.001
CTP-RAPID versus final infarct volume	0.61	0.47–0.73	<0.001
0–6 h time window ( $n = 51$ )			
CNN Output versus final infarct volume	0.38	0.11–0.59	<0.001
CTP-RAPID versus final infarct volume	0.54	0.32–0.71	<0.001
6–24 h time window ( $n = 38$ )			
CNN Output versus final infarct volume	0.58	0.32–0.76	<0.001
CTP-RAPID versus final infarct volume	0.67	0.44–0.82	<0.001

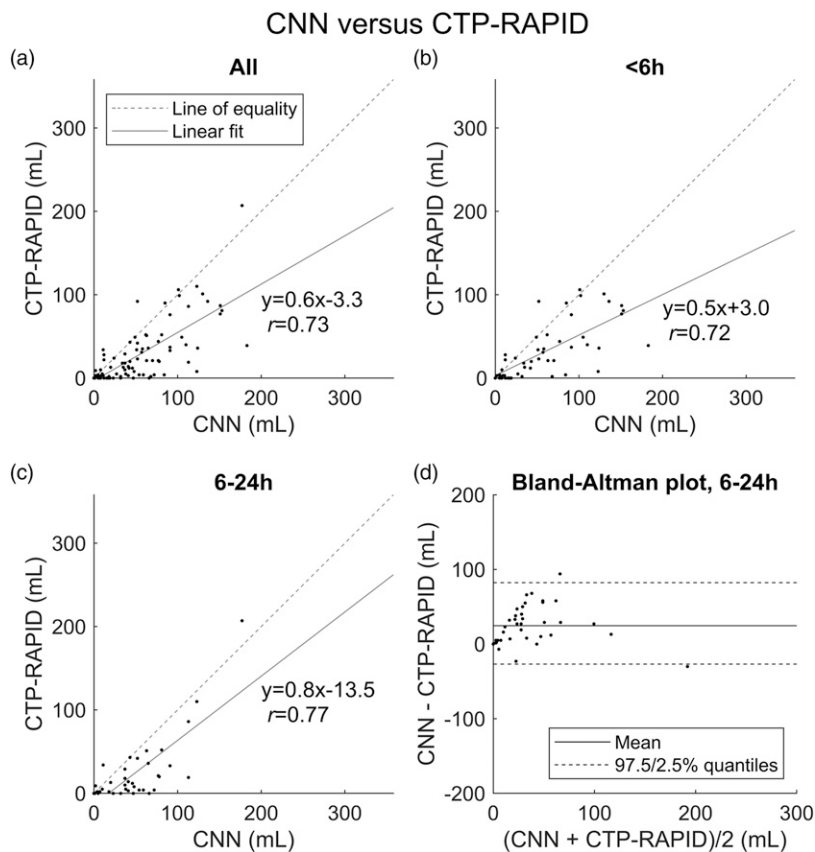
CNN: convolutional neural network, CTP-RAPID: Computed tomography perfusion RAPID.

was 7 mL (95% limits of agreement  $-145$ – $81$  mL). The mean volume difference between CTP-RAPID and final infarct volume was  $-18$  mL (95% limits of agreement  $-152$ – $34$  mL).

Intraclass correlation coefficients for the comparisons above are presented in Table 3 suggesting moderate

reliability between CNN output and final infarct volume in the late time window.

A correlation of  $r = 0.73$  (slope 0.6,  $p < 0.001$ ) was found between CNN and CTP-RAPID outputs when all patients, irrespective of time of symptom onset, were included in the analysis, with the CNN having a tendency to overestimate



**Figure 4.** Lesion volume (mL) correlation between the convolutional neural network (CNN) output and CT perfusion RAPID (CTP-RAPID) ischemic core (CBF <30%). (a) All cases ( $n = 89$ ). (b) Patients imaged <6 h from symptom onset ( $n = 51$ ). (c) Patients imaged 6–24 h from symptom onset ( $n = 38$ ). (d) Bland-Altman plot of agreement between lesion volume estimates of the CNN and CTP-RAPID in the 6–24 h time window.

infarct sizes compared to CTP-RAPID (Figure 4). In subgroup analyses using the early and late time windows, correlations between CNN and CTP-RAPID were  $r = 0.72$  (slope 0.5,  $p < 0.001$ ) and  $r = 0.77$  (slope 0.8,  $p < 0.001$ ), respectively. In the late time window, the mean volume difference between CNN and CTP-RAPID was 25 mL (95% limits of agreement  $-27$ – $82$  mL).

Convolutional neural network accuracy in triaging patients for EVT in the late time window was compared to CTP-RAPID and assessed according to criteria from the DAWN trial as described in the statistical analysis section. Compared to CTP-RAPID, the CNN had a sensitivity of 0.38 with a specificity of 0.89, negative predictive value of 0.31, and positive predictive value of 0.92.

Among all patients, 11 had a chronic cerebral infarct on the same side as the LVO. In three of these cases, the CNN marked parts of the chronic infarct as an acute ischemic lesion. However, this changed the volume prediction only by 0.2–3 mL. A total of 890 ASPECTS regions were evaluated to determine the anatomical accuracy of the CNN. The accuracy, sensitivity, and specificity were 0.62, 0.70,

and 0.57, respectively. The Sørensen-Dice similarity coefficient was 0.60.

## Discussion

Better correlation between our CTA-based CNN outputs and CTP-RAPID core estimates with final infarct volumes were found in the late (6–24 h) versus early (0–6 h) time window. This finding supports the notion of Goyal et al. that extensive use of perfusion imaging in the early time window might not be desirable, and that instead, clinical-imaging mismatch (using NCCT) should be considered to be used for penumbra estimation while using CTA for LVO detection.<sup>12</sup> In the same vein, Lopez-Rivera et al. found an increased likelihood for undergoing EVT in centers with lower CTP utilization, which was not associated with worse clinical outcomes or increased hemorrhage, suggesting under-treatment bias with routine CTP.<sup>13</sup> Also, Boned and Martins have described the “ghost infarct core,” which refers to the tendency of CTP to overestimate

infarct core size in an early time window of <3 h from symptom onset.<sup>14,15</sup>

Correlation with realized infarcts was moderate with the CNN and good with CTP-RAPID in the late time window. Previously, we found a good correlation between CNN outputs and final infarct volumes in patients with anterior circulation AIS that were not treated with EVT.<sup>5</sup> In that study, 55% of patients were treated with supportive care and 45% received thrombolytic therapy, which may explain the differences in performance, as a large proportion of patients can be assumed to have suffered from infarct growth and several studies have shown that CTA is more CBF than CBV weighted including some penumbra.<sup>16–18</sup> On the other hand, Bal et al. found CTA-SI ASPECTS better than NCCT at predicting final infarct extent, especially in a very early 0–90 min time window.<sup>19</sup> In another study, Sallustio et al. found CTA-SI ASPECTS to be a better predictor of outcome than NCCT in patients with stroke treated with EVT.<sup>20</sup> In both of these studies, a fast image acquisition protocol was used for CTA. The conflicting results from these five studies could, at least in part, be explained with different ways of measuring infarct extent as three of the studies used ASPECTS and two used manually segmented lesion volumes to compare performance. These studies also used variable outcome measures for determining CTA performance and inclusion criteria for time from symptom onset to presentation also varied between studies from <3 h to <9 h. As such, findings from these studies may not reflect CTA performance in later time windows.

The observation that in the late time window, CTP-RAPID tended to underestimate core volumes, is in contradiction with numerous previous studies, which have shown a trend for core volume overestimation with CTP.<sup>21–24</sup> These studies, however, used delays of <9 h from symptom onset as inclusion criteria, so these findings might not reflect CTP performance with delays of >9 h. Moreover, large studies have shown the benefit of EVT up to 24 h after symptom onset using CTP-RAPID to guide patient selection.<sup>1,2,25</sup> However, CTP has also received critique in recent history, and it has been questioned, whether it should even be used outside of properly powered clinical trials.<sup>22</sup> Possible problems in using CTP for treatment selection in individual patients include: optimal thresholds varying between vendors and postprocessing platforms, with time after stroke, quality of collateral flow, ischemic preconditioning, and duration of perfusion scans.<sup>12,22,26–29</sup> CTP is also more susceptible to motion artifacts than CTA.

Our CNN method had a positive predictive value of 0.92, that is, if CNN predicted core volume was below the cutoff point, CNN had a probability of 0.92 to correctly classify the patient as eligible for thrombectomy. However, the negative predictive value was only 0.31, which means that a number of patients that could benefit from EVT, would be left outside of treatment if only the CNN method was used.

Our CNN was trained with lesions that were manually segmented from CTA-SI as the ground truth. More accurate final infarct volume prediction may be possible by training the CNN with a different ground truth, such as manually segmented final infarcts from follow-up imaging. This introduces its own challenges, though. For example, there is no one optimal delay for follow-up imaging, as more than 30% relative growth in infarct volume can be witnessed in a significant portion of patients after 24 h<sup>30</sup> and a 3–5 day delay may overestimate infarct size also due to ischemia related edema. Follow-up ischemic lesion volume at 24 h has nevertheless been found to be a valuable secondary outcome measure.<sup>30</sup> The CNN underestimated final infarct volume in 28% of cases. In all of these cases, an underestimation was observed with CTP-RAPID also. This underestimation may be related to infarct progression and edema.

Sheth et al. used CTP-RAPID ischemic core estimates as ground truth for their DeepSymNet algorithm with good results but did not present comparisons to follow-up imaging.<sup>6</sup> The results from their study and our previous study, however, suggest that it may be possible to get reasonably accurate infarct core estimations for triaging purposes using a CTA-based deep learning method. Hilbert et al. have also used CTA-based deep learning models in predicting functional outcome and reperfusion results in AIS.<sup>7</sup> Their approach was quite different in that no lesion segmentations or volumetric data was used for neural network training, but instead, functional outcome and reperfusion measures were used as outcomes and visualization models were used afterward to assess which features or parts of the images the models used for decision making.

In this study, we selected consecutive patients who received EVT to simulate real life performance. This resulted in some limitations to available data, as the exact time from symptom onset to recanalization was unknown for almost half of all patients, albeit they presented  $\leq 24$  h from the time they were last known to be well. This prohibits us from analyzing whether correlation between CNN outputs and final infarct volumes would have been better depending on the delay from symptom onset to recanalization. We used follow-up CT to determine final infarct extent, as a 24 h follow-up CT is the standard protocol in our institution, although it is not as sensitive as DWI. Other limitations of this study include different vendors for CTA imaging and variation in follow-up time and imaging modalities.

In conclusion, a CTA-based CNN is able to detect anterior circulation ischemic strokes with moderate correlation to final infarct volumes in the late time window (6–24 h) in patients successfully treated with EVT.

## Acknowledgments

Eero Salli, PhD (Tec.) and Ulla Wilppu, M.Sc. (Tec.) are acknowledged for their technological assistance and stimulated discussions during the project.



## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study received funding from Helsinki University Hospital (SS: TYH2019253, MK: Y7810046 and Y780021014, LH: Y780020120)

## Ethical approval

Helsinki University Hospital ethical committee approved this retrospective study and patients' informed consent was waived.

## ORCID iD

Lasse Hokkinen  <https://orcid.org/0000-0001-6084-9100>

## References

- Goyal M, Demchuk AM, Menon BK, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. *N Engl J Med* 2015; 372: 1019–1030.
- Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N Engl J Med* 2018; 378: 11–21.
- 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–e418.
- Öman O, Mäkelä T, Salli E, et al. 3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke. *Eur Radiol Exp* 2019; 3: 8.
- Hokkinen L, Mäkelä T, Savolainen S, et al. Evaluation of a CTA-based convolutional neural network for infarct volume prediction in anterior cerebral circulation ischaemic stroke. *Eur Radiol Exp* 2021; 5(25): 25.
- Sheth SA, Lopez-Rivera V, Barman A, et al. Machine learning-enabled automated determination of acute ischemic core from computed tomography angiography. *Stroke* 2019; 50: 3093–3100.
- Hilbert A, Ramos LA, van Os HJA, et al. Data-efficient deep learning of radiological image data for outcome prediction after endovascular treatment of patients with acute ischemic stroke. *Comput Biol Med* 2019; 115: 103516.
- Marcus D, Olsen T, Ramaratnam M, et al. The extensible neuroimaging archive toolkit. *Neuroinformatics* 2007; 5: 11–33.
- Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D slicer as an image computing platform for the quantitative imaging network. *Magn Reson Imaging* 2012; 30: 1323–1341.
- Chollet F, et al. Keras. Available at <https://keras.io> 2015, (accessed 25 January 2021).
- Abadi M, Barham P, Chen J, et al. Tensorflow: a system for large-scale machine learning. In: Proceedings of the 12th USENIX symposium on operating systems design and implementation, Savannah, GA, USA, 2–4 Nov 2016. Usenix Association, 2016, pp. 265–283.
- Goyal M, Menon BK and Derdeyn CP. Perfusion imaging in acute ischemic stroke: let us improve the science before changing clinical practice. *Radiology* 2013; 266: 16–21.
- Lopez-Rivera V, Abdelkhalq R, Yamal JM, et al. Impact of initial imaging protocol on likelihood of endovascular stroke therapy. *Stroke* 2020; 51: 3055–3063.
- Boned S, Padroni M, Rubiera M, et al. Admission CT perfusion may overestimate initial infarct core: the ghost core concept. *J Neurointerv Surg* 2017; 9: 66–69.
- Martins N, Aires A, Mendez B, et al. Ghost infarct core and admission computed tomography perfusion: Redefining the role of neuroimaging in acute ischemic stroke. *Interv Neurol* 2018; 7: 513–521.
- Yoo AJ, Hu R, Hakimelahi R, et al. CT angiography source images acquired with a fast-acquisition protocol overestimate infarct core on diffusion weighted images in acute ischemic stroke. *J Neuroimaging* 2012; 22: 329–335.
- Sharma M, Fox AJ, Symons S, et al. CT angiographic source images: flow- or volume-weighted? *Am J Neuroradiol* 2011; 32: 359–364.
- Mukherjee A, Muthusami P, Mohimen A, et al. Noncontrast computed tomography versus computed tomography angiography source images for predicting final infarct size in anterior circulation acute ischemic stroke: a prospective cohort study. *J Stroke Cerebrovasc Dis* 2017; 26: 339–346.
- Bal S, Bhatia R, Menon BK, et al. Time dependence of reliability of noncontrast computed tomography in comparison to computed tomography angiography source images in acute ischemic stroke. *Int J Stroke* 2015; 10: 55–60.
- Sallustio F, Motta C, Pizzuto S, et al. CT angiography ASPECTS predicts outcome much better than noncontrast CT in patients with stroke treated endovascularly. *Am J Neuroradiol* 2017; 38: 1569–1573.
- Copen WA, Yoo AJ, Rost NS, et al. In patients with suspected acute stroke, CT perfusion-based cerebral blood flow maps cannot substitute for DWI in measuring the ischemic core. *PLoS One* 2017; 12: e0188891.
- Schaefer PW, Souza L, Kamalian S, et al. Limited reliability of computed tomographic perfusion acute infarct volume measurements compared with diffusion-weighted imaging in anterior circulation stroke. *Stroke* 2015; 46: 419–424.
- Austein F, Riedel C, Kerby T, et al. Comparison of perfusion CT software to predict the final infarct volume after thrombectomy. *Stroke* 2016; 47: 2311–2317.
- Tsang ACO, Lenck S, Hilditch C, et al. Automated CT perfusion imaging versus non-contrast CT for ischemic core

- assessment in large vessel occlusion. *Clin Neuroradiol* 2018; 30: 109–114.
25. Albers GW, Marks MP, Kemp S, et al. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N Engl J Med* 2018; 378: 708–718.
  26. Kamalian S, Kamalian S, Maas MB, et al. CT cerebral blood flow maps optimally correlate with admission diffusion-weighted imaging in acute stroke but thresholds vary by postprocessing platform. *Stroke* 2011; 42: 1923–1928.
  27. Kamalian S, Kamalian S, Konstas AA, et al. CT perfusion mean transit time maps optimally distinguish benign oligemia from true “at-risk” ischemic penumbra, but thresholds vary by postprocessing technique. *Am J Neuroradiol* 2012; 33: 545–549.
  28. Copen WA, Deipolyi AR, Schaefer PW, et al. Exposing hidden truncation-related errors in acute stroke perfusion imaging. *Am J Neuroradiol* 2015; 36: 638–645.
  29. Kudo K, Sasaki M, Yamada K, et al. Differences in CT perfusion maps generated by different commercial software: quantitative analysis by using identical source data of acute stroke patients. *Radiology* 2010; 254: 200–209.
  30. Bucker A, Boers AM, Bot JCJ, et al. Association of ischemic lesion volume with functional outcome in patients with acute ischemic stroke: 24-hour versus 1-week imaging. *Stroke* 2017; 48: 1233–1240.