Automated Cerebral Hemorrhage Detection Using RAPID


ABSTRACT

BACKGROUND AND PURPOSE: Intracranial hemorrhage (ICH) is an important event that is diagnosed on head NCCT. Increased NCCT utilization in busy hospitals may limit timely identification of ICH. RAPID ICH is an automated hybrid 2D–3D convolutional neural network application designed to detect ICH that may allow for expedited ICH diagnosis. We determined the accuracy of RAPID ICH for ICH detection and ICH volumetric quantification on NCCT.

MATERIALS AND METHODS: NCCT scans were evaluated for ICH by RAPID ICH. Consensus detection of ICH by 3 neuroradiology experts was used as the criterion standard for RAPID ICH comparison. ICH volume was also automatically determined by RAPID ICH in patients with intraparenchymal or intraventricular hemorrhage and compared with manually segmented ICH volumes by a single neuroradiology expert. ICH detection accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and positive and negative likelihood ratios by RAPID ICH were determined.

RESULTS: We included 308 studies. RAPID ICH correctly identified 151/158 ICH cases and 143/150 ICH-negative cases, which resulted in high sensitivity (0.956, CI: 0.911–0.978), specificity (0.953, CI: 0.907–0.977), positive predictive value (0.956, CI: 0.911–0.978), and negative predictive value (0.953, CI: 0.907–0.977) for ICH detection. The positive likelihood ratio (20.479, CI 9.928–42.245) and negative likelihood ratio (0.046, CI 0.023–0.096) for ICH detection were similarly favorable. RAPID ICH volumetric quantification for intraparenchymal and intraventricular hemorrhages strongly correlated with expert manual segmentation (correlation coefficient r = 0.983); the median absolute error was 3 mL.

CONCLUSIONS: RAPID ICH is highly accurate in the detection of ICH and in the volumetric quantification of intraparenchymal and intraventricular hemorrhages.

ABBREVIATIONS: CNN = convolutional neural network; ICH = intracranial hemorrhage; LR = likelihood ratio; NPV = negative predictive value; PPV = positive predictive value.
increase diagnostic confidence. Deep learning convolutional neural networks (CNNs) represent a technology for automated imaging interpretation that has shown promise in the detection of ICH and other cerebral emergencies. However, this technology has not yet been adopted in widespread clinical practice.

RAPID ICH (iSchemaView) is an artificial intelligence software program developed to identify acute ICH and determine ICH on NCCT studies. RAPID ICH builds on the widely used RAPID software platform for the detection of cerebral ischemia in patients with large-vessel occlusions and is readily adaptable to the RAPID mobile triage platform, which is currently used for triage of patients with ischemic stroke to thrombectomy.

In this study, we tested whether RAPID ICH can accurately detect and volumetrically quantify ICH.

**MATERIALS AND METHODS**

The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Patient Cohort and Image Acquisition**

This retrospective cohort study complied with the Health Insurance Portability and Accountability Act. Institutional review board approval was obtained at each site, and the need for informed consent was waived. NCCTs were obtained from 6 institutions. Only NCCTs that were free of significant motion artifact were included for analysis.

NCCT studies from multiple vendors (Online Table 1) were acquired in the axial plane with section thickness that ranged from 1 to 5 mm. Radiation doses varied by vendor and location, and these variables were not controlled for in this study, which was intended to test whether RAPID ICH can accurately detect and volumetrically quantify ICH.

**Table 1: Patient demographic details**

<table>
<thead>
<tr>
<th>Sex</th>
<th>All</th>
<th>ICH−</th>
<th>ICH+</th>
<th>P Valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, n (%)</td>
<td>117 (38.0)</td>
<td>53 (33.5)</td>
<td>64 (42.7)</td>
<td>&lt;.0002</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>164 (53.3)</td>
<td>81 (51.3)</td>
<td>83 (55.3)</td>
<td></td>
</tr>
<tr>
<td>Unknown, n (%)</td>
<td>27 (8.8)</td>
<td>24 (15.2)</td>
<td>3 (2.0)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2: Performance of RAPID for ICH detection**

<table>
<thead>
<tr>
<th>RAPID vs Consensus Truth</th>
<th>Consensus = Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (positive)</td>
</tr>
<tr>
<td>RAPID 1 (positive)</td>
<td>151</td>
</tr>
<tr>
<td>0 (negative)</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>158</td>
</tr>
</tbody>
</table>

a For a Pearson chi-square test of independence.

**FIG 1.** Representative imaging examples of ICH correctly detected by RAPID. A. Primary intraparenchymal hemorrhage within the right thalamus and posterior limb of the right internal capsule (red outline by RAPID). B. Small extra-axial subdural hematoma overlying the right cerebral hemisphere (red outline by RAPID). C. Subarachnoid hemorrhage (red outline by RAPID). D. Subarachnoid and intraventricular (right and left lateral, third, and fourth ventricles) hemorrhage (red outline by RAPID).
to sample variations in standard radiology practices. Images were not tilt corrected or otherwise manipulated before interpretation by the neuroradiologists in the study. Studies with significant metal or motion artifact were excluded from the analysis.

**RAPID Machine Learning ICH Detection**

A deep CNN with a hybrid 2D–3D architecture was trained on a cohort of 805 NCCT examinations. All head CTs analyzed by RAPID ICH were postprocessed into 5-mm-thick axial slices, and 48 images per study were analyzed by the training dataset. If CT source data were acquired with a thickness of <5 mm, images were merged and averaged into 5-mm thickness before analysis. The training dataset included NCCT with intraparenchymal hemorrhage (245 cases; 30%), intraventricular hemorrhage (83 cases; 10%), extra-axial hemorrhage (70 cases, 9%), subarachnoid hemorrhage (67 cases, 8%), and no hemorrhage (457 cases, 57%).

Regions of ICH were manually outlined by neuroradiology experts and stored as binary masks that were used as ground truth for the training analysis. Ground truth masks were randomly split into training (80%) and testing (20%) groups.

After training, the ICH detection module was prototyped in Python/Keras and implemented on a Linux server (4 CPUs) that was embedded on a larger imaging platform dedicated to ischemic stroke analysis (iSchemaView).

Additional details of the CNN are presented in Online Figure.

**RAPID ICH Validation Study Design**

The RAPID ICH module was validated in a different and independent dataset of 308 adult patients that included 158 patients with ICH (52%) and 150 patients without ICH (48%). The reference standard or truth was determined by consensus among 3 expert neuroradiologists who reviewed NCCT images via Horos (Horos Project, version 3.3) or Osirx (Pixmeo, version 11). Neuroradiologists had full control over window width and levels for all images and were blinded to RAPID ICH results.

The primary end point was the detection of any ICH, which included intraparenchymal, subdural, epidural, subarachnoid, and intraventricular hemorrhages, and ICH presence was recorded in a binary manner (present [ICH+] or absent [ICH−]). Most subdural hemorrhages were acute, but some had subacute or chronic components. Epidural and subdural hemorrhage are collectively referred to as extra-axial hemorrhage.

The secondary end point was ICH volume for isolated intraparenchymal and intraventricular hemorrhages. The reference standard ICH volume was determined by manual segmentation by a single neuroradiologist and verification by a second neuroradiologist. Manual segmentation was performed on axial images. Regions of confluent hyperattenuation that were consistent with intraparenchymal or intraventricular hemorrhage were outlined on all images and volumetrically quantified in Osirix (Pixmeo, version 11).

**Table 3: RAPID ICH performance**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Estimate</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence</td>
<td>0.513</td>
<td>0.457</td>
<td>0.568</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.956</td>
<td>0.911</td>
<td>0.978</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.953</td>
<td>0.907</td>
<td>0.977</td>
</tr>
<tr>
<td>PPV</td>
<td>0.956</td>
<td>0.911</td>
<td>0.978</td>
</tr>
<tr>
<td>NPV</td>
<td>0.953</td>
<td>0.907</td>
<td>0.977</td>
</tr>
<tr>
<td>Positive LR</td>
<td>20.479</td>
<td>9.928</td>
<td>42.245</td>
</tr>
<tr>
<td>Negative LR</td>
<td>0.046</td>
<td>0.023</td>
<td>0.096</td>
</tr>
</tbody>
</table>

**FIG 2.** Representative false-positive and false-negative ICH examples. A. RAPID incorrectly detected ICH (false-positive result) in a patient with volume averaging in the anterior cranial fossa and beam-hardening artifact in the posterior fossa (red outlines by RAPID). B. RAPID failed to detect ICH (false-negative result) in a patient intraparenchymal hemorrhage in the left thalamus (yellow arrow). C. RAPID failed to detect ICH (false-negative result) in a patient with a small amount of subarachnoid and intraventricular hemorrhage (yellow arrows).
Statistical Analysis
Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR), and negative LR were calculated by comparison of RAPID identification of ICH with the reference standard ICH. All statistical analyses were performed using JMP Pro 15 (SAS Institute).

RESULTS
A total of 308 NCCTs were included in the validation study. NCCT was acquired from 5 CT vendors (Online Table 1), distributed across the ICH- and ICH+ studies. After training, the processing time for RAPID’s ICH detection module was <3 minutes per study.

NCCT was derived from 117 female patients (38.0%) and 164 male patients (53.3%), and sex was unknown in 27 patients (8.8%). Most patients were 60 years or older (63.3%), but young patients (ages 20–39 years) were well represented (33 patients, 10.9%). These demographic data are summarized in Table 1.

ICH was detected (ICH+) in 158 NCCT studies (51.2%) by neuroradiology experts, and the distribution of ICH is denoted in Online Table 2. Intraparenchymal hemorrhage was part of most ICH+ NCCT studies (79 cases, 50%), and extra-axial hemorrhage (39 cases, 25%), subarachnoid hemorrhage (25 cases, 16%), intraventricular hemorrhage (10 cases, 6%), and other hemorrhage (5 cases, 3%) were less common.

We compared ICH detection by RAPID with the consensus of 3 expert neuroradiologists (Table 2). RAPID correctly identified 151/158 ICH+ cases and 143/150 ICH– cases (Figs 1 and 2). Therefore, RAPID had a high sensitivity (0.956, CI: 0.911–0.978), specificity (0.953, CI: 0.907–0.977), PPV (0.956, CI: 0.911–0.978), and NPV (0.953, CI: 0.907–0.977) for ICH detection (Table 3). The positive LR (20.479, CI 9.928–42.245) and negative LR (0.046, CI 0.023–0.096) for ICH detection were similarly favorable. NCCT in which RAPID did not detect hemorrhage (false-negative cases) involved cases with small volumes of ICH (<1.5 mL in all instances) that were intraparenchymal, intraventricular, and subdural in location (Fig 2).

We then determined the accuracy of RAPID for ICH volumetric quantification. ICH volume was automatically quantified from the segmentation using Osirix. RAPID quantification of ICH volume demonstrated a strong correlation with the neuroradiology experts (correlation coefficient $r = 0.983$, Fig 3). The mean volumes of ICH in this analysis were 12 mL (RAPID) and 8 mL (expert), and the median absolute error was 3 mL.

DISCUSSION
In this study, we found that the automated artificial intelligence RAPID ICH module is highly accurate for the detection of ICH on NCCT. Moreover, RAPID ICH volumetric quantification of ICH for intraparenchymal and intraventricular hemorrhage was highly accurate. These findings have important implications for the more widespread adoption of artificial intelligence ICH detection into clinical practice.
Other studies have used CNNs to detect the presence of ICH on NCCT. The high sensitivity (96%) and specificity (95%) of RAPID ICH compare favorably with these prior studies that found a sensitivity of 70%–98% and a specificity of 87%–95% for their individual CNN. All 7 false-negative RAPID ICH evaluations consisted of small ICHs that measured <1.5 mL in volume, and 6 of the 7 false-positive cases were attributable to volume averaging. The remaining false-positive study was read as a dural fold by an expert reader. Future iterations of this platform might improve the detection of these small volumes of ICH and the exclusion of false-positive ICHs with artifact reduction approaches. The use of studies from multiple centers and the use of neuroradiology expert interpretations as the criterion standard rather than radiology reports likely increase the generalizability and accuracy of our study compared with prior studies.

Computer-assisted detection of findings on imaging studies has been previously used for cancer detection in mammography, pulmonary nodule detection on chest CT, and colonic polyp detection, but these technologies are not based on CNNs and have limited utility in routine clinical practice. By contrast, CNN technology has been successfully applied to nonradiology studies, such as the detection of diabetic retinopathy and skin cancer detection, which has prompted efforts to move this technology into routine medical practice.

The progressive and substantial increase in the amount of diagnostic imaging studies places particular stress on the timely interpretation of NCCT. Delays in interpretation may lead to delayed identification of ICH, resulting in nonexpeditious patient care that could lead to poor patient outcomes. We speculate that the adoption of CNN technology for ICH detection into routine clinical practice will speed NCCT interpretations, which must still be performed by a board-certified radiologist, and patient treatment decisions. Future studies should be designed to test how this technology changes radiology workflows and patient care.

The use of automated image processing in acute ischemic stroke, which is another cerebrovascular emergency that relies on prompt diagnosis and treatment, is already in widespread clinical practice after several randomized studies demonstrated the effectiveness of this approach in selecting patients for endovascular thrombectomy treatment. We expect that the adoption of CNN technology for ICH detection into routine clinical practice will speed NCCT interpretations, which must still be performed by a board-certified radiologist, and patient treatment decisions. Future studies should be designed to test how this technology changes radiology workflows and patient care.

Our study has several limitations. The retrospective design may introduce bias, and the inclusion of studies from a limited number of CT vendors and locations may limit the generalizability of our findings. Future prospective studies that include a larger number of sites are required for further validation of our findings. It is also possible that beam-hardening artifact, particularly within the posterior fossa, may further limit the sensitivity of RAPID ICH. We also note that although RAPID ICH was highly accurate in this study, automated ICH detection should not preclude interpretation by a trained radiologist because even small undetected hemorrhages might impact outcomes in patients, such as those with ischemic strokes who are being considered for IV thrombolysis.

CONCLUSIONS

RAPID ICH is highly accurate in the detection of ICH and in the volumetric quantification of intraparenchymal and intraventricular hemorrhages. The overall robustness of this CNN approach suggests that automated ICH detection is sufficiently developed for introduction into routine clinical practice.