Can Artificial Intelligence based Automated CT Brain Interpretation Software help Early Clinical Decision Making for Stroke Patients in Real World Resource Limited Settings with Non-Specialist Physicians? An Interrupted Time Series Study from Tezpur, Assam, India

Justy Antony Chiramal

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2023

Committee:

Stephen J. Mooney

Paul K. Drain

Program Authorized to Offer Degree:
Department of Epidemiology

University of Washington

Abstract

Can Artificial Intelligence based Automated CT Brain Interpretation Software help Early Clinical Decision Making for Stroke Patients in Real World Resource Limited Settings with Non-Specialist Physicians? An Interrupted Time Series Study from Tezpur, Assam, India.

Justy Antony Chiramal

Chair of the Supervisory Committee:

Stephen J. Mooney

Department of Epidemiology

We evaluated the impact of an AI (Artificial Intelligence) based software for automated interpretation of CT brain through a retrospective interrupted time series study at a rural hospital in Tezpur, India, that is managed by non-specialist physicians and supported by teleradiology services. We compared the diagnostic accuracy of the software to detect an abnormality against the teleradiologist report and evaluated the impact in stroke patients by comparing the time from CT imaging to significant intervention from before the deployment to the timestamps after the deployment. The specificity and negative predictive value were remarkably high for most of the findings, but the sensitivity and positive predictive value were low for subdural and subarachnoid hemorrhage (n=531). The median time to intervention was significantly lower at 59 minutes in the post deployment phase (IQR: 30.5, 128) than 83 minutes before the deployment (IQR: 57, 144) for acute stroke patients (n=176). Our study showed that it is feasible and impactful to deploy AI based software in resource limited hospitals to help the physicians to make early decisions on life saving interventions in critically ill patients.

TABLE OF CONTENTS

List of Figures

List of Tables

ACKNOWLEDGEMENTS

I am so grateful for this opportunity to work with my mentors and thesis committee members Dr. Steve Mooney and Dr. Paul Drain who are faculty at the University of Washington. They gave me the space to develop my own work, but always guided me when I was in doubt about the research methods or analysis. They never hesitated to have a meeting with me to discuss my questions and helped me develop the best version of the write up.

I would also like to thank the physicians: Dr. Jemin Webster, Dr. Jacob Johnson and Dr. Asolie Chase and the research staff especially Nurse Rachel at the Baptist Christian Hospital, Tezpur who were involved primarily with the software deployment and the data collection. I am also grateful to the Qure.ai team- Dr. Prashant Warier, Dr. Dennis Robert, Saniya Pawar, Pranav, Ayan, and Tejaswini who developed and deployed this impactful solution, so that it could benefit many resource-limited settings across the world. Without their passion, participation and input, the study could not have been successfully conducted.

I would also like to acknowledge Prof. Dr. Steve Hawes, the Chair of the Department of Epidemiology at the School of Public Health at the University of Washington for supporting me pursue my ambition to work with this project that I was involved with in India.

Last but not the least, I want to express my very profound gratitude to my husband (Dominic Jose) and to my daughters (Anaika and Agatha) for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Justy Antony Chiramal

1. Introduction

Stroke is the second leading cause of death across the world and the third leading cause of disability globally (1). Strokes can be either hemorrhagic with an intracranial bleed or ischemic with an interruption of blood supply. A patient can lose up to 32,000 neurons and 230 million synapses every second after the insult (2); time to appropriate treatment sharply affects the disability, morbidity, and mortality in stroke patients. The American Heart Association and American Stroke Association (AHA/ASA) guidelines recommend completing a clinical evaluation in 15 minutes, analyzing brain imaging within 45 minutes, and initiating treatment (if required) within <60 minutes.

However, selecting the appropriate treatment requires distinguishing between hemorrhagic and ischemic stroke. Treatment strategies for ischemic stroke includes thrombolysis (clot dissolving reperfusion therapy) with intravenous tPA (Tissue Plasminogen Activator) like Alteplase within 3-4.5 hours of symptom onset, and further interventions like mechanical thrombectomy (by interventional radiologists), antiplatelet agents, or neurosurgical interventions for decompression if there is raised intracranial pressure. The treatment for hemorrhagic stroke includes administration of mannitol to reduce the intracranial pressure and patients may require further neurosurgical interventions (3). It is therefore vital that physicians have ready access to diagnostic imaging, ideally a non-contrast Computed Tomography (CT) brain scan, and an accurate read of that scan.

These technologies are common in high income settings but are not widespread in low- and middle-income settings. Yet incidence and prevalence of stroke cases has been increasing in LMICs in the past few years (4–6). Worse, lack of resources including radiology and neurology expertise (7,8), has resulted in gaps in healthcare resources and lack of awareness and identification of symptoms of stroke, which delay the diagnosis and treatment (6,9–13). Delay in imaging interpretation in addition to the delay in arrival to the hospital and lack of expertise in management, has led to underutilization of thrombolysis in the developing countries (14,15). There have been multiple strategies that have been implemented to overcome these barriers like teleradiology, tele stroke and telemedicine platforms. Despite these efforts, recent studies have shown dismal access to and utilization of stroke interventions like thrombolysis and mechanical thrombectomy in developing countries (16,17).

AI (Artificial Intelligence) based software have emerged as a potential strategy in reducing the time to intervention in the context of many illnesses like intracranial hemorrhage (18), Traumatic brain injury (19), Tuberculosis (20,21). The goal of this study is to estimate the accuracy and impact of an artificial intelligence (AI) based software for automated interpretation of CT brain imaging on the time to intervention when deployed in a setting with no radiology or neurology expertise. The main objectives of this study were to evaluate the diagnostic accuracy of the software against the radiologist report as ground truth and to assess the impact of software on operational outcomes by comparing the time to intervention before the deployment of software to the time to intervention after the deployment of the software among stroke patients.

2. Methods

2.1 Study Setting

The study was conducted at the BCH (Baptist Christian Hospital), Tezpur, Assam which is a 130 bedded mission hospital under the Emmanuel Health Association (EHA) headquartered in New Delhi, India. The hospital has specialists from internal medicine, surgery, orthopedics, pediatrics and about 5-6 junior medical officers. Patients required interventional radiology or neurology, or neurosurgical treatment were referred to the district hospital. A newly developed trauma center now treats most of the trauma cases who are referred from BCH. During both the phases, the hospital has been dependent for regular reporting of CT scans on teleradiology services provided by the Columbia Asia Hospital in Bangalore in Southern India (3000 km away). During both phases, the number of health care staff in the ER remained the same. Each ER shift of 8 hours was covered by one junior medical officer and one internal medicine physician on duty along with 3 shift nurses and 2 support staff.

There were no neurologists in the hospital, but as a part of physician-based stroke unit model development, they have tele stroke assistance from the Christian Medical College, Ludhiana (22). The standard of care at the hospital was that the physicians would look at the CT images and confirm the findings with the teleradiologist report. If the patient was critically ill and they were not confident of the interpretation from scans, they would call up the teleradiologist to get the report early. Since CT Brain is the only imaging option available in that setting, it was advised for a variety of indications like stroke, head trauma, seizure, loss of consciousness or altered sensorium, severe headache or to check for raised intracranial pressure prior to lumbar puncture. On an average, there are 10-15 patients who undergo CT imaging of the brain on a daily basis.

2.2 Study design

We used an interrupted time series design to compare the time to intervention in acute stroke before and after the deployment of the AI based CT brain software, in addition to the diagnostic accuracy evaluation for all scans. This was the best possible study design that could evaluate both the diagnostic accuracy of the software and the impact of the software. We obtained data through electronic health records and chart reviews for the two phases which spanned for 3 months each.

We obtained Institutional Ethics Committee approval from the Emmanuel Health Association, New Delhi (Protocol No:255 (Version 3)). Additionally, we also obtained approval from the IRB board of the UW (University of Washington) under the IRB ID: STUDY00017314. All the data points were retrospectively collected through a chart review, Electronic Health Records (EHR) and Picture Archiving and communication System (PACS) review. The data was collected by a research staff nurse at the hospital on Microsoft Excel spreadsheets and were shared after removing identifiers for the patients except for the hospital number.

2.3 Study population

Study subjects were consecutive adult patients who underwent CT brain for any indication in BCH during the two phases of study period. The inclusion criteria for the study were that the patients should have undergone a CT scan of the head done for any indication, during the study period and that they should be of age ≥18 years, since the software is FDA approved for >18 years. For the objective of diagnostic accuracy evaluation, we only included patients who had both the radiologist report and the qER report available for comparison. For the objective of operational outcomes, we included only patients who a) were diagnosed with stroke and b) had a significant intervention determined based on CT scan. We excluded patients who had missed significant intervention or if the time of intervention was missing. This could happen in case the patients had come to BCH solely for the CT scan but were under treatment at some other hospital or if the family took the patient home against medical advice immediately after imaging.

2.4 Intervention

qER is an FDA approved and CE certified software which can detect 5 target findings -intracranial hemorrhage (along with 5 subtypes- intraparenchymal, intraventricular, subarachnoid, subdural, and extradural hemorrhage), hypodensities suggestive of infarct, mass effect, midline shift and cranial fracture. It can also quantify volume of hemorrhage and infarct in ml (milliliters) as well as midline shift in mm (millimeters). The software has been demonstrated to have great accuracy when compared to reports of three radiologists as ground truth (23).

The qER software auto pushes the CT images to the cloud server and the qER report is generated and sent back to the system and the Qure.ai app, that in turn notifies and alerts the physicians in about 3-5 minutes of the CT imaging. qER is hardware agnostic since it has been trained on >300,000 CT scans from more than 22 different CT machine models. In addition to offering a pre-populated radiology report, the HIPAA compliant qER solution is also able to label and annotate the abnormalities on the key slices. qER notification contains the details of abnormalities including localization with image of key slices and annotation of abnormalities, including quantification of bleed, midline shift in mm (Figure 1).

2.5 Data Collection

Pre-qER Phase: November 2020 to January 2021

We included all consecutive patients who had undergone CT brain imaging between 1st of November 2020 and 31st of January 2021 for the phase before the deployment of the software qER. There were no radiologists at the hospital in this phase. The standard of care at the hospital was that the physicians would look at the CT images and if they were not confident and if the patient was critically ill, they would call up the radiologist to get the report early or wait for the radiologist report in non-critically ill patients.

Post-qER Phase: September 2022 to November 2022

We included all patients after the deployment of the software qER between 1st of September 2022 and 30th November 2022. In this phase, the clinicians had the additional input from the qER report through the notification system in addition to the standard of care. Though there was a junior radiologist posted at the hospital, they still depended on the teleradiology services for CT reporting. If there was discrepancy in their interpretation of the CT images and the qER report, they would confirm the findings with either the in-house radiologist or the teleradiologist. The notification system would ensure to alert the physicians to look at the scans early, when they are overwhelmed with a large number of patients in the ER for care. Our hypothesis was that these factors would reduce the time from imaging to intervention. The teleradiologists did not have access to the qER report at any point in time. Though qER was deployed in February 2021, auto push of the CT scans was not implemented for more than a year because of the travel restrictions due to pandemic. Thus, our phase 2 data was planned for the latter months in 2022 to analyze the operational outcomes after the deployment of the software with auto push that enabled automatic uploading of the CT images to the cloud server. Each of the phases with the corresponding period and data collection and sources are depicted in Figure 2.

The data collected for all the patients during the study period included demographic details, clinical history including co-morbidities, time stamps and reports from teleradiology for all patients and qER report for patients in the post-qER phase. We also assessed the discharge outcomes in terms of alive, deceased or referred status. All the data was anonymized, i.e., any non-essential personal details (e.g., name, contact details, etc.) were removed. We retained the hospital number to enable tracking the patients for qER reports.

2.6 Data Analysis

After the data was shared, the primary investigator from UW independently evaluated every patient to identify cases of stroke after looking at the symptom profile, time of onset, comorbidities, radiology report, final diagnosis, and interventions. As a quality check the physician from the study site also independently evaluated patients for stroke or other diagnosis in a subset of 50 cases from each phase. We analyzed the data using R (R Foundation for Statistical Computing, Vienna, Austria) and developed plots using the R package ggplot2. Data analysis was done after data cleaning and processing of the data from excel sheets. The baseline descriptive analysis of indications and findings for all CT scans was done for data on co-morbidities, indications for CT and demographic data like age and gender for sub populations for diagnostic accuracy and for the stroke cohort. We used descriptive statistics to summarize continuous variables as medians with interquartile ranges (IQRs) and mean +/- standard deviation (SD) and categorical variables using proportions and numbers.

For the evaluation of diagnostic accuracy, the 5 target findings were compared from the software against the radiologist report as ground truth. We coded for presence or absence for each of the target abnormalities for both qER report and radiologist report. A single patient may have multiple abnormalities (e.g., the same patient could have findings of infarct, hemorrhage, and mass effect in the CT scan). We calculated sensitivity, specificity, positive and negative predictive values with 95% confidence intervals for every target abnormality using the epi.tests function from the epiR package that also generated a 2x2 table for each of the findings. We also presented these estimates for the presence of any target abnormality in the scan overall. Conventionally, diagnostic accuracy should be compared against the diagnosis by a panel of radiologists. Since our study was not budgeted for that, we also decided to compute kappa statistics to estimate the agreement between the radiologist report and the qER report.

For the objective of operational outcomes in stroke patients, we evaluated segments of time from onset of symptoms to entry to hospital, entry to hospital to CT imaging, CT imaging to the teleradiology report, CT imaging to significant intervention and CT imaging to disposition from ER. Our primary goal was to estimate the time to intervention which was calculated from the time of obtaining the CT scan to the time when the significant intervention that was dependent on the CT report would be given. Significant intervention was defined as administration of either thrombolysis or antiplatelet agents like Aspirin in ischemic strokes or could be decompressing medication like mannitol for hemorrhagic strokes. The time from CT acquisition to teleradiologist report was not expected to be affected by the software, since the radiologists did not have access to the qER report. But we wanted to estimate that since that would be equivalent to the standard of care. We also evaluated the time from CT acquisition to the time of disposition of the patient from the ER, to analyze if the software would make a difference in time to admission or referral. Times to treatment were right-skewed, so we computed median and IQR as summary measures. We also divided the timeline in intervals of +/- 30 minutes and compared the distribution of patients by proportions in pre-qER and post-qER phases. We compared the pre-qER timeline to the timeline from the post-qER phase to look for any statistically significant change, at 5% significance level using the Wilcoxon Rank Sum test. For post qER patients, we also evaluated the time from CT acquisition to qER report in terms of median and IQR. This was analyzed in two segments: a) time from CT acquisition to the time the images were acquired by

Qure and b) time from when the images were acquired by Qure to qER report.

3. Results

3.1 Baseline patient characteristics of the study populations

There were 983 adult patients who had CT scans over the period of 6 months from both pre-qER and post-qER phases. Of these, 323 patients were in the pre-qER phase and 660 in the post-qER phase. Thirty-seven percent were female. The mean age was 50.2 and majority of the patients were in the 50-70 age group. For the diagnostic accuracy evaluation, we had a total of 531 patients who had both the radiologist and qER reports. Though there were 660 patients in the post-qER phase, 170 patients were excluded because they had either the radiologist or the qER report that were missing. For the operational outcomes in the stroke cohort, there were a total of 73 patients in pre-qER phase and 103 patients in the post-qER phase who had acute stroke, and who had a significant intervention after CT acquisition. The flowchart in Figure 3 depicts step by step how we arrived at these numbers of patients in the stroke cohort.

The baseline characteristics of the subjects are described in Table 1 for the sub populations that correspond to each of the objectives- one for the diagnostic accuracy and the second for the operational outcomes in stroke patients. Among the diagnostic accuracy cohort, 41% were female. 40% of the patients had hypertension, 12% with diabetes and 10% were consumers of alcohol. The most common indication for CT imaging in the diagnostic accuracy cohort was trauma, closely followed by acute onset weakness. Stroke attributed to about 19% of this cohort. The diagnostic accuracy cohort had 59% of the scans with abnormalities and infarct was the most common finding. 1.5% of the cohort had the outcome of death.

Among the entire study population, there were 176 patients with acute stroke who had a significant intervention after CT acquisition; 57% of the patients had ischemic stroke and the rest had hemorrhagic stroke. 39% of the stroke cohort were females. Young stroke (<50 years age) accounted for 20% of the stroke and only one patient <=30 years of age who had an acute stroke. Among the stroke patients, 73% had hypertension, 20% had diabetes and 10% were consumers of alcohol. Stroke patients also had relatively higher prevalence of smoking, prior history of stroke and IHD (ischemic heart disease). The most common indication for CT imaging for stroke patients was acute onset of weakness, which was among 50% of the patients and that could include hemiparesis (weakness of one half of the body), monoparesis (weakness of one limb) or facial deviation. CT scan was reported as abnormal among 93% of the stroke patients with infarct (both acute and chronic included) as the most common finding. 5% of the patients in the stroke cohort had the outcome of death.

3.2 Diagnostic accuracy of qER

For the post-qER phase, we compared the 5 critical findings that were detected by qER against the gold truth of teleradiologist reports for 531 patients who had both the radiologist and the qER reports available. The overall sensitivity and specificity were 81% and 87% respectively for the presence of any of the target abnormalities. The finding of midline shift had the highest sensitivity and specificity of 94% and 96% respectively. Among the various intracranial hemorrhages, intraparenchymal hemorrhage had a better sensitivity and specificity at 90% and 99.6% respectively. The specificity and the negative predictive value were >90% for all the findings except for infarct and the presence of any one target abnormality. Intraparenchymal hemorrhage had the highest positive predictive value of 96% followed by intracranial hemorrhage with 93%. The kappa statistic values suggest that there is almost perfect agreement between the radiologist and qER for findings of intracranial hemorrhage, intraparenchymal hemorrhage and intraventricular hemorrhage. The kappa statistic was suggestive of substantial agreement between the reports for infarct, midline shift and presence of any one of the target abnormalities. The agreement was the lowest for subarachnoid hemorrhage. The sensitivity and specificity, along with the positive and negative predictive values, and Cohen's kappa statistics for each finding are detailed in Table 2.

3.3 Timeline analysis in Stroke patients

68.3% of the stroke patients presented to the ER between 8 am and 8 pm. The timelines measured in minutes had a wide range with many outliers. The detailed breakdown of the timelines for each of the steps from CT imaging to disposition from ER and the proportion of patients in each of the 30-minute intervals for the pre-qER and post-qER phases are described in Table 3. In the pre-qER phase only one patient received thrombolysis with Alteplase and in the post-qER phase this improved to 7 patients receiving Alteplase.

The median time from CT imaging to teleradiology report was 69 minutes in the pre-qER phase and was only 47 minutes in the post qER phase. Majority of the scans are reported by the teleradiologist from 30-90 minutes of the CT imaging. The time segment of interest was the time from CT imaging to time of significant intervention and there was a significant reduction from 83 minutes (IQR: 57,144) to 59 minutes (IQR: 30.5, 128) with a p-value of 0.0078 by the Wilcoxon rank sum test. The box plot in Figure 4 depicts the difference in the time from imaging to intervention in both the phases, after removing the outliers. The proportion of patients who received significant intervention in <30 minutes had increased from 4.1% in pre-qER phase to 25.2% in the post-qER phase. The density plot distribution in Figure 5 shows the distribution of the time in minutes from CT imaging to significant intervention across both phases. Majority of the patients in the pre-qER phase received their significant intervention within 30-90 minutes, whereas in the post-qER phase most of them received it within the first 60 minutes after CT imaging. The time from CT imaging to disposition from the ER showed an increase when comparing pre-qER and post-qER phases. In the post-qER phase, the median time from CT imaging to the time of qER report was 16 minutes (IQR: 12, 19). This time was further divided into time from CT imaging to acquisition by qER software and the time from acquisition by qER software to qER report. The median time for CT imaging to acquisition by qER software was 12 minutes (IQR: 10,13) and the median time from CT acquisition by qER to the qER report was 4 minutes (IQR: 2.25, 5).

4. Discussion

In this retrospective interrupted time series study, using an AI-based software for automated interpretation of CT brain was feasible and accurate for assisting physicians in diagnosing strokes in a resource-limited setting. On comparing the performance of software against the radiologist diagnosis as ground truth, the specificity and negative predictive values were remarkably high and most of the target abnormalities had high sensitivity and positive predictive value. Among stroke patients, there was a significant reduction in time to significant intervention in the post-qER phase compared to the pre-qER phase.

Anecdotal reports and hospital-based studies indicate a huge burden of young stroke and hemorrhagic type of stroke in Assam. In our stroke cohort there were 20% patients who were <

50 years of age with acute stroke. A recent study showed that hemorrhagic strokes accounted for more than 50% of the cases in Assam, compared to only about 20% of the strokes in the rest of India (24). In our stroke cohort, hemorrhagic strokes were seen in about 43% of the patients, which is closer to the reports. One of the potential factors for higher incidence of hemorrhagic strokes among the Assamese indigenous population and the tea garden workers could be a higher prevalence of hypertension reportedly between 33% to 60.8%, which is the largest single risk factor of stroke (25–27) and our stroke cohort also had a high prevalence of hypertension at 73%.

There have been studies done on the impact of AI based automated interpretation of CT/ MRI/ Perfusion imaging for strokes (18,28), but most of the studies have been done in hospitals in developed countries in settings that had neurology and radiology expertise available. Schmitt etal. In Germany evaluated the accuracy of a CT brain software in Germany to identify intracranial hemorrhage suspect hyper densities in non- contrast CT brain images among stroke patients (28). Seyam etal. conducted a similar study to evaluate the performance of AI software to detect intracranial hemorrhage and compared the turnaround time to report a critical finding in the post intervention phase to the pre-intervention phase (18). This is the first of its kind study done in an LMIC setting with no specialists.

The diagnostic accuracy evaluation showed that the qER software had a great accuracy to diagnose some of the intracranial hemorrhages, mass effect and midline shift with radiologist report as ground truth. The specificity and the negative predictive value of the software were very high to detect most of the target abnormalities. This is very important especially in patients with acute stroke since the high negative predictive value for hemorrhage could be the only factor in ischemic stroke patients for the decision to administer thrombolysis if the patient presents in 4.5 hours of symptom onset or administer antiplatelet agents if presenting later. In ischemic strokes the hyperacute infarct may not be visible in CT scan until many hours after the insult. DWI (Diffusion Weighted Imaging) sequence of MRI brain is the earliest modality to pick up acute infarcts and they are not accessible in these settings. The sensitivity was lower than the recorded performance of the qER software in other settings possibly due to a smaller sample size for each of the target abnormalities. The sensitivity and specificity of qER to detect intracranial hemorrhage were comparable to what was observed in the study by Schmitt et al., where sensitivity and specificity to detect ICH was 0.91 and 0.89 respectively (28). Similar measures were observed in the study by Seyam et al., who recorded the sensitivity and specificity to be 0.87 and 0.94 to detect intracranial hemorrhage respectively (18). They recorded a similarly high negative predictive value which was 0.98 for intracranial hemorrhage. The performance of the qER software in terms of kappa statistics was great for most of the target abnormalities except for subarachnoid and extradural hemorrhage. The kappa statistic of the software in that study was 0.80 which was comparable to the kappa statistic from our study.

In the timeline analysis, though qER was not expected to have an impact on the time from CT imaging to the teleradiologist report we saw a reduction in time by 22 minutes in the post-qER phase than the pre-qER phase. This could be attributed to the workload reduction in the post-qER phase when the healthcare systems were less overwhelmed and the teleradiology services had most of the work force back at work. The most crucial finding from our study was the significant reduction in the time from CT imaging to significant intervention in the post-qER phase when compared to the pre-qER phase. Though there has been a junior radiologist posted at the hospital since the post-qER phase, the standard of care continues where the CT scans are still interpreted by the teleradiologists. The hospital team confirmed that the in-house radiologist is only called to see the CT images if there was a discrepancy between the scan interpretation by the physician and the qER report. Thus, this reduction could be substantially attributed to the software. There was also a significant increase in the proportion of patients who got early intervention after the imaging in the post-qER phase when compared to the pre-qER phase. We expected a subsequent reduction in the time from CT imaging to disposition from ER in the post-qER phase, but in reality, though a decision could be made on patient's admission or referral, there are many factors that could delay the process like lack of availability of in-patient/ ICU bed or delay in ambulance services for referral. But this result serves as a ground reality that there was no escalation in ER services and infrastructure during both the phases though they were many months apart. Seyam et al. reported a reduction in the time to reporting of a critical finding in the post intervention phase by 7 minutes from 70 minutes in the pre-intervention to 63 minutes (95% CI: 55,71) in the post-intervention phase in a setting where radiologists and neurologists were present (18).

In the post-qER phase, the median time from CT imaging to qER report was 16 minutes which was substantially higher than the time recorded in other settings where auto acquisition of the images by the software has been enabled. We found that the main limiting factor at BCH, Tezpur was the low internet speed, thus making the time from imaging to CT acquisition by qER the main bottleneck with a median time of 12 minutes. Once the software acquired the images the report generation was superfast with a median time of 4 minutes, which was in line with our observations from other settings. This could be a factor for improvement to be considered for BCH as well as other hospitals in similar settings to improve the performance and impact of the software.

One of the main strengths of the software is the potential to improve the number of ischemic stroke patients who could be treated by thrombolysis. In the pre-qER phase only one patient out of 73 got thrombolysis and in the post-qER phase there were 7 patients who received thrombolysis. This could not be entirely attributed to the software because Alteplase which is used for thrombolysis is usually expensive and most of the patients cannot afford it, whereas in post-qER phase there was a mobile stroke unit study that was going on that covered the cost of Alteplase for stroke patients. But it is important to note that if the time from imaging to intervention is reduced, there could be more acute stroke patients who could benefit from receiving thrombolysis which could reduce the overall morbidity and mortality with stroke. Though underpowered, our pilot study demonstrated that more patients received significant intervention in <30 minutes from CT imaging in the post-qER phase compared to the pre-qER phase.

One of the main limitations of our study was that the pre-qER and post-qER phases were separated by 19 months, whereas ideally the pre and post timelines should have been separated by a much shorter time period. An added complexity was that the pre-qER phase was during the 1st year of COVID when Indian healthcare systems especially in rural areas were significantly impacted. Because of this there could have been factors related to the pandemic that could affect the pre-qER phase like the number of patients and healthcare workers that could be different between both phases. The hospital team confirmed that the number of healthcare workers in the ER shift remained the same in both phases. The number of CT scans in the post-qER phase were almost double the number of scans in pre-qER phase. This could be because the hospital was overwhelmed with COVID patients and did not have the capacity to admit patients with other emergencies, which was a familiar scene across the world at that time. Another limitation was that diagnostic accuracy is ideally planned against the diagnosis of a panel of radiologists as ground truth rather than the report from a single radiologist. But, for this study we did not have the resources to undertake that, so the sensitivity as expected was lower than what has been recorded earlier (23).

We acknowledge that our study may be underpowered, since the number of patients with scans for both diagnostic accuracy and operational outcomes in stroke may be limited. Despite this limitation we could still observe meaningful outcomes from this pilot project, since there have not been any publications that have evaluated the usefulness of this kind of technology in resource limited healthcare settings in LICs and LMICs. Though we assumed time to intervention could be confounded by factors like time of arrival to the ER, we did not do stratified analysis since we anticipated very few numbers of patients in each stratum. We also analyzed the impact of the software on operational outcomes in patients with acute stroke using the pre and post comparison. Though ideally, clinical outcomes in terms of Glasgow Coma Scale (GCS), National Institute of Health Stroke Scale (NIHSS) and Modified Rankin Scale (mRS) scores could have been analyzed for impact, we decide to focus on more operational than clinical outcomes since there was a lot of missing data due to lack of resources. Also, the guidelines and literature has already shown that the outcomes are better when the time to intervention is shortened, we focused on the impact on time to intervention for our study.

qER is instructed to be used only as an additional triaging tool, while the physician is still expected to interpret the scan and corroborate the findings with the clinical picture and contact the radiologist when in doubt. Thus, it still serves as a great tool in assisting the physician through the alert on their mobile phone which helps bring their attention to the patient's critical condition. The qER report empowers the physicians as an added boost to their confidence in making decisions for interventions. This study demonstrated that it is feasible to deploy a software like this in remote and rural areas with basic internet speed and without the need for expensive hardware and resources. The physicians at BCH validated the usefulness of the software and appreciated the impact it made in their process of decision making. Due to the vagueness of symptoms and lack of accessibility and affordability, most of the stroke patients end up in smaller rural hospitals across the world with lack of specialists. A software like qER could empower the non-specialist physicians to decide to either thrombolyse the patient or to refer the patient to a stroke center at the earliest, which is crucial to reduce the morbidity and mortality associated with stroke.

5. Bibliography

1. Feigin VL, Stark BA, Johnson CO, Roth GA, Bisignano C, Abady GG, et al. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol [Internet]. 2021 Oct;20(10):795–820. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1474442221002520
2. Saver JL. Time is brain - Quantified. Vol. 37, Stroke. 2006. p. 263–6.
3. Powers WJ, Rabinstein AA, Ackerson T, Adeoye OM, Bambakidis NC, Becker K, 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. Vol. 50, Stroke. Lippincott Williams and Wilkins; 2019. p. E344–418.
4. Damasceno A, Gomes J, Azevedo A, al. et. An epidemiological study of stroke hospitalizations in Maputo, Mozambique: a high burden of disease in a resource-poor country. Stroke. 2010;41:2463–9.
5. Connor M, Thorogood M, Casserly B, Dobson C, Warlow C. Prevalence of stroke survivors in rural South Africa: results from the Southern Africa Stroke Prevention Initiative (SASPI) Agincourt field site. Stroke. 2004;35:627–32.
6. Pandian JD, Kalkonde Y, Sebastian IA, Felix C, Urimubenshi G, Bosch J. Stroke systems of care in low-income and middle-income countries: challenges and opportunities. The Lancet [Internet]. 2020 Oct 31 [cited 2023 Mar 26];396(10260):1443–51. Available from: http://www.thelancet.com/article/S014067362031374X/fulltext
7. Frija G, Blažić I, Frush DP, Hierath M, Kawooya M, Donoso-Bach L, et al. How to improve access to medical imaging in low- and middle-income countries ? EClinicalMedicine [Internet]. 2021 Aug 1 [cited 2023 May 4];38:101034. Available from: http://www.thelancet.com/article/S258953702100314X/fulltext
8. Hricak H, Abdel-Wahab M, Atun R, Lette MM, Paez D, Brink JA, et al. Medical imaging and nuclear medicine: a Lancet Oncology Commission. Lancet Oncol [Internet]. 2021 Apr 1 [cited 2023 May 4];22(4):e136–72. Available from: http://www.thelancet.com/article/S1470204520307518/fulltext
9. Menon B, Swaroop J, Deepika H, Conjeevaram J, Munisusmitha K. Poor awareness of stroke—a hospital-based study from South India: an urgent need for awareness programs. J Stroke Cerebrovasc Dis. 2014;23:2091–8.
10. Baatiema L, Aikins A de G, Sav A, Mnatzaganian G, Chan C, Somerset S. Barriers to evidence-based acute stroke care in Ghana: a qualitative study on the perspectives of stroke care professionals. BMJ Open. 2017;7.
11. Kaddumukasa M, Kayima J, Kaddumukasa M, al. et. Knowledge, attitudes and perceptions of stroke: a cross-sectional survey in rural and urban Uganda. BMC Res Notes. 2015;8:819.
12. Jenkins C, Ovbiagele B, Arulogun O, al. et. Knowledge, attitudes and practices related to stroke in Ghana and Nigeria: a SIREN call to action. PLoS One. 2018;13.
13. Kalkonde Y, Deshmukh S, Gokhale C, Jacob M, Bang A. ‘Stroke happens suddenly so it cannot be prevented’: a qualitative study to understand knowledge, attitudes and practices about stroke in rural Gadchiroli, Maharashtra, India. J Neurosci Rural Pract. 2020;11:53–62.
14. Berkowitz A, Mittal M, McLane H, al. et. Worldwide reported use of IV tissue plasminogen activator for acute ischemic stroke. Int J Stroke. 2014;9:349–55.
15. Ogbole G, Owolabi M, Ogun O, Ogunseyinde O, Ogunniyi A. Time of presentation of stroke patients for CT imaging in a Nigerian tertiary hospital. Ann Ib Postgrad Med. 2015;13:23–8.

16. Sarfo FS, Akpa OM, Ovbiagele B, Akpalu A, Wahab K, Obiako R, et al. Patient-level and system-level determinants of stroke fatality across 16 large hospitals in Ghana and Nigeria: a prospective cohort study. Lancet Glob Health. 2023 Apr 1;
17. Asif KS, Fadar ;, Otite O, Desai SM, Herial N, Inoa V, et al. Mechanical Thrombectomy Global Access For Stroke (MT-GLASS): A Mission Thrombectomy (MT-2020 Plus) Study. Circulation [Internet]. 2023 Mar 8 [cited 2023 Mar 9]; Available from: https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.122.063366
18. Seyam M, Weikert T, Sauter A, Brehm A, Psychogios MN, Blackham KA. Utilization of Artificial Intelligence–based Intracranial Hemorrhage Detection on Emergent Noncontrast CT Images in Clinical Workflow. Radiol Artif Intell. 2022 Mar 1;4(2).
19. Hibi A, Jaberipour M, Cusimano MD, Bilbily A, Krishnan RG, Aviv RI, et al. Automated identification and quantification of traumatic brain injury from CT scans: Are we there yet? Vol. 101, Medicine (United States). Lippincott Williams and Wilkins; 2022. p. E31848.
20. Ciecierski-Holmes T, Singh R, Axt M, Brenner S, Barteit S. Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: a systematic scoping review. Vol. 5, npj Digital Medicine. Nature Research; 2022.
21. MacPherson P, Webb EL, Kamchedzera W, Joekes E, Mjoli G, Lalloo DG, et al. Computer-aided X-ray screening for tuberculosis and HIV testing among adults with cough in Malawi (the PROSPECT study): A randomised trial and cost-effectiveness analysis. PLoS Med. 2021 Sep 1;18(9).
22. John L, William A, Dawar D, Khatter H, Singh P, Andrias A, et al. Implementation of a Physician-Based Stroke Unit in a Remote Hospital of North-East India-Tezpur Model. J Neurosci Rural Pract. 2021 Apr 1;12(2):356–61.
23. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. The Lancet. 2018 Dec 1;392(10162):2388–96.
24. Mahanta BN, Mahanta TG, Gogoi P. Clinico-epidemiological profile of stroke patients admitted in a tertiary care Hospital of Assam. Clin Epidemiol Glob Health. 2018 Sep 1;6(3):122–9.
25. Hazarika NC, Biswas D, Narain K, Kalita HC, Mahanta J. Hypertension and its risk factors in tea garden workers of Assam. 2002;
26. Biswas D, Narain K, Mahanta J. Hypertension in the native rural population of Assam. 2003 Nov; Available from: https://www.researchgate.net/publication/7996484
27. Anchala R, Kannuri NK, Pant H, Khan H, Franco OH, Di Angelantonio E, et al. Hypertension in India: A systematic review and meta-analysis of prevalence, awareness, and control of hypertension. Vol. 32, Journal of Hypertension. Lippincott Williams and Wilkins; 2014. p. 1170–7.
28. Schmitt N, Mokli Y, Weyland CS, Gerry S, Herweh C, Ringleb PA, et al. Automated detection and segmentation of intracranial hemorrhage suspect hyperdensities in non-contrast-enhanced CT scans of acute stroke patients. Eur Radiol. 2022 Apr 1;32(4):2246–54.

6. Tables and Figures

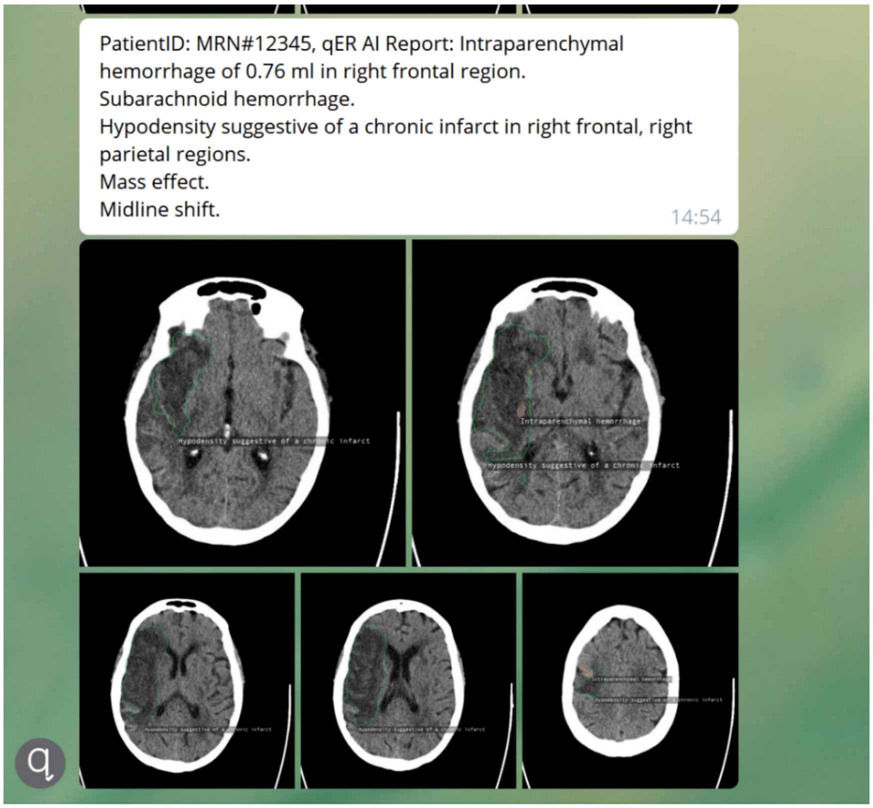

Figure 1. qER notification showing the report with key slices and annotations of abnormalities.

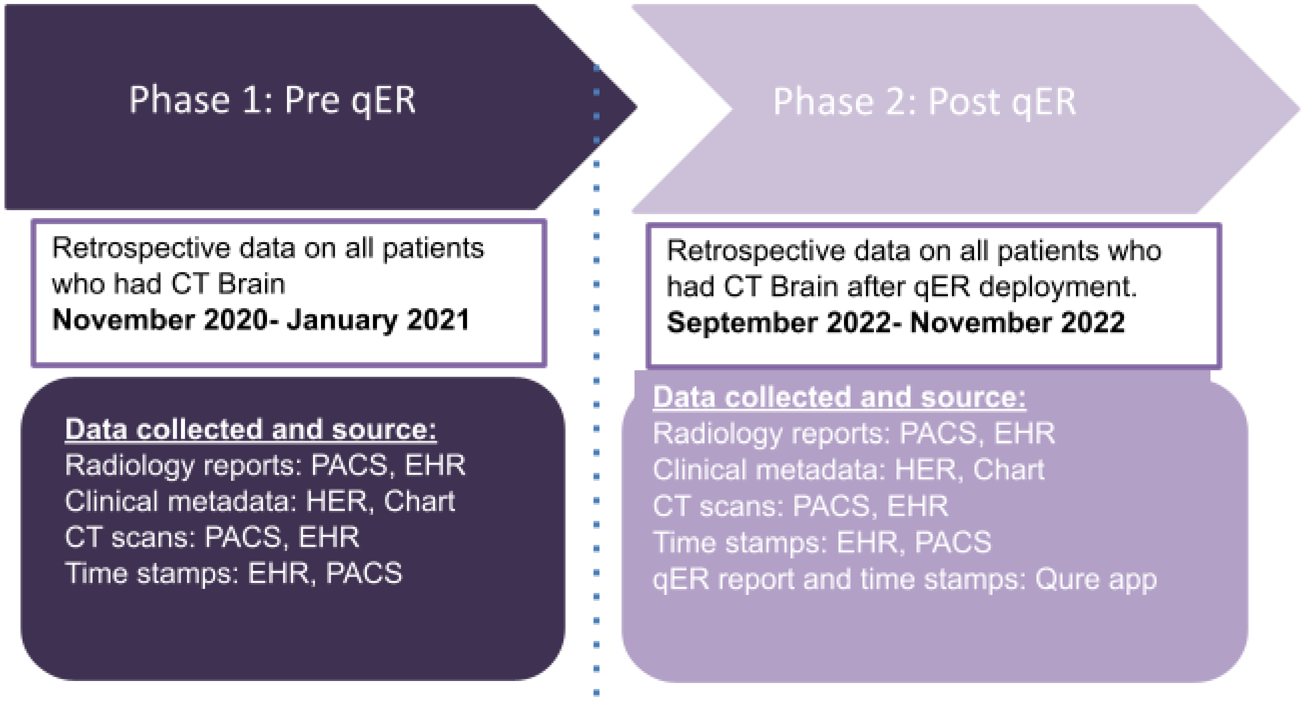

Figure 2. Description of pre-qER and post-qER phases of the study with the data collected and their resources.

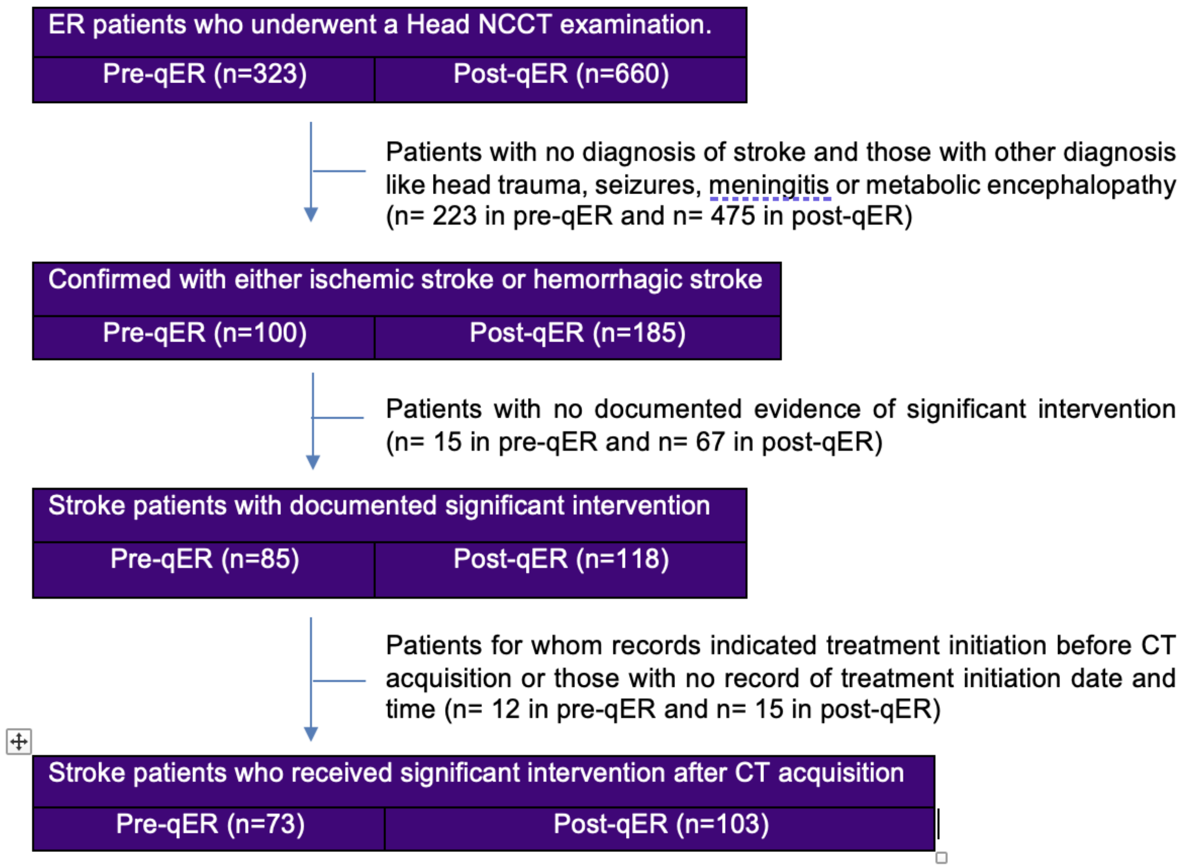

Figure 3. Flowchart depicting step by step how we finalized the number of acute stroke patients for both pre-qER and post-qER phases.

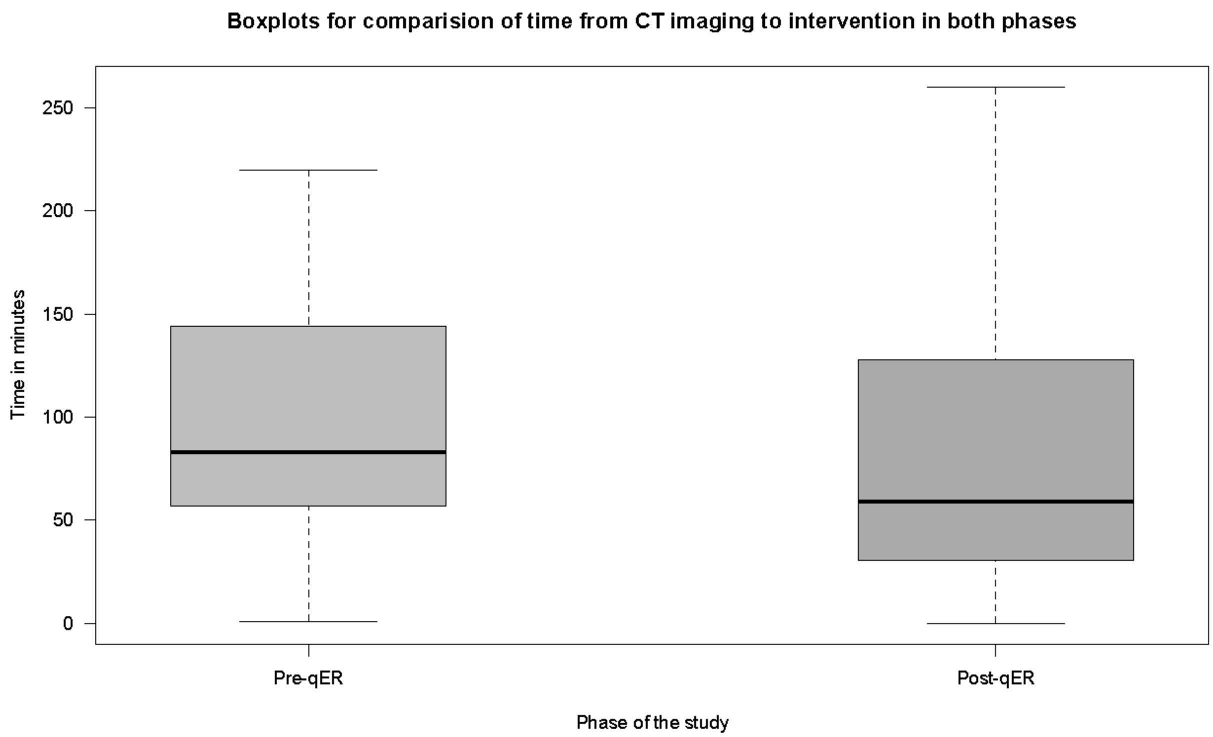

Figure 4. Box Plot showing the distribution of the time in minutes from CT imaging to intervention in both phases.

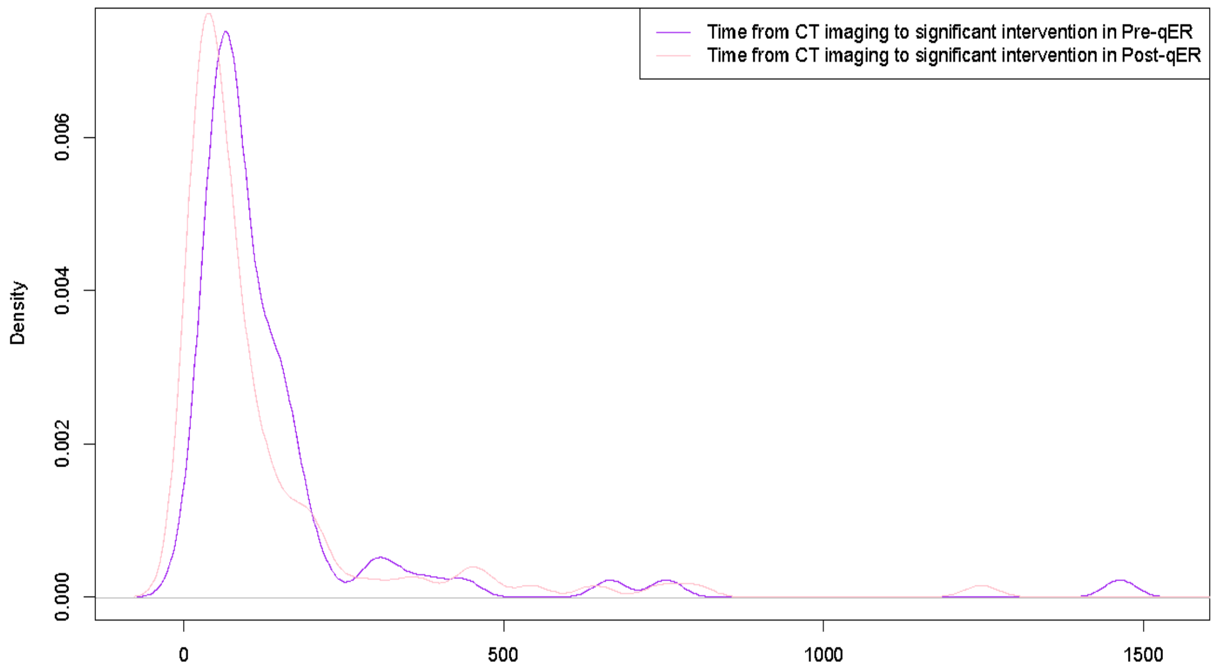

Figure 5. Density plot showing the distribution of time in minutes from CT imaging to significant intervention for both phases.

Table 1. Baseline patient characteristics of the entire study population who had CT brain at BCH Tezpur.

	Number of patients, n (%)	
Characteristic	**Sub population for diagnostic accuracy evaluation (n= 531)**	**Sub-population with diagnosed Stroke (n= 176)**
Phase of the study		
Pre (Phase 1)	NA	73 (41.5)
Post (Phase 2)	531 (100)	103 (58.5)
Age		
<= 30 years	71 (13.4)	1 (0.6)
31-40 years	71 (13.4)	10 (5.7)
41-50 years	83 (15.6)	24 (13.6)
51-60 years	112 (21.1)	43 (24.4)
61-70 years	110 (20.7)	53 (30.1)
>70 years	84 (15.8)	45 (25.6)
Gender		
Female	216 (40.7)	68 (38.6)
Male	315 (59.3)	108 (61.4)
Co-morbidities		
Diabetes mellitus	62 (11.7)	34 (19.3)
Hypertension	212 (39.9)	129 (73.3)
Previous stroke	21(4)	11 (6.3)
Previous IHD	6 (1.1)	5 (2.8)
Smoking	8 (1.5)	7 (4)
Alcohol consumption	54 (10.2)	17 (9.7)
Indication for CT brain		
Acute onset of weakness	103 (19.4)	88 (50)
Slurring of speech	13 (2.4)	17 (9.7)
Trauma	106 (20)	NA
Seizure	34 (6.4)	7 (4)
Loss of consciousness or altered sensorium	59 (11.1)	38 (21.6)
Giddiness	72 (13.6)	16 (9.1)
Headache	53 (10)	8 (4.5)
Abnormalities on CT brain		
Abnormal	314 (59.1)	164 (93.2)
Intracranial hemorrhage	75 (14.1)	76 (43.2)
Infarct	199 (37.5)	90 (51.1)
Cranial fracture	24 (4.5)	0
Mass effect	37 (7)	43 (24.4)
Midline shift	36 (6.8)	33 (18.8)
Type of Stroke		
Ischemic	61 (11.5)	101 (57.4)
Hemorrhagic	42 (7.9)	75 (42.6)

No stroke	428 (80.6)	0
Status at discharge		
Alive	351 (66.1)	164 (93.2)
Died	8 (1.5)	9 (5.1)
Referred	60 (11.3)	3 (1.7)
Missing	112 (21.1)	0

Table 2. Diagnostic accuracy of the findings for qER from the after-intervention phase against radiologist report as gold standard (n=531).

Finding	Number of abnormal scans	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)	Cohen's kappa statistic (95% CI)
Infarct	197	0.73 (0.66, 0.79)	0.89 (0.85, 0.92)	0.79 (0.72, 0.85)	0.85 (0.81, 0.88)	0.63 (0.56, 0.70)
Intracranial hemorrhage	80	0.84 (0.74, 0.91)	0.99 (0.97,1.00)	0.93 (0.85, 0.98)	0.97 (0.95, 0.98)	0.86 (0.80, 0.92)
Intraparenchymal	58	0.90 (0.79, 0.96)	0.996 (0.98,1.00)	0.96 (0.87, 1.00)	0.99 (0.97, 1.00)	0.92 (0.87, 0.97)
Intraventricular	31	0.87 (0.70, 0.96)	0.98 (0.97, 0.99)	0.77 (0.60, 0.90)	0.99 (0.98, 1.00)	0.81 (0.70, 0.91)
Subdural	18	0.22 (0.06, 0.48)	0.96 (0.95, 0.98)	0.18 (0.05, 0.40)	0.97 (0.95, 0.98)	0.58 (0.39, 0.77)
Subarachnoid	9	0.44 (0.14, 0.79)	0.97 (0.95, 0.98)	0.18 (0.05, 0.40)	0.99 (0.98, 1.00)	0.24 (0.036, 0.44)
Extradural	3	0.67 (0.09, 0.99)	0.99 (0.98, 1.00)	0.40 (0.05, 0.85)	0.998 (0.99, 1.00)	0.5 (0.07, 0.92)
Mass effect	37	0.86 (0.71, 0.95)	0.92 (0.89, 0.94)	0.44 (0.32, 0.56)	0.99 (0.97, 1.00)	0.54 (0.42, 0.65)
Midline shift	36	0.94 (0.81, 0.99)	0.96 (0.94, 0.98)	0.63 (0.49, 0.76)	0.996 (0.98, 1.00)	0.73 (0.63, 0.84)
Cranial fracture	24	0.63 (0.41, 0.81)	0.98 (0.96, 0.99)	0.60 (0.39, 0.79)	0.98 (0.97, 0.99)	0.59 (0.43, 0.76)
Presence of any one of the target abnormalities	272	0.81 (0.76, 0.85)	0.87 (0.83, 0.91)	0.87 (0.82, 0.91)	0.81 (0.76, 0.86)	0.68 (0.62, 0.74)

Table 3. Timeline analysis of acute stroke patients at BCH Tezpur.

	Time in minutes, median (IQR), n (%)	
Time segments	**Pre-qER (n=73)**	**Post-qER (n=103)**
CT imaging to Teleradiology report	69 (58,104)	47 (34,75)
< 30 minutes	0	13 (12.6)
30-60 minutes	24 (32.9)	52 (50.5)
60-90 minutes	28 (38.4)	17 (16.5)
90-120 minutes	9 (12.3)	10 (9.7)
120-150 minutes	4 (5.5)	3 (2.9)
150-180 minutes	2 (2.7)	4 (3.9)
>180 minutes	6 (8.2)	4 (3.9)
CT imaging to significant intervention	83 (57,144)	59 (30.5, 128)
< 30 minutes	3 (4.1)	26 (25.2)
30-60 minutes	16 (21.9)	27 (26.2)
60-90 minutes	20 (27.4)	16 (15.5)
90-120 minutes	9 (12.3)	7 (6.8)
120-150 minutes	10 (13.7)	5 (4.9)
150-180 minutes	3 (4.1)	3 (2.9)
>180 minutes	12 (16.4)	19 (18.4)
CT imaging to disposition from ER	136 (106, 214)	159.5 (102, 217)
< 30 minutes	1 (1.4)	1 (1)
30-60 minutes	1 (1.4)	5 (4.9)
60-90 minutes	8 (11)	10 (9.7)
90-120 minutes	14 (19.2)	10 (9.7)
120-150 minutes	14 (19.2)	12 (11.7)
150-180 minutes	10 (13.7)	15 (14.6)
>180 minutes	22 (28.9)	40 (38.8)
CT imaging to qER report	NA	16 (12,19)
0-5 minutes	NA	0
5-15 minutes	NA	37 (35.9)
15-30 minutes	NA	45 (43.7)
30-45 minutes	NA	1 (1)
45-60 minutes	NA	2 (1.9)
>60 minutes	NA	6 (5.8)