Intra and Inter-Observer Reliability of Mobile Tablet PACS Viewer System vs. Standard PACS Viewing Station-Diagnosis of Acute Central Nervous System Events

Aashim Bhatia, Sandip Patel, Gustavo Pantol, Yen-Ying Wu, Michael Plitnikas, Christopher Hancock*
Department of Radiology, Mount Sinai Medical Center, Miami Beach, USA
Email: *chancockmd@gmail.com

Received July 1, 2012; revised February 16, 2013; accepted February 24, 2013

Copyright © 2013 Aashim Bhatia et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background and Purpose: The display resolution of the Apple iPad® is 1024 × 768 pixels, which is greater than that required for generating the typical CT or MRI images. The purpose of this study is to determine if specific CT and MR sequences can be interpreted accurately on mobile device/PACS software platforms when compared to a traditional stationary high resolution monitor/PACS radiological workstation. If so, this allows radiologists to provide comparable interpretation as if they were onsite at an imaging center or hospital. Materials and Methods: This study is an investigator initiated, single site, retrospective, nonrandomized, IRB approved study. Five radiologists were included in this study. Each independently interpreted specific CT and MR sequences on traditional high-resolution LCD monitors via eFilm® software as well as an iPad® mobile device using Osirix® software program. Repeat interpretations were performed, with 4 weeks minimum interval between interpretations of each patient. This investigation included: 50 patients with CTA perfusion imaging, 50 patients with MRI of the brain, and 50 patients with MRI of the spine, which were image study orders generated through emergency room requests. Subsequently, interpretive results of each radiologist for each patient were statistically compared to evaluate for intra-observer and inter-observer reliability. Results: The parameters set within the CTA perfusion brain studies demonstrated excellent intra-observer variability. All of the parameters within the MRI brain studies demonstrated excellent intra-observer variability with a Cohen’s kappa value > 0.75. The Cohen’s kappa values for the board certified neuroradiologist demonstrated excellent variability for all parameters; the resident radiologists had good variability, with a majority of kappa values near 0.75. Conclusions: The data and statistical analysis demonstrated that portable mobile devices such as the Apple iPad® can display adequate resolution of CT and MRI sequences to accurately diagnose acute central nervous system injuries and other non-acute pathology.

Keywords: Teleradiology; Telemedicine; Cloud Network; PACS; Mobile Device; Emergency; Acute Stroke

1. Introduction

Mobile devices with their many applications have made it possible to view radiological images on mobile tablet devices such as the Apple iPad® [1,2]. One of the commonly used applications includes the recently FDA approved mobile application, Mobile MIM® [3]. Mobile tablet/PACS viewer systems, may be used in a cloud environment on Windows® based personal computers and Apple platforms [4]. The first comprehensive mobile medical imaging picture archiving and communication system (PACS) freeware application, Osirix® (Figure 1), can be used on the iPad® and iPhone® [1,2,4,5].

Previous studies have shown the potential for evaluating emergent imaging findings with mobile devices [6-9]. Early studies evaluated personal computers for diagnosing imaging findings before mobile devices were technologically sufficient to view radiological images [10]. Wrist fractures on digital radiographs were evaluated on a personal computer linked 17-inch color monitors, with a display of 1024 × 768 pixels [10]. The traditional workstation measured 21 inches with a display of 1728 × 2304 lines, demonstrating no difference in accuracy among observers [10]. Another more recent study evaluated orthopedic injuries and intracranial hemorrhage on a personal digital assistant and Apple iPod Touch versus non-standard workstation monitors showing the potential
for reading emergent studies on handheld devices [11]. Nonemergent findings such as pulmonary nodules have been compared on the iPad versus the LCD standard display demonstrating no significant difference in diagnostic variability [12]. A study interrogating the utility of analyzing emergent CT brain studies was performed using the Apple iPad as a mobile tablet viewing device with promising results suggesting further evaluation [13]. These studies support the use of the mobile viewer as an accurate, reliable tool for viewing acute and non-acute medical images in various clinical settings.

If specific CT and MR sequences can be interpreted accurately on mobile device/PACS software platforms when compared to a traditional stationary high resolution monitor/PACS workstation, then a radiologist can provide the same turnaround time as if onsite at an imaging center or hospital [14]. This improved interpretation time will result in more timely evaluation and treatment, even when the radiologist is off site [15-17]. This will improve treatment times at a significant number of institutions as many rely on off-site interpretation during non-business hours. A diagnostically accurate mobile medical viewer application is useful for additional primary and secondary interpretations by radiologists, for surgeons who wish to see images in the operating room, for physician discussions with patients and other consulting physicians, for academic presentations, for research, and to facilitate teleradiology/telemedicine endeavors [6,17].

We hypothesize the iPad and other similar mobile device display resolutions are adequate to make critical findings in acute CNS events, in which a delay in diagnosis can prolong necessary intervention and treatment [18]. Recent work supports our hypothesis that mobile tablet and smart phones indeed provide more than adequate resolution required for diagnostic purposes while maintaining expected accuracy [19].

2. Methods

This study is an investigator initiated, single site, retrospective, nonrandomized, IRB approved study. Radiologists included 5 observers consisting of: one American College of Radiology Board Certified Diagnostic & Certificate of Added Qualification-Neuroradiologist, three 4th Year radiology residents (2 entering neuroradiology fellowships), and one 2nd year radiology resident. Each radiologist independently interpreted specific CT and MR sequences on traditional high-resolution LCD monitors via eFilm software as well as on an iPad using Osirix mobile software [1,4,20,21]. The study included anonymized patient images obtained from July 2007 through October 2010. The images consisted of: 50 patients with MRI of the spine, 50 patients with MRI of the brain, and 50 patients with CTA perfusion imaging. The MRI of the spines included cervical, thoracic and lumbar spines. There were patients that had two studies; two different patients had two MRIs of the spine, a lumbar and thoracic spine. One patient had two MRI of the spines including a lumbar and cervical spine. There was a patient, which had two CTA stroke evaluations on successive days included in our study. The CT perfusion study included a plain brain with a brain and bone window; in addition two CT perfusion slabs were included. The MRI of the spine included 3 mm slice thickness sequences: T2 sagittal 3 mm TE 102/TR 4300, T1 sagittal TE 12/TR 500, STIR sagittal TE 49/TR 4360 and T2 axial. MRI of the brain included 5 mm slice thickness sequences: T2 flair axial TE 102/TR 8500, gradient axial TE 26/TR 444, T1 axial TE 10/TR 555, DWI axial with correlating ADC axial maps. Orders were generated through the emergency department for an acute CNS event requiring acute imaging. Each observer reviewed the studies independently on both platforms. Repeat interpretations were performed, with an interval of at least 4 weeks, on each respective software and workstation vs. tablet platform. The blinded interpreting radiologist physicians received access to file folders containing 50 anonymized patient image studies each; organized 1 - 50 for emergent CTA Brain perfusion, 1 - 50 for emergent MRI Brain, and 1 - 50 for emergent MRI Spine. The physicians were specifically blinded to the patient’s medical history simulating how many acute central nervous system events arrive to the emergency room with little, incomplete, or no history. Our institution is a certified comprehensive stroke center. Thus, CTA brain perfusion studies generated through the emergency room are performed to assess for acute ischemic events, and thus this history would be assumed.

Computed tomography angiogram (CTA) brain perfusion study sequences utilized CT brain, CTA maximum intensity projection 30 mm slab reconstructions, as well as colorized and post-processed CTA perfusion blood flow, blood volume, and mean transit time (MTT) renderings [22-25]. Image viewing analysis for parameters of magnification, window panel number, contrast, brightness, and same study comparison were not hindered by software independent default settings, allowing the interpreter free reign to evaluate images he/she would during a standard interpretive session [26-28]. MRI sequences for the brain included DWI, ADC, T2, and gradient. MRI sequences for the spine consisted of T2 sagittal, T1 sagittal, STIR sagittal, and T2 axial. The MRI brain and spine studies in our emergency room may be ordered in any emergency situation deemed clinically appropriate such as trauma, infection, tumor, or acute neurological deficit for common etiologies such infarction, disk herniation and compressed nerve roots [29-33].

The contrast and brightness of the two platforms were
compared. They were found to use different algorithms when applied to the selected series as many proprietary viewers have pre-determined their preferred default settings for rendering image display [26,27]. The Osirix® system locates the middle image of each series, calculates the best W/L for that image, and applies those settings to the entire series [26,34]. The eFilm® image viewer as used on the standard workstation calculates the best W/L for each image, and applies it to that specific image only [26,35]. When comparing images on both systems, no qualitative visual difference was appreciated or reported and there was no fixed default setting. Specifically, interpreting radiologists were encouraged to change window level brightness and contrast variables on both the Osirix and eFilm PACS viewers as would typically be done during interpretation.

We chose the iPad mobile device as our tablet due to its compatibility with the globally popular Osirix® freeware viewer application for the Mac operating system [36]. Osirix® mobile is an application used as a DICOM viewer for the iPad® and iPhone® allowing for medical study image viewing [4,37]. All personal identifiers were removed from the data spreadsheets such that there were no links between the study sequence folders and the individual patients as shown in Figure 1.

Each radiologist documented his results into positive or negative restricted diffusion for infarction via analysis of restricted DWI & ADC mapping, blooming artifact seen on gradient sequencing, perfusion deficit as determined by analysis of cerebral blood flow, cerebral blood volume, and mean transit time [38-43]. In addition, other CNS pathologic etiologies were assayed including: intracranial hemorrhage, spinal cord compression, fracture, spinal canal stenosis, neural foraminal stenoses, disk herniation, and annular fissuring [24,44-47].

Subsequently, the results of each radiologist for each patient were compared to evaluate for intra-observer and inter-observer reliability [48]. The gold standard use for comparison of accuracy of results was based on the findings viewed by the board certified neuroradiologist on the high resolution LCD monitors/eFilm PACS workstation. Results were plotted for the board certified neuroradiologist individually due to wide difference in experience between reviewers, and then for all resident radiologists as a group including the neuroradiologist. The statistical method according to Cohen’s kappa test was employed [49-54]. Cohen’s kappa score range is −1 to 1 [55]. Fleiss’s arbitrary guidelines characterize Kappa’s values > 0.75 as excellent, Kappa’s values = 0.40 to 0.75 as fair to good, and Kappa’s values < 0.40 as poor [55-58].

\[
\text{Kappa} = O - E / 1 - E
\]

\( O \) = Proportion of ratings where scores are in agreement

\( E \) = Proportion of agreements expected “by chance” if scored randomly

The iPad’s display resolution provides an acceptable display to make critical findings in acute CNS events as shown in Table 1.

### 3. Results

Statistical evaluation of intra-observer variation with Cohen’s kappa test demonstrated excellent correlation in a majority of the parameters evaluated, as demonstrated in Table 1.

<table>
<thead>
<tr>
<th>Device/Sequence</th>
<th>Resolution/Pixels Per Inch</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPad</td>
<td>1024 × 768 (132)</td>
</tr>
<tr>
<td>iPhone 4</td>
<td>960 × 640 (326)</td>
</tr>
<tr>
<td>High Resolution Monitor</td>
<td>1200 × 1600</td>
</tr>
<tr>
<td>MRI Brain</td>
<td>512 × 512</td>
</tr>
<tr>
<td>MRI Spine</td>
<td>512 × 512</td>
</tr>
<tr>
<td>CTA Brain</td>
<td>512 × 512</td>
</tr>
</tbody>
</table>

Figure 1. (Osirix® User Interface). Legend: User interface of Osirix® application on the Apple iPad® displaying a list of studies on the left of the screen and preview of series on the right.
with a kappa value > 0.75 in CTA Brain, MRI Brain, and MRI Spine evaluations.

The CTA brain perfusion analysis parameters evaluated included: Intracranial hemorrhage, bright MCA sign, hypo-density sign (<1/3 MCA infarction zone, PCA, ACA, and posterior fossa arterial distributions) and perfusion measurements (cerebral blood flow, cerebral blood volume, and mean transit time). 24 of the 50 CTA perfusion studies demonstrated no findings and were considered negative. Parameters within the CTA brain studies demonstrated excellent intra-observer variability for evaluation of the middle cerebral artery infarction zones as shown in Figure 2. Statistical comparison of the anterior and posterior circulation ischemia was limited to low incidence. For example, the anterior cerebral artery (ACA) territory remote infarct/ischemic change comparison demonstrated a Cohen’s kappa of 0.66 with a small remote ischemic region seen on post analysis, but the statistical evaluation was limited due to the low n=1 obtained in this research study. CT perfusion analysis, as frequently seen at our certified comprehensive stroke center, was considered to be highly reliable for comparison between image viewer system platforms. Notice the conspicuous color delineation seen in the mobile tablet research arm; as is depicted on the iPad® screen shot in a normal case as shown in Figure 3.

MRI brain analysis parameters included: acute, subacute, and chronic features of hemorrhage, restricted diffusion positivity, T2/T2 flair hyperintense lesions, and gradient blooming artifact as shown in Figure 4. All of the parameters within the MRI brain studies demonstrated excellent intra-observer variability with a kappa value > 0.75. Focal infarcts were visualized on the iPad with good intra-observer reliability. Diffusion restriction showing acute infarctions were readily observed, as were vascular thromboses, which were seen conspicuously on gradient sequencing; Figure 5.

MRI spine analysis parameters included: fracture (<50% vertebral body height loss, anterior/mid/posterior column fracture involvement, spinal cord integrity, moderate stenosis, severe stenosis, annular fissuring, and disk herniation; Figure 6. The Cohen’s kappa values for the neuroradiologist demonstrated excellent variability for all parameters >0.75; the resident radiologists had good variability, Cohen’s kappa values between 0.4 and 0.75, and with a majority of Cohen’s kappa values close to 0.75. Spinal cord lesions were seen on the iPad® by all radiologists with excellent reliability. Disk extrusions and annular fissures were easily discernible on both the standard viewing workstation and the mobile viewer platform as shown on iPad® screen shots Figure 7.

4. Discussion

The parameters set within the CTA perfusion brain studies demonstrated excellent intra-observer variability. CTA brain perfusion analysis demonstrated high reliabil-
A. BHATIA ET AL. 95

ity for intra-observer and inter-observer analysis of cerebral blood flow, blood volume, and mean transit time. Statistical evaluation of the MRI brain study research arm demonstrated excellent variability in all parameters for the neuroradiologist and resident radiologists combined. MRI of the spine studies demonstrated the widest variability for the group of resident radiologists; however the neuroradiologist showed excellent variability. This is likely due to relative inexperience of residents and subsequent inconsistency in reading studies. The highest kappa values can be contributed to findings that are more evident to examiners due to relatively conspicuous findings such as greater than 1/3 infarct, DWI positivity, and vertebral body fracture greater than 50%. One limitation of the study is that only one American Board of Radiology board certified with certificate of added qualification (CAQ) Neuroradiologist was included in the comparison analysis. Our institution only employs one such radiologist on the faculty. Future studies may expand the IRB to include multiple centers. HIPAA and image transfer feasibility considerations will need to be addressed at that time. A follow-up study across multiple academic centers would be beneficial to further support our findings.

Another limitation of the study was the range of experience in examiners reading the various studies. However, even with the lack of experience in residents reading complex cases, such as the MRI spine and CT perfusion sequences, the majority of parameters demonstrated excellent intra-observer variability according to Cohen’s kappa score as a key metric of this study.

These results indicate that the iPad and other suitable mobile viewing/PACS platforms can be used by radiologists for diagnostic interpretation. Future studies will focus on application to real world situations, such as on call cases. Items for study include the feasibility for wireless upload of studies to Osirix® (or other software platform), forwarding images to internet mobile software applications via DICOM push/pull algorithms, and subsequent downloading onto mobile viewers for diagnosis. Bandwidth cost and wireless coverage will need to be investigated to determine the practicality of this concept.

Advances continue to be made in the mobile interpretation arena. The MIM® viewer system was recently approved by the Food and Drug Administration (FDA) as a mobile application suitable for interpretation. Other mobile PACS/viewer applications are also under development and FDA approval.

Future studies may be employed to corroborate if similar variability is obtained using other standard workstation monitor systems and divergent PACS software.

5. Conclusion

The data and statistical analysis from our study demonstrate that the Apple iPad®/Osirix® PACS mobile view-
er platform displays more than adequate resolution for typical standard CT and MRI images. This validation allows the interpreting radiologist to confidently interpret and accurately diagnose acute central nervous system injuries and other pathology in a remote setting to determine if acute, subacute, and or non-acute management is necessary. This research supports referring physicians who wish to view critical medical images themselves, surgeons and other interventionalists who need to view images in their operating suites that might not otherwise be available, advances teaching and research options, allows for expansion of telemedicine initiatives, and provides physicians the ability to display and discuss medical images with their patients around the world.

6. Acknowledgements

The authors would like to acknowledge Dr. Jennifer Hui, MD for her editing efforts of this manuscript prior to submission and during the review process.

REFERENCES


Abbreviation

CTA = Computed Tomography Angiogram
MRI = Magnetic Resonance Imaging
ACA = Anterior Cerebral Artery
PCA = Posterior Cerebral Artery
CTA = Computed Tomographic Angiogram