Lesion detection performance: comparative analysis of low-dose CT data of the chest on two hybrid imaging systems

Maryam Jessop*, John D Thompson¹, Joanne Coward¹, Audun Sanderud³, José Jorge⁴, Martijn de Groot⁵, Luís Lança⁶, Peter Hogg¹.

¹University of Salford, Manchester, UK; ²Brighton and Sussex University Hospitals NHS Trust, Brighton, UK; ³Oslo and Akershus University College of Applied Sciences, Oslo, Norway; ⁴Haute École de Santé Vaud, HES-SO, Lausanne, Switzerland; ⁵Hanze University of Applied Sciences, Groningen, The Netherlands; ⁶Escola Superior de Tecnologia da Saúde de Lisboa (ESTeSL), Lisbon, Portugal.

*Corresponding author: Maryam Jessop
Chief Nuclear Medicine Technologist
Department of Imaging
Royal Sussex County Hospital
Eastern Road
Brighton
East Sussex BN2 5BE
UK

e-mail: Maryam.Jessop@bsuh.nhs.uk
Phone: +44 1273 696955 #4382 / #3689
Fax: +44 1273 664923
Abstract:

**Purpose:** Incidental findings, discovered in low-dose CT images obtained during hybrid imaging, are an increasing phenomenon with advancing CT technology. Understanding their diagnostic value along with technical limitations is therefore important when reporting images and recommending follow-up, which may result in additional radiation dose from further diagnostic imaging and an increase in patient anxiety. This study assesses lesion detection in CT images obtained during attenuation correction acquisitions on two SPECT/CT systems.

**Methods:** An anthropomorphic chest phantom, containing simulated lesions of varying size and density, was imaged on an Infinia Hawkeye 4 and a Symbia T6 with low-dose CT settings used during attenuation correction acquisitions in myocardial perfusion imaging. Twenty-two readers completed a lesion detection task, assessing 46 images (15 normal, 31 abnormal containing 41 lesions) from each SPECT/CT system. Data was evaluated using a jackknife alternative free-response receiver operating characteristic (JAFROC) analysis.

**Results:** JAFROC analysis showed a significant difference ($p<0.0001$) in lesion detection with figures of merit 0.599 (95% CI 0.568, 0.631) and 0.810 (95% CI 0.781, 0.839) for the Infinia Hawkeye 4 and Symbia T6 respectively. Lesion detection on the Infinia Hawkeye 4 was generally limited to larger, higher density lesions. The Symbia T6 images allowed improved detection rates with mid-sized lesions and some lower density lesions. However, readers struggled to detect small (5mm) lesions on both image sets, irrespective of density.

**Conclusions:** Lesion detection is more reliable in low-dose CT images from the Symbia T6 than those from the Infinia Hawkeye 4. This phantom based study gives an indication of potential lesion detection in the clinical context as shown by two commonly used SPECT/CT systems, which may assist the clinician in determining if further diagnostic imaging is justified.

**Keywords:** JAFROC, observer performance, low-dose CT, CTAC, incidental findings.
INTRODUCTION

This research comes about with the use of hybrid technology (SPECT/CT) and the subsequent phenomenon of incidental findings in low-dose x-ray computed tomography (CT) images obtained when performing CT attenuation correction (CTAC) acquisitions for myocardial perfusion imaging (MPI). Performance of attenuation correction on single photon emission computed tomography (SPECT) data obtained in MPI acquisitions is needed because of artefactual resemblance of perfusion defects, which vary with each patient, and are the result of attenuation of soft tissue in the chest musculature and breasts, and in some instances by the diaphragm (1). Current use of low-dose CT in this context produces image data that is then converted into an attenuation correction map (μmap), and subsequently applied to the SPECT data with the intent of producing more accurate images for interpretation with increased specificity and improved diagnostic performance (2-3).

The existence of the CT information and the extent of its use in the detection of incidental pathology has become cause for controversy, with discussions on ethical issues and concern regarding added radiation dose in resultant follow-up examinations (4-5). Understanding technical limitations of the images that are produced is a requirement of those who are reporting the images, while also influencing the confidence level upon which any recommendations for further examinations is based (6-7). Incidental findings have been referred to as ‘unsought information generated in the seeking of the information one desires’ (8). While such information can be beneficial, it can also be of detriment to the patient (as can a false positive report) by causing increased anxiety, excessive diagnostic intervention and extra cost (4). If reported or not reported, incidental findings or ‘incidentalomas’ can also lead to medicolegal dilemmas (9). Whatever one’s perspective, the reality is that low-dose CT images now have the potential to produce incidental findings, especially with newer hybrid imaging systems containing ‘state of the art’ CT technology (5,10).

Goetze et al. reported results from a study using a Millennium VG Hawkeye SPECT/CT system (GE Healthcare, Waukesha, WI, USA), that produced ‘potentially significant abnormal findings’ in 10.5% of their patients, and advocated that the CTAC images should be routinely assessed for abnormalities (11). Tootell et al. indicated that 8.1%-18% of CTAC images obtained during MPI may demonstrate some abnormality (1.4% possibly significant, 0.3% significant), and also commented that low image quality of the CTAC images made detection of small lesions very difficult (5). However, a multi-vendor study of lesion detection in CTAC images of a range of SPECT/CT systems implied that the clinical detection of incidental findings may be highly dependent on the CT acquisition parameters used for attenuation correction, and thus the type and age of system used to perform the acquisition (fixed parameters or fully diagnostic) (12). The current work aims to contribute to the understanding of incidental findings and their detection in low-dose CT images produced during attenuation correction in MPI.

METHOD

Recommended manufacturers’ acquisition protocols for CTAC in MPI were used to acquire images of an anthropomorphic chest phantom in two SPECT/CT systems. Simulated lesions of varying size and density were placed in clinically relevant positions throughout the phantom to simulate proximity or distance from structures and complexity of surroundings, while involving upper, middle and lower zones of the lung. Appropriate image data was then analysed in a free-response observer performance study.

Image acquisition

CTAC acquisitions were obtained on both the GE Infinia Hawkeye 4 and the Siemens Symbia T6 using an anthropomorphic chest phantom (LUNGMAN Multipurpose Chest Phantom N1, Kyoto Kagaku Company Ltd, Kyoto, Japan) which contained a removable mediastinum and pulmonary vessel structure, and included three sets of simulated tumour lesions of differing sizes and densities. Those used were 5mm, 8mm, 10mm, and 12mm in size and of densities (Hounsfield units) +100HU, -630HU, and -800HU. Positioning of the simulated lesions was achieved using four configurations, resulting in a varied placement of density and size throughout the lung fields. A diagnostic quality CT scan, performed on the Symbia T6 for each set of lesion positions, acted as a lesion reference map for the ‘truth’ in the observer performance study. Standard manufacturer’s CT quality control was performed on the imaging equipment prior to acquisition to ensure performance levels fell within tolerance thus ensuring validity of subsequent image data.

Unlike the GE Infinia Hawkeye 4, the Siemens Symbia T6 offered various reconstruction kernels able to be set within a CTAC acquisition. Three reconstruction kernels were recommended by Siemens Healthcare and these were a very smooth kernel (B08s) which is a dedicated kernel preparing data for attenuation correction, and two standard higher resolution body kernels (B30s and B60s). The B30s is considered a medium smooth standard body kernel, and the B60s is considered a sharp standard body kernel (13). The images reconstructed...
with the B60s kernel were used in this lesion detection study since they provided optimised images for evaluation of the simulated lung fields and lesions. Acquisition settings are seen in table 1, with the display field of view of the Symbia T6 defined to the sides of the phantom, allowing greater spatial resolution in the reconstructed images.

Observer performance study:
22 readers completed an observer performance study under the free-response receiver operating characteristic (FROC) paradigm. All image evaluations were completed using ‘ROCView’ (14). A total of 46 single CT images were evaluated for each SPECT/CT system. These included fifteen normal images and 31 abnormal images containing 41 lesions, of which 8 contained more than one lesion. The images were randomized for each evaluation. Image viewing stations to be used by readers were assessed for compliance with minimum standards set out by the Royal College of Radiologists in IT guidance documents for image viewing screens (15). Room lighting was dimmed and constant. Readers were trained and accessed the ROCView website via unique usernames. Readers were asked not to restart the evaluation unless genuine mistakes or misunderstandings arose when using the software. They were however permitted to stop and start as they wished, resuming their evaluation at their convenience.

Each reader was required to search the images for lesions and localize (mark) them using mouse clicks. A confidence rating was then applied to each region marked using a slider-bar confidence scale. Responses were recorded on a 10-point (1-10) confidence scale. Data was analyzed using freely available jackknife alternative free-response receiver operating characteristic (JAFROC) software (Version 4.2, www.devchakraborty.com) where the JAFROC figure of merit (0) defines the probability that a lesion rating is higher than any rating on normal images (16-17). A difference in lesion detection performance would be considered significant at \( p = 0.05 \) and the F statistic equal or greater than the critical value (\( \alpha=0.05 \) ) (18-20).

RESULTS
Image appearance was examined from both hybrid imaging systems, and statistical data evaluated. Reader performance within the image evaluations was examined, and the effects of CT viewing experience assessed. Finally, lesion detection in terms of size and density was examined. It is also worth noting that dose modulation was used on the Symbia T6 contributing to lower exposure doses (21), with a Dose length Product (DLP)=97mGy*cm for the Infinia Hawkeye 4 compared to DLP=44mGy*cm on the Symbia T6.

Image appearance
The Symbia T6 images are clearer and have greater contrast and spatial resolution. Differences in clarity of the images are apparent in figure 1.

JAFROC analysis
JAFROC analysis employing Dorfman-Berbaum-Metz-Multi Reader Multi-Case (DBM-MRMC) significance testing found a statistically significant difference in lesion detection performance: \( F(1,21) = 224.1 \) (critical value= 4.3248), \( p<0.0001 \). As this was a phantom study, the results could only be classified as ‘fixed case’ therefore the results relate to a ‘Random Readers and Fixed Cases’ analysis. The area under the alternative free-response receiver operating characteristic (AFROC) curves, plotted in figure 2, is equivalent to the JAFROC 0, thus providing the figure of merit (FOM) value (17).

Reader averaged FOM results are presented in tables 2 and 3. The low standard deviation demonstrated consistent performance by readers on both imaging systems. The results in table 3 show the difference between the two imaging systems. Also, on examination of the 95% confidence intervals in table 3, one can see that these do not include zero which shows a statistical significance in the treatment pairing (22).

Readers
The readers had a wide range of CT experience from 0 to 23 years (average 4.7 years), and some had extensive experience in nuclear medicine ranging from 0 to 13 years (average 3 years). Regression analyses demonstrated no relation between the readers’ experience in viewing CT and their figure of merit for both imaging systems. Additional regression analysis looking at CT viewing experience with regards to lesion localization (LL) and non-lesion localizations (NL), and showed no good relations present in these instances.

Regression analyses were then used to determine any relation between lesion localization (LL) and non-lesion Localizations (NL) on both imaging systems. These results showed that in approximately 60% of cases there was consistent evaluation of images from the two imaging systems by the individual readers, when looking separately at lesion localization (LL) and non-lesion localization (NL). How well the individual reader...
detected lesions on one set of images corresponded to how well they detected lesions on the set of images from the other imaging system, in approximately 60% of cases. Any propensity for false positives was also consistent between both imaging systems for individual readers in approximately 60% of cases, but this was not linked to lesion detection.

**Lesion detection**

‘ROCView’ recorded readers’ detection of lesions from randomized images, some of which contained more than one lesion. The number of readers that detected the lesions in each case were totalled and the true positive (TP) confidence ratings in each case were averaged, for both imaging systems. Corresponding lesion position data was examined and whether cases contained multiple lesions.

Overall averages, as related to specific lesion size and density, and True Positive (TP) confidence ratings, are presented in table 4 in absolute numbers. Graphs in figures 3 and 4, which represent this data expressed as a percentage, show that lesion detection on the Symbia T6 was more dependent on size. However on the Infinia Hawkeye 4 it appears that lesion detection, while dependent on size was also dependent on density where the +100HU lesions were detected more effectively (with the exception of the 8mm lesions where the -630HU lesions were detected with greater frequency). 5mm lesions were not detected reliably on either imaging system.

When examining lesion detection on a case by case basis, there were obvious differences in detection of the same lesion in some instances. These differences occurred more commonly in cases with multiple lesions but not exclusively so. On examination of the images, the reasons for these differences could be identified to some degree, and included things such as similarity of the lesion to blood vessels or surrounding structures, partial volume effect, and complexity of lesion surroundings.

**DISCUSSION**

It is important to remember that the original purpose of the low-dose CT, in this instance, is to provide attenuation correction for myocardial perfusion imaging. However, as the CT images are available, evaluation of them should be considered, and some might say required (23-24). While controversy abounds surrounding the extent of reporting low-dose CT acquisitions, the fact remains that incidental findings do occur. There are a number of factors affecting reliability of these findings, including quality of image, lesion size and lesion density. For the purposes of this research, identifying limitations of lesion detectability on the hybrid imaging systems used is an important objective, especially in the clinical context.

**Equipment**

CT rotation time is not relevant to this study, because there is no breathing artefact to be taken into account as the phantom is static. However the advanced technology of the Symbia T6, including UFC detectors enabling more effective utilization of x-ray exposure (25), and dose modulation, has resulted in radiation exposure doses that are half that of the Infinia Hawkeye 4. The larger matrix size used by the Symbia T6, and the adjustable display field of view, facilitates an increase in resolution. The Symbia T6 also has the advantage of multiple reconstruction kernels set at acquisition which enables low-dose CT acquisition data to be easily optimized for both attenuation correction and image viewing. While the Infinia Hawkeye 4 uses reconstruction algorithms optimized for a low-dose CT regime (26), it is unable to achieve image quality produced by the newer technology of the Symbia T6.

**JAFROC analysis and lesion detection**

There is a significant difference (p<0.05) in lesion detection clearly demonstrated between the low-dose CT images produced by the two hybrid imaging systems (p<0.0001). More lesions were detected with more confidence on the Symbia T6 as reflected in the higher FOM seen in table 2. While detection of lesions on the Symbia T6 appears to be more dependent on size, the effect of both size and density on lesion detection on the Infinia Hawkeye 4 is more apparent, as demonstrated in the graph in figure 4. A satisfaction of search effect may have been seen in images with multiple lesions, as was the effect of partial voluming, and complexity of surroundings, despite using images that showed either single or multiple lesions at their maximum visibility.

**Application to clinical context**

When breathing artefact is factored into lesion visibility in the clinical setting, it is understandable that some reporters, in the author’s experience, may seem reluctant to report lesions in chest images from the Infinia Hawkeye 4. Conversely, the clarity of images able to be produced on the Symbia T6 during low-dose CT
advances in CT technology affecting image acquisition and reconstruction appear to be significant. The limits demonstrated in this study for reliable lesion detection on the hybrid imaging systems used, would be useful for those reporting images, both in their understanding of technical limitations and reliability of lesion detection in this context. This may increase confidence of some reporters to attach greater significance to their findings and recommend appropriate follow-up investigations. Conversely, greater understanding of specifics in regard to technical limitations and reliability of lesion detection may result in fewer follow-up investigations being recommended therefore less added radiation dose to the patient. More reliable data concerning this may ease the medical, ethical and legal dilemmas that have arisen. Therefore confidence levels, as used in observer performance, are clinically relevant and are important in image interpretation where characteristics of patient, imaging system, and image reporter intertwine.

CONCLUSION

The aim of this observer performance study was to contribute to the understanding of incidental findings and their detection in low-dose CT images obtained during MPI CT based attenuation correction acquisitions. Evaluation of lesion detection, in this context, was carried out on two commonly used SPECT/CT hybrid imaging systems. Advances in CT technology affecting image acquisition and reconstruction appear to be significant in the detection of simulated lesions.

While only phantom based, the results obtained are indicative of potential lesion detection within the clinical context, which may assist the clinician in determining if further diagnostic imaging is justified. Continuing research into the phenomenon of incidental findings is needed, specifically in determining limitations for the low-dose CT images and hybrid imaging systems from where these findings originate.

CONFLICT OF INTEREST STATEMENT

The authors have nothing to declare.

ACKNOWLEDGEMENTS

This work was supported by the University of Salford. The authors are appreciative of the kindness shown by Cumbria University for the loan of the LUNGMAN N1 anthropomorphic chest phantom, and are grateful to the Central Manchester Nuclear Medicine Centre for their kind assistance.

Results of this study were presented at the 2014 World Congress of the International Society of Radiographers and Radiographic Technologists.

REFERENCES

http://download.journals.elsevierhealth.com/pdfs/journals/00338389/PIIS0033838911000029.pdf
[13] Siemens AG Medical Solutions. SOMATOM Sensation 10/16 application guide. [Product information: C2-023.630.19.01.02]


Table 1 Acquisition settings used for CTAC

Figure 1 (A) Symbia T6 and (B) Infinia Hawkeye 4 images

Figure 2 Reader averaged AFROC curves for both treatments

Table 2 JAFROC FOM with standard deviations (SD) and 95% confidence interval (CI) for each treatment

Table 3 Treatment Difference between the two imaging systems and 95% confidence interval

Table 4 Lesion detection data (averaged over cases) where the total number of readers was 22. (S-T6 = Symbia T6 and IH-4 = Infinia Hawkeye 4)

Figure 3 Symbia T6 lesion detection data expressed as a % of readers that detected the lesion (averaged over cases) with overall average TP confidence ratings for each. Lesion identification (ID) on the x-axis describes lesion density (HU) and size groupings

Figure 4 Infinia Hawkeye 4 lesion detection data expressed as a % of readers that detected the lesion (averaged over cases) with overall average TP confidence ratings for each. Lesion identification (ID) on the x-axis describes lesion density (HU) and size groupings
Figure 1 (A) Symbia T6 and (B) Infinia Hawkeye 4 images

Figure 2 Reader averaged AFROC curves for both treatments
Table 1: Acquisition settings used for CTAC
<table>
<thead>
<tr>
<th>Treatment</th>
<th>JAFROC FOM and SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Siemens Symbia T6</td>
<td>0.810 (SD = 0.014)</td>
<td>(0.781 , 0.839)</td>
</tr>
<tr>
<td>2: GE Infinia Hawkeye 4</td>
<td>0.599 (SD = 0.015)</td>
<td>(0.568 , 0.631)</td>
</tr>
</tbody>
</table>

**Table 2** JAFROC FOM with standard deviations (SD) and 95% confidence interval (CI) for each treatment

<table>
<thead>
<tr>
<th>Treatment Difference (Treatment pairing 1-2)</th>
<th>JAFROC FOM &amp; SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Siemens Symbia T6 2: GE Infinia Hawkeye 4</td>
<td>0.211 (SD = .014)</td>
<td>(0.182 , 0.240)</td>
</tr>
</tbody>
</table>

**Table 3** Treatment Difference between the two imaging systems and 95% confidence interval

<table>
<thead>
<tr>
<th>Lesion Details</th>
<th>No. readers (S-T6)</th>
<th>TP rating (S-T6)</th>
<th>No. readers (IH-4)</th>
<th>TP rating (IH-4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12mm +100HU</td>
<td>21.667</td>
<td>8.61</td>
<td>21.667</td>
<td>7.477</td>
</tr>
<tr>
<td>12mm -630HU</td>
<td>21</td>
<td>7.968</td>
<td>16</td>
<td>6.415</td>
</tr>
<tr>
<td>12mm -800HU</td>
<td>21.333</td>
<td>7.97</td>
<td>3.667</td>
<td>3.637</td>
</tr>
<tr>
<td>10mm +100HU</td>
<td>21</td>
<td>8.08</td>
<td>21</td>
<td>6.567</td>
</tr>
<tr>
<td>10mm -630HU</td>
<td>19.75</td>
<td>6.933</td>
<td>7.75</td>
<td>3.43</td>
</tr>
<tr>
<td>10mm -800HU</td>
<td>18</td>
<td>7.035</td>
<td>8</td>
<td>2.765</td>
</tr>
<tr>
<td>8mm +100HU</td>
<td>15.5</td>
<td>5.85</td>
<td>8</td>
<td>1.75</td>
</tr>
<tr>
<td>8mm -630HU</td>
<td>18</td>
<td>6.682</td>
<td>9.6</td>
<td>2.488</td>
</tr>
<tr>
<td>8mm -800HU</td>
<td>15.333</td>
<td>6.407</td>
<td>1</td>
<td>2.667</td>
</tr>
<tr>
<td>5mm +100HU</td>
<td>8.333</td>
<td>4.957</td>
<td>3.667</td>
<td>1.15</td>
</tr>
<tr>
<td>5mm -630HU</td>
<td>9.75</td>
<td>4.2</td>
<td>3.25</td>
<td>1.308</td>
</tr>
<tr>
<td>5mm -800HU</td>
<td>3.666</td>
<td>1.24</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 4** Lesion detection data (averaged over cases) where the total number of readers was 22.  (S-T6 = Symbia T6 and IH-4 = Infinia Hawkeye 4)