Visual function assessment in medical imaging research

Lanca, C, Lanca, L, Thompson, John D. and Hogg, P

<table>
<thead>
<tr>
<th>Title</th>
<th>Visual function assessment in medical imaging research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>Lanca, C, Lanca, L, Thompson, John D. and Hogg, P</td>
</tr>
<tr>
<td>Type</td>
<td>Article</td>
</tr>
<tr>
<td>URL</td>
<td>This version is available at: <a href="http://usir.salford.ac.uk/36948/">http://usir.salford.ac.uk/36948/</a></td>
</tr>
<tr>
<td>Published Date</td>
<td>2014</td>
</tr>
</tbody>
</table>

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: usir@salford.ac.uk.
Visual Function Assessment in Medical Imaging Research
Carla Lança, PhD

<<abstract>>

Background Medical image perception research relies on visual data to study the diagnostic relationship between observers and medical images. A consistent method to assess visual function for participants in medical imaging research has not been developed and represents a significant gap in existing research.

Methods Three visual assessment factors appropriate to observer studies were identified: visual acuity, contrast sensitivity, and stereopsis. A test was designed for each, and 30 radiography observers (mean age 31 ± 11.7 years) participated in each test.

Results Mean binocular visual acuity for distance was 20/14 for all observers. The difference between observers who did and did not use corrective lenses was not statistically significant ($P = 0.12$). All subjects had a normal value for near visual acuity and stereoacuity. Contrast sensitivity was better than the population norms.

Conclusion All observers had normal visual function and could participate in medical imaging visual analysis studies. Protocols of evaluation and populations norms are provided. Further studies are necessary to fully understand the relationship between visual performance on tests and diagnostic accuracy in practice.

Introduction

Medical image quality needs to be assessed in both clinical and research settings, and image quality testing in the clinical setting must comply with regulations and best practice. Interpretation of radiographs depends on the clarity of visual patterns within the image in addition to the neurological and psychological factors that affect the observer’s analysis. Traditional approaches of assessing image quality involve only physical measures such as noise, resolution, and contrast. These measures can be useful, but they do not predict the combined diagnostic performance of system and observer.

Observer performance assessment is typically performed to quantify system and observer performance. This was typically done using receiver operating characteristic (ROC) analysis, but the location sensitive free-response ROC (FROC) method now offers improved statistical power. The ROC curve is graphical representation of binary classifier that is varied by a decision threshold. In this type of study an observer is
typically asked to decide whether an image is diseased or not and then state confidence in this decision using a rating scale. The observer is not required to indicate the anatomical location of disease with this method. Diagnostic performance is typically quantified as the area under the ROC curve (AUC). ROC analysis has been an important tool in medical imaging, and the development of the FROC paradigm has enabled researchers to perform studies that measure results closer to the clinical reality. FROC analysis removes the uncertainty in ROC methods caused by a lack of location data, where the observer is now required to localize suspicious areas within the image. ROC methods are limited to a single rating per image, whereas FROC methods allow multiple true (lesion localisation) and false (non-lesion localisation) localisations in each image.

A robust approach to assessing image quality should include physical measures (noise, resolution, contrast etc) and observer performance (ROC/FROC). However, we also feel that visual assessment of the observer is also an important stage in the diagnostic process.

This study seeks to identify and characterize suitable measures of observer visual function in the context of medical image evaluation. A series of visual function tests were applied to a group of observers. The intention was to reveal how visual function data should be analysed and how conclusions about visual function can be drawn.

Errors in Medical Imaging

Diagnostic errors in medical imaging have been reported since 1947. In 2013, Lee et al reported an error rate for radiologic examinations of approximately 30%, with some techniques particularly prone to errors. For example, 20% to 50% of chest radiographs are misdiagnosed. Most missed tumors occurred in the apices, paramediastinal, and hilar areas. Difficulty in separating normal structures from signs of early-stage lung cancer was the apparent cause of these errors. Factors contributing to errors in interpretation are complex and hard to isolate. One potential source of error could be due to changes in visual function that decrease an observer’s ability to correctly identify small, solitary pulmonary nodules. This seems to hold true for medical students. However, among radiology residents and board-certified
radiologists, no correlation was found between visual performance and the ability to correctly locate pulmonary nodules. These findings suggest that factors other than visual perception determine a radiologist’s ability to correctly identify solitary pulmonary nodules. Also, some have argued that experience brings an economy of effort and greater efficiency, which can improve visual performance.

Increased workload, cognitive biases, and equipment and technique problems (eg, referring physician failing to adequately communicate the reason for performing the examination) all are factors that contribute to diagnostic errors in radiology. Perception research in medical imaging relies on data to quantify the relationship between visual stimuli and observer recognition. Contrast sensitivity and visual acuity are fundamental quality measures for visual systems. These factors play a major role in perception and affect the ability of an observer to detect pathophysiology on a medical image.

Visual Function

An observer’s ability to process visual information is one of the fundamental links in the diagnostic imaging chain. Visual function is the primary tool through which imaging information is gathered for processing into concrete data. Decreased visual acuity could significantly increase the threshold contrast required to identify high-frequency diagnostic information. Contrast sensitivity is an indicator of visual pattern-detection for stimuli of various sizes. Low-contrast objects are difficult to evaluate and are one of the greatest challenges for observers reviewing images. Contrast sensitivity across all spatial frequencies declines with age. This decline normally starts at 45 years of age, and higher spatial frequencies are more affected than lower frequencies. However, this topic has been subject to very little scrutiny in medical imaging. Quaghebeur et al found that 71% of radiologists felt that regular monitoring of visual acuity should be required for practice and 82% agreed to undergoing such testing.

Measurement of Visual Function

The Snellen Visual Acuity test is a common standard in the measurement of vision. Results from this test take the format 20/x. In this system, the numerator (20) is the distance at which the subject recognizes an optotype (the letters or symbols on the chart), and the denominator (x) is the distance at which a person with standard visual
acuity would recognize the optotype. In this system, 20/10 vision is excellent, because the observer could see at 20 feet what a person with standard vision could see at 10 feet. 20/100 vision is poor, because the observer could only see at 20 feet what a person with standard vision could see at 100 feet.

Numerous charts are used for visual acuity testing, but the ETDRS chart is preferred for vision testing in clinical trials. This study used a Vector Vision ETDRS Chart – CSV 1000 (see Figure 1). Logarithm of the Minimum Angle of Resolution (LogMAR) is a more precise method of calculating visual acuity. In LogMAR notation, lower scores correspond to better vision, and as acuity becomes worse, the value of the LogMAR increases. When converting LogMAR to a Snellen visual acuity measurement, the following equation can be used:

\[
\text{Decimal acuity} = \text{antilog} \left( -\log_{10} \text{LogMAR} \right) = 10^{-\log_{10} \text{LogMAR}}
\]  

This example assumes a LogMAR result of 0.15, which corresponds to a Snellen visual acuity measurement of approximately 20/14 using equation 1:

\[
\begin{align*}
\text{Decimal visual acuity} &= 10^{-0.15} = 1.41 \\
\text{Snellen visual acuity denominator} &= 20/1.41 \\
\text{Snellen visual acuity} &= 20/14
\end{align*}
\]

Grating contrast sensitivity is an important measure of visual function. It measures the ability of an observer to perceive slight changes in luminance between regions which are not separated by definite borders. This is a significant function in imaging analysis of low-contrast targets such as isoechoic lesions on ultrasonography or isodense lesions on computed tomography (CT). These lesions can be recognised only indirectly, through contour irregularities or displacement of identifiable adjacent structures. The perception of complex patterns in mammograms also is linked to the contrast sensitivity of the observer. Pattern-detection is determined by eye contrast sensitivity for stimuli of various sizes. If an observer has abnormal contrast sensitivity, low-contrast targets might be difficult to identify and can be missed. Figure 2 shows an anthropomorphic chest phantom containing a small simulated nodule (A) which is only distinguished because of the irregular shape of the vessel (B).
Grating contrast sensitivity is measured by detection of sinusoidal gratings, which are patterns of parallel light and dark bars. Contrast sensitivity testing measures the eye’s sensitivity over a range of spatial frequencies represented by bar widths. The spatial frequency, measured in cycles per degree (cpd), is a measure of how often sinusoidal components of the structure repeat per unit of distance. Normal contrast sensitivity maximizes at a spatial frequency of about 6 cpd and declines at both higher and lower spatial frequencies (Figure X).

Digital images could be represented both in spatial and frequency domains. The spatial domain refers to a matrix of grey level intensities in a 2-D spatial plane. The frequency domain refers to the rate of change of intensity in an image in terms of sinusoidal intensity profile.

Stereopsis is a measure of an observer’s ability to perceive 3-D features (ie, the ability to perceive the depth of an object) and thus obtain binocular single vision. Binocular single vision is the coordination of both eyes, fusing 2 slightly differing images into a whole image with 3-D perception. Although the radiographic image is a 2-D depiction, it represents 3-D anatomy. Radiographs are created from the shadows of the x-ray absorption pattern as they pass through the body, including information from multiple planes in 3-D space. The observer must translate the image into a 3-D representation to properly analyse and localize structures.

Recent investigations indicate that stereopsis is an advantage, and probably could be an impact factor in the comprehension of complex radiographs with many visible structures. Stereopsis is an important function that helps the observer when reviewing spatial information from medical imaging. Figure 3 shows a 3-D reconstruction of a chest CT scan in which the anatomic structures are volumetrically represented. Stereopsis allows the observers to determine the depth of objects in the central visual area, enhancing vision quality.

Stereopsis is measured using a unit called minutes/seconds of arc, which is a measure of angular distance. A minute of arc is equal to 1/60th of one degree; a second of arc is equal 1/3600th of one degree. This is a measure of the angle of separation between the two eyes, and the distance from the eyes to the object. The smaller the angle is, the better the observer’s ability. Some standard stereopsis test measures from 3500 to 20 seconds of arc, which can provide an idea of the kind of tiny angles the human eye is capable of recognizing.
Methods

Thirty observers with a mean age of 31.6 ± 11.7 years (17-57 years) agreed to be visual assessed before evaluating plain radiographs and CT scans. The observer group included radiographers, student radiographers, and medical physicists. Ethical standards for the study complied with Lisbon School of Health Technology requirements, and observers were given feedback on the results of the visual function tests they participated in.

The observers received a visual assessment that included visual acuity, contrast sensitivity, and stereopsis tests. These visual function tests measured abilities necessary for assessing medical image quality (see Table 1). The most commonly measured aspect of visual function is visual acuity. Visual acuity describes the ability of the eye to resolve the size of an object. In radiology, this function is important for an observer’s ability to identify small, solitary pulmonary nodules, for example. A nodule could be missed because the visual acuity of the observer is not sufficient to resolve the size of the nodule (see Figure 4).

Eye charts are used to measure visual acuity. These charts consist of uppercase letters arranged in rows, with the largest letters at the top with progressively smaller letters toward the bottom. The observers’ visual acuity for distance was assessed in low-light conditions at a distance of 8 feet with an illuminated ETDRS chart in a backlit stand (see Figure 1).

The CSV chart incorporates LED light source technology and auto-calibrate the test luminance to 85 candela per square meter (cd/m²). The CSV1000 chart has the advantage of having 5 letters on every row, equal spacing of the rows on a log scale (separated by 0.1 log unit), equal spacing of the letters on a log scale, and letter difficulty balanced for each row. Vision testing begins with the left-most letter on the top row of the chart. Visual acuity was recorded at the last line on which the observer correctly identified at least 3 of the 5 letters. Visual acuity is worse than average above 0.0 LogMAR or when the denominator of the Snellen visual acuity measurement is greater than 20 (eg, 20/100). Near visual acuity was assessed in well-lit conditions with both eyes at a distance of 40 cm using a LogMAR chart (Good-Lite), (see Figure 5). Visual acuity was recorded at the last line on which the observer correctly identified at least 3 of the 5
letters. Visual acuity was considered abnormal when greater than 1M (the M-unit is the unit of letter size). In this study, contrast sensitivity was assessed in low-light conditions with the CSV1000E contrast chart. The chart consists of a matrix of sinusoidal gratings: circles filled with dark and light bars. Spatial frequency increased from top to bottom, and contrast decreased from left to right (see Figure 6). For the purpose of our study, spatial frequency was divided into low (3 cpd), medium (6 cpd) or high (12 and 18 cpd) categories. The contrast level of the last circle the observer correctly identified on each row was recorded as the contrast sensitivity score for that row. The procedure was repeated for each row in descending order. Visual contrast was considered abnormal when less than 1.61 for 3 cpd; less than 1.66 for 6 cpd; less than 1.08 for 12 cpd; and less than 0.56 for 18 cpd.

Stereoacuity was assessed with a Stereo Butterfly test (Stereo Optical Company) at 40 cm. For this test, a card with superimposed images of circles was shown to the observer to measure the ability to detect the elevation of the circles above the plane of the card (see Figure 7). The circles indicate a stereopsis level ranging from 800 to 40 seconds of arc. The standard stereopsis test has been applied with results equal or shorter than 50 seconds of arc considered normal stereoscopic acuity. All subjects who usually wear corrective lenses were asked to wear them during vision testing to assure that refractive errors were corrected. Refractive error refers to the amount of myopia, hyperopia, or astigmatism that might affect the testing. These conditions require correction with glasses or contact lenses to achieve the best possible corrected visual acuity. Because the observers typically would use both eyes to perform image evaluation, the 3 visual functions were measured with both eyes open.

Results

Of the 30 observers, 3 (10%) could not recall ever having their vision examined and 5 (16.7%) were examined approximately 5 years before the study began. Eleven observers (36.7%) reported that their most recent vision examination had been performed within the previous 2 years.

Visual acuity, contrast sensitivity, and stereoacuity distributions were recorded. The mean visual acuity for distance was 20/14, with a minimum of 20/10 and a maximum of 20/20. All subjects had maximal visual acuities of 20/20 (LogMAR, 0.0)
The mean visual acuity of female observers (n = 15) was 20/14 (LogMAR, -0.15) and for male observers (n = 15) it was 20/14 (LogMAR, -0.16). Subgroup analyses by sex revealed no significant differences (P = 0.46).

Seven subjects wore corrective lenses. They had a mean visual acuity of 20/16 (LogMAR, -0.11). Participants who did not wear corrective lenses had a mean visual acuity of 20/14 (LogMAR, -0.17). The difference between observers who used corrective lenses and those who did not was not statistically significant (P = 0.12).

All subjects had normal near visual acuity and near normal stereoaucuity (40.67 ± 2.54 seconds of arc). The log average values of contrast sensitivity for each spatial frequency were better than the population norms (see Table 3). Of the 30 observers, one had low contrast sensitivity for spatial frequencies of 6 cpd (1.55) and of 12 cpd (0.31). However, this observer’s contrast sensitivity was improved after receiving a new optical prescription to 6 cpd (2.14) and 12 cpd (1.25).

The difference between observers who used corrective lenses and those who did not was significant for the higher spatial frequencies (18 cpd spatial frequency P = 0.012). Observers who used corrective lenses of any kind had a lower log average value of contrast sensitivity (1.34 ± 0.15) than those who did not (1.52 ± 0.08). Subgroup analyses by sex revealed no significant differences (P > 0.05) for all spatial frequencies.

**Discussion**

For this group of observers, the minimum criteria for participation in medical imaging studies were met. All observers’ visual abilities were fit for the study, although one observer required a new optical prescription. All subjects achieved the normative range and had acceptable results for the 3 visual functions. These 3 abilities are necessary for medical image evaluation in 2 codependent tasks: detection (visual acuity and contrast sensitivity) and localisation (stereopsis). For these preliminary results, we used population norms as a measure of adequacy and supposed that these norms would apply to medical imaging. This assumption needs to be supported with more research in this field.

The mean binocular visual acuity of participants for distance was 20/14. Our findings are comparable with one previous study, which reported the mean acuity of radiologists as 20/15.12 In the present study visual acuity was not significantly different when comparing observers by sex and all subjects had maximal visual acuities measuring
LogMAR 0.0 (20/20) or better for distance. Visual acuity influences the ability to detect nodules which makes the assessment of this function an important measure for medical imaging. All subjects had normal values for contrast sensitivity. Only one observer had low contrast sensitivity for spatial frequencies of 6 cpd and 12 cpd. The observer’s glasses were updated to a new prescription, which resolved this problem.

We detected a statistically significant difference between observers who used corrective lenses and those who did not in the log average values for higher spatial frequencies (18cpd). This difference favored the participants who did not use corrective lenses. One cannot conclude, however, that there would be any resultant clinical effect on medical image observer studies. Only 7 observers used corrective lenses, and although a statistically significant difference was detected, both values were within the normal range. It is unlikely that this difference would adversely influence clinical performance.

In the present study, we measured visual function with different charts for each function. However, in future studies these tests could be standardized and displayed on the same monitor used for medical image display. This will provide more accurate information about visual function for the observer’s actual practice and could be applied to the set of visual tests in this study.

Three observers could not recall ever having their vision examined and 5 were examined approximately 5 years ago. The elapsed time since the last reported eye examination raises the question of whether regular examinations are important to medical imaging. No strict international recommendations in medical imaging exist, although it is recommended that even those with no signs or risk factors for eye disease receive a comprehensive eye evaluation at 40 years of age. More research is necessary to identify norms and guidelines for visual performance evaluation in medical imaging.

Observers without prescribed corrective lenses should be tested, and when visual anomalies are detected they must be excluded from studies that involve image interpretation unless the corrective lenses are used. Observers with corrective lenses should have routine eye examinations and, if necessary, an updated prescription to ensure they maintain maximal visual performance.

The authors plan to continue the evaluation of the role of visual function in image interpretation to determine the level of decreased eye function that could lead to errors in image interpretation. Future challenges are related to the visual function tests,
which could be correlated with the screen viewing distance used by observers to provide more relevant information about diagnostic accuracy.

An assessment of visual function of medical image observers is absent in the literature. Although quality control programs have been implemented for the performance of digital displays, similar attention has not been devoted to quality control for radiologists and other health care professionals who examine the results of medical imaging systems. Our study proposes a range of visual tests suitable for assessing observers before participation in studies on medical imaging. We propose that these visual function tests be conducted on potential observers prior to conducting medical imaging research using perceptual methodologies.

**Conclusion**

Quality standards for visual assessment should be implemented to decide if an observer has adequate eye function to participate in medical imaging observer studies. A method has been provided for visual function assessment of observers prior to medical imaging perceptual research studies. Normal visual function should be assured prior to performing vision-based tasks in medical imaging. Protocols of evaluation and population norms have been provided with the assessment of three functions (visual acuity, contrast sensitivity, and stereopsis). Observers with visual function anomalies should be excluded from observer studies that involve image evaluation and interpretation, unless corrective lenses are prescribed and used. Further studies are necessary to clarify the relationship between visual function and diagnostic performance.

<<Author Bio>>

*Carla Lança, PhD* is lecturer in Orthoptics at the Orthoptic Department in Lisbon School of Health Technology in Portugal. *Her current research interest is in visual function and it’s assessment in medical imaging research.*

*John D. Thompson, PhD* is an associate lecturer at the University of Salford. He is a diagnostic radiographer specializing in nuclear medicine and his current interest is in observer performance, with a particular focus on free-response receiver operating characteristic (FROC) methods and jackknife alternative FROC (JAFROC) analysis. Most recent research has assessed nodule detection performance in low-resolution CT
Luis Lança, PhD is a senior lecturer in Radiography at Lisbon School of Health Technology and lead for Radiography Programme. He is also affiliated Senior Research Specialist in Radiography at the Department of Clinical Science Intervention and Technology at Karolinska Institutet, Sweden.

Peter Hogg is Chair/full Professor at University of Salford, Lead for Diagnostic Imaging Research Programme, Research Dean and Director of the Health Sciences Research Centre. He is also affiliated Senior Research Specialist in Radiography at the Department of Clinical Science Intervention and Technology at Karolinska Institutet, Sweden.

Received January 15, 2015; accepted after revision March 11, 2015.
Reprint requests may be mailed to the American Society of Radiologic Technologists, Communications Department, at 15000 Central Ave SE, Albuquerque, NM 87123-3909, or e-mailed to communications@asrt.org.
© 2015 American Society of Radiologic Technologists

References


Table 1. Visual functions Necessary for Assessing Medical Image Quality

<table>
<thead>
<tr>
<th>Visual Function</th>
<th>Basis for Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual acuity</td>
<td>Makes possible the accurate detection of size\textsuperscript{16,17} radiologic anatomic structures.</td>
</tr>
<tr>
<td>Contrast sensitivity</td>
<td>Makes possible the discrimination of low-contrast and high-contrast frequency information.\textsuperscript{19,20}</td>
</tr>
<tr>
<td>Stereopsis</td>
<td>Reduces the amount of visual scanning necessary to extract spatial information, which sustains comprehension of complex visual experiences.\textsuperscript{25,27} Most importantly, this function provides visual memory with a 3-D interpretation.</td>
</tr>
</tbody>
</table>

Table 2. Descriptive Statistics Data From the 3 Visual Function Tests

<table>
<thead>
<tr>
<th>Visual Function Test</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Snellen LogMAR ± SD</td>
<td>Snellen LogMAR</td>
<td>Snellen LogMAR</td>
</tr>
<tr>
<td>Distance acuity</td>
<td>20/14 0.15 ± 0.07</td>
<td>20/10 0.30</td>
<td>20/20 0.00</td>
</tr>
<tr>
<td>Near acuity</td>
<td>20/20 0.39 ± 0.02</td>
<td>0.32</td>
<td>0.40</td>
</tr>
<tr>
<td>Stereopsis (seconds)</td>
<td>40.67 ± 2.54</td>
<td>40</td>
<td>50</td>
</tr>
</tbody>
</table>

Log Average Values of Contrast Sensitivity

| (3 cpd) | 1.85 ± 0.09 | 1.63 | 2.08 |
| (6 cpd) | 2.16 ± 0.15 | 1.70 | 2.29 |
| (12 cpd) | 1.89 ± 0.16 | 1.25 | 1.99 |
| (18 cpd) | 1.47 ± 0.12 | 1.25 | 1.55 |

Abbreviations: cpd, cycles per degree; SD, standard deviation.

Table 3. Normative Data for Analysing Visual Function Tests Results

<table>
<thead>
<tr>
<th>Visual functions</th>
<th>Population Norms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance acuity</td>
<td>0.0 LogMAR\textsuperscript{28}</td>
</tr>
<tr>
<td>Near acuity</td>
<td>1M\textsuperscript{29}</td>
</tr>
<tr>
<td>Stereopsis</td>
<td>50 seconds\textsuperscript{25}</td>
</tr>
<tr>
<td><strong>Contrast Sensitivity</strong></td>
<td></td>
</tr>
<tr>
<td>3 cpd</td>
<td>1.61 ± 0.21\textsuperscript{10}</td>
</tr>
</tbody>
</table>
Abbreviation: cpd, cycles per degree.

<table>
<thead>
<tr>
<th>cpd</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 cpd</td>
<td>$1.66 \pm 0.23$</td>
</tr>
<tr>
<td>12 cpd</td>
<td>$1.08 \pm 0.32$</td>
</tr>
<tr>
<td>18 cpd</td>
<td>$0.56 \pm 0.35$</td>
</tr>
</tbody>
</table>

Figure 1. ETDRS – CSV 1000 Vector Vision Chart. Image courtesy of author CL.

Figure 2. A simulated 5-mm pulmonary nodule (A) that is only distinguishable from simulated pulmonary vessels (B) by the shape of the object. Image courtesy of author JDT.

Figure 3. A 3-D reconstruction of a chest computed tomography acquisition. Image courtesy of author JDT.

Figure 4. The 3 arrows in the magnified box indicate a 5-mm simulated solitary pulmonary nodule in a chest phantom model. Image courtesy of author JDT.

Figure 5. LogMar Good-Lite chart. Image courtesy of author CL.

Figure 6. CSV1000E Vector Vision contrast chart. Image courtesy of author CL.

Figure 7. The Stereo Butterfly test (Stereo Optical Company). Image courtesy of author CL.