Diagnostic Accuracy and Visual Search Efficiency: Single 8 MP vs. Dual 5 MP Displays

Elizabeth A. Krupinski¹

Published online: 31 October 2016
© Society for Imaging Informatics in Medicine 2016

Abstract This study compared a single 8 MP vs. dual 5 MP displays for diagnostic accuracy, reading time, number of times the readers zoomed/panned images, and visual search. Six radiologists viewed 60 mammographic cases, once on each display. A sub-set of 15 cases was viewed in a secondary study using eye-tracking. For viewing time, there was significant difference ($F = 13.901$, $p = 0.0002$), with 8 MP taking less time (62.04 vs. 68.99 s). There was no significant difference ($F = 0.254$, $p = 0.6145$) in zoom/pan use (1.94 vs. 1.89). Total number of fixations was significantly ($F = 4.073$, $p = 0.0466$) lower with 8 MP (134.47 vs. 154.29). Number of times readers scanned between images was significantly fewer ($F = 10.305$, $p = 0.0018$) with 8 MP (6.83 vs. 8.22). Time to first fixate lesion did not differ ($F = 0.126$, $p = 0.7240$). It did not take any longer to detect the lesion as a function of the display configuration. Total time spent on lesion did not differ ($F = 0.097$, $p = 0.7567$) (8.59 vs. 8.39). Overall, the single 8 MP display yielded the same diagnostic accuracy as the dual 5 MP displays. The lower resolution did not appear to influence the readers' ability to detect and view the lesion details, as the eye-position study showed no differences in time to first fixate or total time on the lesions. Nor did the lower resolution result in significant differences in the amount of zooming and panning that the readers did while viewing the cases.

Keywords Digital display · Image perception · Observer performance

✉ Elizabeth A. Krupinski
ekrupin@emory.edu

¹ Department of Radiology & Imaging Sciences, Emory University, 1364 Clifton Rd NE Room D107, Atlanta, GA 30322, USA

Introduction

A recent study [1] evaluated 14 primary clinical display monitors in a large clinical setting using the American Association of Physicists in Medicine Task Group 18 [2] methods and guidelines (measuring ambient light, display uniformity, luminance ratio, luminance response, maximum luminance, spatial resolution). They found that nearly 30 % of the displays failed to meet at least one of the guideline minimum acceptability levels. It was noted that the problem can be particularly impactful as typical radiology workstations use two monitors and even slight differences in appearance as a result of not meeting the minimum standards could result in images looking very different, potentially leading to interpretation errors. These results are not unique [3, 4]. Additionally, there has been much published in recent years on the digital reading environment and the need to pay attention to ergonomic and human factors concerns and the daily use of electronic displays [5–8]. Part of the concern lies in growing awareness of physical complaints from radiologists like neck, back, and shoulder pain as they view images on multiple monitors [9, 10]. Part of the concern also lies in the increasing awareness of how fatigued radiologists are today with the growing number of images that are required to be reviewed in a given period of time [11–13]; and the fact that fatigue has been shown to negatively impact detection accuracy [14–16].

One potential solution to maintain image quality and reduce physical strain issues is to adopt the use of single displays rather than multiple, at least for the viewing of clinical images (a side display for accessing patient records or reporting systems may still be required). This option is quite feasible as there are a number of medical-grade display manufacturers developing and marketing high-resolution wide screen displays. There have, however, not been any major studies assessing the impact of viewing multiple images on a single display vs. two side-by-

side displays. There are reasons why performance could differ. For high-resolution images such as mammography, the single screen displays are generally close but not the same resolution as the two side-by-side displays. On the flip side, the dual display format, no matter how closely together they are positioned, will always have a gap between the actual display areas as by design necessity the monitors require bezels. A single display does not have this problem.

The goal of this study, therefore, was to assess diagnostic accuracy and reader efficiency as a function of display and display layout. This study had two reading conditions: single 8 MP (no bezel) display vs. dual 5 MP (with bezel) displays. There were three aspects of reader performance that were studied: diagnostic accuracy as measured by receiver operating characteristic (ROC) analysis, reading time, and number of times the readers zoomed/panned the images.

Methods

The study was carried out using display monitors (Eizo, Hakusan, Japan) that were as equivalent except for the parameter of interest (8 MP no bezel (RadiForce RX850) vs. dual 5 MP (RadiForce GX540) with bezel). They were set up with as equivalent as possible with respect to white points, maximum/minimum luminance, and black levels, and were calibrated to the DICOM GSDF (Digital Imaging and Communications in Medicine Grayscale Standard Display Function). Ambient room lights were set to 25 lx. The study protocol was approved by the University of Arizona's Human Subjects office for IRB approval prior to the start of the study.

Six mammographers were recruited as observers—four Board-Certified mammographers, one mammography Fellow, and one senior resident about to enter a mammography Fellowship. Each observer viewed 60 mammographic cases, once on each display format using a counterbalanced design in which half of the observers viewed the cases in one condition first and the other half viewed them on the alternative condition first. At least 3 weeks passed between sessions to promote forgetting of the cases. The observers were asked to first decide if each case was normal or abnormal, then report their confidence in that decision as possible, probable or definite, yielding a six-point response scale. Diagnostic performance was then measured using the confidence ratings in a multi-reader multi-case (MRMC) receiver operating characteristic analysis [17].

Forty of the cases contained a single malignant lesion (20 masses and 20 microcalcification clusters), and 20 were lesion free. All cases with lesions had been verified by biopsy regarding lesion status, and the lesion-free cases had at least 2-years follow-up without change in status. Lesions ranged from subtle to moderately subtle (as judged by an independent mammographer not participating in the study) and were located throughout the breast. Cases were chosen to represent a range

of breast densities, but approximately 2/3 of the lesion-containing and lesion-free cases were classified in the original reports as dense in order to represent challenging cases.

An additional study with eye-position recording, using a sub-set of 15 cases (5 mass, 5 calcification, 5 normal), was also conducted. The ASL SU4000 Eye-Tracker system (Applied Science Labs, Bedford, MA) was used to record eye-position. The eye-position data were analyzed using standard methods. Parameters analyzed were total viewing time, number of fixations generated, time to first hit lesion (in either image), total time on lesion (sum for both views), and number of times scanned from one image to the other. For this study, the same six observers participated, at least 3 months after the main diagnostic study.

Results

Diagnostic Performance Study

The MRMC ROC analysis revealed no statistically significant difference in diagnostic performance between the two display conditions ($F = 4.43$, $p = 0.0891$). Average Az with the dual 5 MP displays was 0.849 and with the single 8 MP display was 0.823 (see Table 1). For viewing time, there was a significant difference ($F = 13.901$, $p = 0.0002$) between the two display formats, with the single 8 MP display taking less time on average than the dual 5 MP format. On average, viewing time with the single display format was 62.04 s (sd = 24.09) and for the dual format was 68.99 (sd = 25.87). There were no differences as a function of type of lesion ($F = 0.144$, $p = 0.8657$). Figure 1 shows the average reading time for each of the six readers in the two viewing conditions.

There was no significant difference ($F = 0.254$, $p = 0.6145$) in the number of times zoom/pan was used by the readers as a function of single vs. dual monitors. On average, frequency of zoom/pan use with the single display format was 1.94 times per case (sd = 1.40) and for the dual format was 1.89 times per case (sd = 1.38). There was no difference as a function of lesion type ($F = 0.292$, $p = 0.7467$). Figure 2 shows the average zoom/pan use for each of the six readers in the two viewing conditions.

Table 1 Diagnostic performance between the two display conditions

Reader	Dual 5 MP	Single 8 MP
1	0.869	0.840
2	0.819	0.779
3	0.816	0.767
4	0.817	0.825
5	0.894	0.849
6	0.876	0.890
Mean	0.849	0.823

Fig. 1 Viewing times for each of the readers in each viewing condition

Eye-Position Study

As in the larger study, the total viewing time was significantly shorter ($F = 4.372$, $p = 0.0394$) with the single display (mean = 54.65, sd = 24.09) than with the dual displays (mean = 62.86, sd = 27.58). The total number of fixations generated with the single display was also significantly ($F = 4.073$, $p = 0.0466$) shorter (mean = 134.47, sd = 65.14) than with the dual display (mean = 154.29, sd = 65.09). The number of times they scanned from one image to the other (Fig. 3) was significantly fewer ($F = 10.305$, $p = 0.0018$) with the single (mean = 6.83, sd = 2.58) than with the dual display (mean = 8.22, sd = 2.99).

The time to first fixate the lesion (in either image) did not differ significantly ($F = 0.126$, $p = 0.7240$) between the single (mean = 6.83, sd = 3.47) and dual displays (mean = 6.63, sd = 3.64). In other words, it did not take any longer to detect the lesion as a function of the display configuration. The total time spent on the lesion (in both views) did not differ

Fig. 2 Frequency of use of zoom/pan for each of the readers in each viewing condition

Fig. 3 Number of times scanned from one image to the other for each of the readers in each viewing condition

significantly ($F = 0.097$, $p = 0.7567$) between the single (mean = 8.59, sd = 4.43) and dual displays (mean = 8.39, sd = 4.18).

Discussion

Overall, the single 8 MP display yielded the same diagnostic accuracy as the dual 5 MP displays. The lower resolution did not appear to influence the readers' ability to detect and view the lesion details, as the eye-position study showed no differences in time to first fixate or total time on the lesions. Nor did the lower resolution result in significant differences in the amount of zooming and panning that the readers did while viewing the cases. The use of these wide-screen high-resolution monitors seems to be quite feasible for use in primary radiographic interpretation.

Interestingly, there were significant improvements in reader efficiency with the single 8 MP display compared to the two 5 MP displays. This was seen in terms of reduced overall time spent viewing the images in both the main and eye-position studies. From the eye-position study, it appears that the gain in efficiency is not due to detecting or spending any different time on the lesions, but in terms of not scanning back-and-forth from image to image as much. The reason for this may be the presence of the bezels around each of the two 5 MP displays—creating a physical separation between the two images, while the single 8 MP display has the two images abutting each other without anything between them. To confirm this, additional studies are required.

Limitations of this study include the use of only one type of image (mammograms) and readers (mammographers). We chose these specifically, however, as mammograms are the highest resolution images in radiology, and one would expect

the results to generalize to lower resolution images. The laboratory setting may also have influenced the results as it does not replicate entirely the clinical environment in the sense that there are often additional monitors than just those for viewing images (e.g., side monitors for accessing patient data or dictations), and the number and types of images (e.g., priors) is typically greater.

Conclusion

A single 8 MP display, properly calibrated and used in the appropriate radiology viewing environment, is equivalent to two side-by-side 5 MP displays for accurately interpreting mammographic exams. This configuration may actually present some advantages in terms of reduced visual search time and more efficient visual search.

References

1. Thompson DP, Koller CJ, Eatough JP: Practical assessment of the display performance of radiology workstations. *Br J Radiol* 80: 256–260, 2007
2. Assessment of display performance for medical imaging systems: executive summary of the AAPM TG18 report. *Med Phys* 32: 1205–1225, 2005
3. Wade C, Brennan PC: Assessment of monitor conditions for the display of radiological diagnostic images and ambient lighting. *Br J Radiol* 77:465–471, 2004
4. Butt A, Mahoney M, Savage NW: The impact of computer display performance on the quality of digital radiographs: a review. *Aust Dent J* 57:16–23, 2012
5. Prabhu SP, Gandhi S, Goddard PR: Ergonomics of digital imaging. *Br J Radiol* 78:582–586, 2005
6. Krupinski EA, Kallergi M: Choosing a radiology workstation: technical and clinical considerations. *Radiology* 242:671–682, 2007
7. Hirschorn DS, Krupinski EA, Flynn MJ: PACS displays: how to select the right display technology. *J Am Coll Radiol* 11:1270–1276, 2014
8. Kagadis GC, Walz-Flannigan A, Krupinski EA, Nagy PG, Katsanos K, Diamantopoulos A, Langer SG: Medical imaging displays and their use in image interpretation. *RadioGraphics* 33:275–290, 2013
9. Thompson AC, Kremer MJ, Biswal S, Rebner M, Rebner RE, Thomas WR, Edwards SD, Thompson MO, Ikeda DM: Factors associated with repetitive strain, and strategies to reduce injury among breast-imaging radiologists. *J Am Coll Radiol* 11:1074–1079, 2014
10. Boisselle PM, Levine D, Horwich PJ, Barbaras L, Siegal D, Shillue K, Affeln D: Repetitive stress symptoms in radiology: prevalence and response to ergonomic interventions. *J Am Coll Radiol* 5:919–923, 2008
11. Krupinski E, Reiner BI: Real-time occupational stress and fatigue measurement in medical imaging practice. *J Digit Imaging* 25:319–324, 2012
12. Reiner BI, Krupinski E: The insidious problem of fatigue in medical imaging practice. *J Digit Imaging* 25:3–6, 2012
13. Reiner BI, Krupinski E: Demystifying occupational stress and fatigue through the creation of an adaptive end-user profiling system. *J Digit Imaging* 25:201–205, 2012
14. Krupinski EA, Berbaum KS, Caldwell RT, Schartz KM, Kim J: Long radiology workdays reduce detection and accommodation accuracy. *J Am Coll Radiol* 7:698–704, 2010
15. Krupinski EA, Berbaum KS, Caldwell RT, Schartz KM, Madsen MT, Kramer DJ: Do long radiology workdays affect nodule detection in dynamic CT interpretation? *J Am Coll Radiol* 9:191–198, 2012
16. Taylor-Phillips S, Elze MC, Krupinski EA, Dennick K, Gale AG, Clarke A, Mello-Thoms C: Retrospective review of the drop in observer detection performance over time in lesion-enriched experimental studies. *J Digit Imaging* 28:32–40, 2015
17. Dorfman DD, Berbaum KS, Metz CE: Receiver operating characteristic rating analysis: generalization to the population of readers and patients with the jackknife method. *Investig Radiol* 27:723–731, 1992