

Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Image toggling saves time in mammography

Trafton Drew
Avi M. Aizenman
Matthew B. Thompson
Mark D. Kovacs
Michael Trambert
Murray A. Reicher
Jeremy M. Wolfe

SPIE.

Image toggling saves time in mammography

Trafton Drew,^{a,*} Avi M. Aizenman,^b Matthew B. Thompson,^{b,c} Mark D. Kovacs,^d Michael Trambert,^{e,f,g} Murray A. Reicher,^g and Jeremy M. Wolfe^{b,h}

^aUniversity of Utah, Department of Psychology, Salt Lake City, Utah 84122, United States

^bBrigham and Women's Hospital, Department of Surgery, Cambridge, Massachusetts 02139, United States

^cThe University of Queensland, School of Psychology, Brisbane, Queensland 4072, Australia

^dMedical University of South Carolina, Department of Radiology, Charleston, South Carolina 29425, United States

^eCottage Health System, Department of Radiology, Santa Barbara, California 93110, United States

^fThe Sansum Clinic, Department of Radiology, Santa Barbara, California 93110, United States

^gMerge Healthcare, San Diego, California 92121, United States

^hHarvard Medical School, Department of Surgery, Cambridge, Massachusetts 02139, United States

Abstract. When two images are perfectly aligned, even subtle differences are readily detected when the images are “toggled” back and forth in the same location. However, substantial changes between two photographs can be missed if the images are misaligned (“change blindness”). Nevertheless, recent work from our lab, testing nonradiologists, suggests that toggling misaligned photographs leads to superior performance compared to side-by-side viewing (SBS). In order to determine if a benefit of toggling misaligned images may be observed in clinical mammography, we developed an image toggling technique where pairs of new and prior breast imaging exam images could be efficiently toggled back and forth. Twenty-three radiologists read 10 mammograms evenly divided in toggle and SBS modes. The toggle mode led to a 6-s benefit in reaching a decision [$t(22) = 5.11, p < .05$]. The toggle viewing mode also led to a 5% improvement in diagnostic accuracy, though in our small sample this effect was not statistically reliable. Time savings were found even though successive mammograms were not perfectly aligned. Given the ever-increasing caseload for radiologists, this simple manipulation of how the images are viewed could save valuable time in clinical practice, allowing radiologists to read more cases or spend more time on difficult cases. © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.3.1.011003]

Keywords: detection; visual search; perceptual errors; display technology; change blindness.

Paper 15138SSR received Jul. 8, 2015; accepted for publication Sep. 9, 2015; published online Oct. 12, 2015.

1 Introduction

Breast cancer is a leading cause of cancer mortality, and early detection reduces the mortality rate.¹ Breast cancer screening in the United States has resulted in an increase in the number of early stage breast cancers that are detected and decreased the number of late-stage cancers detected.^{1,2} However, there is still substantial room for improvement, as there is an ongoing debate over whether the costs of screening mammography (overdiagnosis, monetary cost) outweigh the benefits (reduced mortality rate).^{2–4} It is also well established that diagnostic performance in digital mammography is imperfect: studies indicated that up to 20% to 30% of breast cancers are missed.⁵

The medical imaging field has risen to this challenge by providing a number of technological advances that are designed to improve diagnostic accuracy in mammography. For example, due to evidence that full-field digital mammography leads to better outcomes than screen-film mammograms, most practices in the United States now utilize the digital modality.⁶ It is less clear whether computer-aided detection (CAD) leads to reliable benefits in screening mammography. CAD systems use computer algorithms to mark potential abnormalities for the radiologist to evaluate. Studies of the influence of CAD on diagnostic performance typically demonstrate that CAD leads to a higher

proportion of cancers detected.⁷ However, in some large studies, this benefit also results in an accompanied increase in false positives.^{8–10} Despite these findings, approximately 75% of all mammograms in America are read with the help of CAD.¹¹ More recently, digital breast tomosynthesis (DBT) has shown great promise. A number of studies have demonstrated that adoption of DBT leads to superior diagnostic accuracy to full-field digital mammography.^{12,13} In particular, the primary benefit of DBT appears to be a reduction in false positives without any change in the rate of cancers detected.

While each of these advances has led to improvements in diagnostic performance under some circumstances, each also comes along with significant costs. For example, DBT increases the number of images in a screening mammogram from four to generally between 250 and 300, thus increasing network/archive costs and required professional time. In addition, a practice seeking to implement DBT for even a single device generally must invest >\$350,000 (USD) in equipment, installation, and training costs. This translates into significantly higher healthcare costs. According to the American College of Radiology, the CMS reimbursement as of 2015 for a screening film-based mammogram is \$82.59, for a digital mammogram is \$134.80, and for a digital mammogram with tomosynthesis is \$190.93.¹⁴ Furthermore, the use of DBT in addition to digital mammography substantially increases the radiation dose to

*Address all correspondence to: Trafton Drew, E-mail: trafton.drew@psych.utah.edu

the patient,¹² though this concern may be mitigated by creating synthetically reconstructed digital mammograms from the DBT images.¹⁵

The purpose of the current study was to evaluate the promise of a different sort of technological advance in mammography. Rather than creating a new method of imaging the breast, we were interested in examining the influence of a different way of viewing pre-existing images. Our interest in examining the benefits of a simple alternative method for viewing pairs of medical images grows from basic research in visual perception and visual attention. Over the past 20 years, there has been considerable academic interest in the phenomenon of “change blindness.” Change blindness is a failure to detect differences between two successive images.^{1,16,17} This phenomenon is most commonly produced by introducing a brief blank interval between consecutive presentations of images.^{2-4,16} Generally, the images are identical except for a critical change, typically the disappearance of an object. Without a blink interval, it is easy to detect even a small change between two otherwise identical images. As the images are alternated, most of the contents are static, but the change produces a local transient that immediately attracts attention. This was the basis for the “blink comparator,” used by astronomers to compare two views of the same patch of sky—a method used to discover Pluto.^{5,18}

Eye movements between pairs of images presented side-by-side (SBS) work like blank intervals in toggling to hide changes.^{6,19} Images from the current and past exams are typically compared SBS. If they were toggled, one after the other, in the same location, might detection of cancer work like detection of Pluto? Unfortunately, we know that even subtle differences or misalignments between images can induce change blindness as effectively as the blank interval by hiding relevant transient changes amidst other irrelevant ones. For instance, introducing a small, completely irrelevant “mud splash” to an image led to dramatic decreases in the ability to detect changes in a photographic target image.²⁰

Prior work on change blindness/change detection involved the toggling either of scenes that were identical except for the target change or toggling under conditions designed to produce strong change blindness. There has been little work on comparing SBS presentation to toggling in slightly misaligned images. Josephs et al.²¹ directly compared change detection performance when naïve observers viewed the two images in either SBS or successively in the same location. The stimuli for this study were photographs where one version of the photograph either did or did not contain a critical object. Observers were asked to detect whether there was a change as quickly and accurately as possible. They manipulated the degree of misalignment between the two photographs. When the changes were small (up to a 4% lateral shift in viewpoint), there was a large time benefit of the toggle viewing mode over the SBS mode with no difference in accuracy. This time benefit disappeared when the difference between the two images was larger. While these results were promising, it is an open question whether the changes between this year’s and last year’s exams are big enough to produce change blindness or small enough to permit toggling to produce an advantage over SBS viewing.

The current relatively small study should be considered to be a pilot assessment of radiologists’ ability to detect subtle signs of cancer in mammograms when viewing in either SBS or toggle mode. Based on the results from the previous work comparing performances in these viewing modes, we hypothesized that

mammographers would be able to perform the task more quickly with no cost, and perhaps bring a benefit in diagnostic accuracy when viewing the mammograms in toggle mode.

2 Methods

Approval from the Harvard Medical School Institutional Review Board (Protocol no. FWA0000484) was obtained to allow collection of observer data in this project. The readers in the study provided informed consent prior to participation.

2.1 Case Preparation

A senior radiologist (MAR) selected exams from a private practice employing full-field digital mammography at multiple locations, identifying four patients aged 55 to 68 (mean: 61) who had screening mammograms reported as normal previously (no findings or benign findings) and no reported findings in a subsequent exam. The prior exams were obtained an average of 24 months previously (range 12 to 54 months). The same physician then randomly selected abnormal screening mammograms that led to a biopsy proven breast cancer on eight patients (mean age: 64, range 53 to 71) whose more recent previous screening mammogram was interpreted as normal. The prior exams were obtained an average of 18 months previously (range 9 to 22 months; see Table 1). Therefore, the abnormal exams represented newly incident and presumably subtle cancers. Abnormal exams each contained a single finding: either a mass or calcifications. All cases were fully anonymized prior to use in the study. We did not control for the level of misalignment between past and present cases. As a result, the level of misalignment varied across cases, but should be representative of the variance observed in clinical practice (see Fig. 1).

2.2 Radiologist Observers

The experiment took place at two large radiology meetings (Radiological Society of North America and American Roentgen Ray Society) with observers who volunteered to participate in a brief study. The data from three participants were excluded from further analysis because they did not fully complete all 10 experimental cases. Data from one medical physicist were excluded in order to ensure all observers had received similar clinical training. The remaining 23 readers (17 ABR approved radiologists and 6 residents who had completed or were in the process of completing a mammography rotation) participated in the study. Average age of the included participants was 46 (stdev:12). The ABR-approved radiologists had an average of 16 (stdev: 12) years of experience in radiology. These radiologists estimated that they read an average of >4500 (stdev: 3900) mammograms per year. All readers who participated were offered a chance to win an Apple iPad in exchange for their participation.

2.3 Reading Environment and Experimental Procedure

The reading evaluation occurred on an Food and Drugs Administration approved DR systems (now Merge Healthcare) radiology information system (RIS)/picture archiving and communications system (PACS) system. Images were displayed on one of the two five megapixel monitors. A third monitor was used for case navigation. Display monitors were set to a monochrome Dome E5 setting. White pixels were set

Table 1 Details of the 12 cases used in the study.

Age at most recent screening	Time between exams (months)	Image type	Trial type	Cancer type
57	39	Abnormal	Practice	Speculated mass—invasive ductal CA
68	22	Abnormal	Practice	1 cm mass right breast—infiltrating ductal carcinoma
67	12	Abnormal	Experimental	Subareolar calcifications—high grade DCIS with comedonecrosis
66	12	Abnormal	Experimental	Left retroareolar calcifications—intraductal carcinoma intermediate grade
46	12	Abnormal	Experimental	Calcifications—upper outer left breast—intermediate to high grade DCIS
62	54	Abnormal	Experimental	5 mm cluster of calcifications—invasive ductal carcinoma
55	13	Abnormal	Experimental	8-mm posterior medial mass—invasive ductal carcinoma
68	24	Abnormal	Experimental	1 cm mass—infiltrating ductal carcinoma
53	21	Normal	Experimental	Normal
70	9	Normal	Experimental	Normal
71	22	Normal	Experimental	Normal
60	21	Normal	Experimental	Normal

to 500 nits (cd/m^2) and black pixels set to 0 nits. The experiment took place in a conference hall where ambient lighting was greater than is generally present in radiology reading rooms.

Once each participant had consented to participate in the study, they were told that the purpose of this study was to evaluate different methods of viewing mammograms. They were instructed to approach the experiment as if they were engaged

in screening mammography with an enriched sample of cancer. Each radiologist read 12 studies: 2 practice and 10 experimental. Each full-field digital mammography study included medial lateral oblique and craniocaudal views of both breasts. Each study included two cases: past and present. Participants were instructed to inspect each case as they would in practice. They were instructed to complete the caseload as quickly as possible without sacrificing accuracy. Once they had reached a



Fig. 1 Example of a practice case used in the study. There is a mass in the present exam. Note the misalignment between cases: (a) RML O past and (b) RML O present.

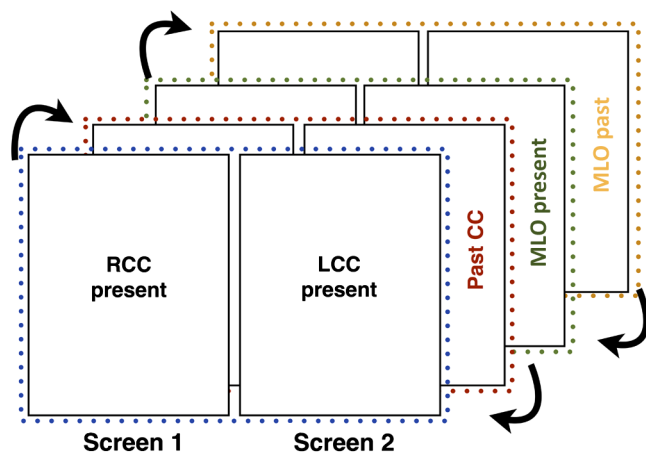


Fig. 2 Schematic illustration of the toggle viewing method. Participants used the mouse wheel to quickly “toggle” between different views.

decision, they were instructed to close the case and then rate it on a modified breast imaging-reporting and data system scale (BIRADS) that forced the radiologists to score each case between 1 (negative assessment) and 5 (highly suspicious of breast cancer). Zeros scores were eliminated to reduce ambiguity in diagnostic accuracy. In order to measure how much time was spent on each case, we took the difference between the time when the case was opened and when the case was closed. The DR system PACS software logged this information. This enabled timing precision that was accurate within 1 s.

If a lesion was noted, participants were instructed to circle the approximate (within a given quadrant of the breast) lesion location on a paper scoring sheet that was provided. If more than one suspicious lesion was identified, the participants were instructed to mark only the most suspicious. Once noted, participants labeled lesions as a mass, calcification, or both.

The experiment was divided into two halves, the order of which was counter balanced across participants. In one-half of the experiment, participants were instructed to “toggle” between past and present cases using the mouse wheel. In this viewing mode, all four views of one of the two cases are simultaneously displayed on the two high-resolution monitors. By sliding the mouse wheel, all four views “toggled” between the past or present cases. This was accomplished using a meta-file that associated each image so that past and present cases could be sequentially viewed very rapidly (see Fig. 2).

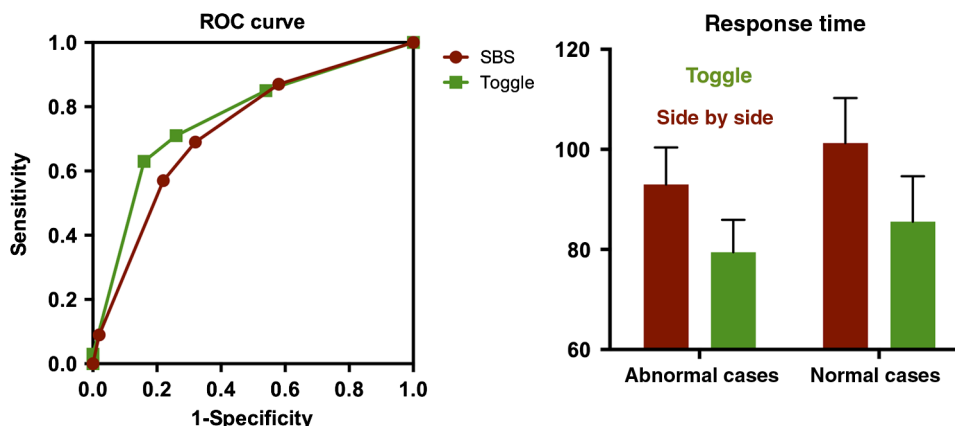


Fig. 3 Behavioral performance on the task. Error bars represent standard error of the mean.

In the other half of the experiment, participants were instructed to read the mammograms as they would normally. They were allowed to arrange views of the images in any manner they preferred, provided they did not toggle between past and present views. The hanging protocols could be quickly and easily accessed via keyboard shortcuts and included many options for presenting the new and prior exams, with multiple images per monitor or a single image per monitor. The options included the ability to automatically sort images so that new and prior images of the same view could be automatically displayed SBS.

To give the participants time to familiarize themselves with these viewing modes (particularly the toggle mode, which was novel to almost all participants), each block of trials began with a single practice case. Participants were generally much more familiar with the SBS method, based on the customary methods used in their clinical practice. The experimenter explained how to navigate through different views while the participant examined this case. Each block of five experimental cases included three cases with pathology-proven cancerous abnormalities and two cases with no abnormalities. Both blocks were preceded by a single practice case that contained an abnormality. Due to our counterbalancing procedure, half of the participants completed the toggle trials first, while the others completed the SBS trials first. We also counter balanced the viewing method associated with each case so that half of the participants viewed experimental cases 1 to 5 in the toggle mode, while the other half viewed these cases in the SBS mode. This design allowed us to compute within subject comparisons while minimizing the influence of order and case difficulty.

3 Results

Based on the previous results that compared performance in the toggle and SBS modes, the three outcome variables we analyzed were reaction time (RT), diagnostic accuracy, and area under the curve (AUC) as computed from empirical receiver operating characteristic (ROC) curves. The results are summarized in Fig. 3. We computed diagnostic accuracy based on the modified BIRADS rating. For diagnostic accuracy, ratings above 2 were marked as correct in cases that contained an abnormality. Ratings below 2 were marked as correct in cases that did not contain an abnormality. AUC was computed using the modified BIRADS scale as well. Both measures were compared using a two-tailed paired-samples *t* test.

While the participants were slightly more accurate when viewing in the toggle mode (75% accurate compared to 69%), this difference was not statistically reliable with this relatively small sample size. We also computed AUC based on the BIRADS scores. These AUC estimates should be interpreted with caution given the low number of cases per condition. The results here mirror the diagnostic accuracy results. While performance in toggle viewing was nominally higher (AUC = .78) than in SBS viewing (AUC = .72), the difference was not reliable [$t(24) = 0.32, p = 1.0$].

RT was calculated based on the time between when the case was first opened, and when the participant closed the case (prior to recording the BIRADS rating), thereby indicating that they had finished reading the case. Two trials were excluded from further analyses because, in both cases, the participant took much longer than on any other case in the experiment. While the overall average RT for the experiment was 90s/case, while the two trials treated as outliers both lasted over 4 min. Despite a small sample size and few cases per participant, we observed a significant benefit in RT for the condition [$F(1,22) = 5.11, p < .05$]. There was no effect of abnormality presence [$F(1,22) = 2.74, p = .11$] and the two factors (viewing mode and abnormality presence) did not significantly interact [$F(1,22) = 0.03, p = .85$]. Overall, when participants completed the task in toggle mode, they came to a conclusion an average of 6-s faster than in SBS mode. The same pattern is present (and is, in fact, slightly larger) if we do not exclude the abnormally long trials.

4 Discussion

Radiologists spend thousands of hours training to detect subtle abnormalities in image details that would be meaningless to the untrained eye. This skill saves lives but, even with extensive training, certification, and advanced technology, mammographers miss as many as 20% to 30% of breast cancers and many of these mistakes are caused by perceptual errors.^{5,22,23} Thus, there remains strong motivation to increase reader accuracy and efficiency.¹ While most research in breast imaging has focused on developing new technologies to improve imaging, the current pilot study adopts a very different approach, focusing instead on optimizing image presentation in an attempt to improve human perception and cognition. Based on the prior basic research outside the realm of radiology,²¹ we hypothesized that altering the viewing mode would improve the performance.

There is precedence for this sort of simple manipulation leading to substantial benefits in diagnostic radiology. Traditionally, chest computed tomography (CT) scans were examined using film-based viewing where individual slices of the chest were displayed one adjacent slice after another, SBS on the light box display. The advent of spiral CT technology allowed the entire chest to be imaged during one breath hold, reducing the issues with misalignment due to respiratory activity. Despite this great improvement in the degree of misalignment between images, most radiologists continued to view individual images separately on a light box. Seltzer et al.²⁴ compared this traditional method of viewing to “cine-viewing” where the images were organized into a stack of images that were coregistered as long as the patient did not move between slices, and the radiologists could scroll or page through the series on a computer screen. They found that cine-viewing led to higher diagnostic accuracy in a nodule detection task. This experience has carried

over to many applications of CT scanning, positron emission tomography and magnetic resonance imaging, and eventually led to the wide-spread adoption of this viewing method in clinics all over the world, for all of these modalities. In fact, the utility of this technique over the previous viewing method shares some characteristics with the toggle viewing method. In both cases, changes appear to be more easily detected when images are overlaid on top of one another as the radiologists either scrolls through slices of the CT, or toggles through mammograms.^{25,26}

The power of this demonstration was that the authors showed that changing how images were viewed provided a sizable benefit in performance. It was, therefore, relatively simple for radiologists to try this alternative method of viewing in order to determine whether they also observed superior performance. As a result, the standard of care changed, and today most physicians read cross-sectional imaging exams by paging through images that are virtually stacked.

Our study found no significant accuracy benefit of toggling as compared to SBS viewing. Our small sample size makes it difficult to determine whether this was due to a lack of any real difference in accuracy or insufficient power. However, if we hypothesize that the observed 5% advantage for toggle mode reflects a real advantage, we can calculate that it would take 180 radiologists reading 10 cases to have the statistical power to detect such a difference with 80% confidence.²⁷ Based on this calculation, the effort required for a larger study seems worthwhile, given that the potential benefit of this simple behavioral intervention could be on the same order as the reported benefit of DBT.

The toggle viewing method produced a large decrease in the amount of time spent on each case. Given the ever-increasing number of cases, and images within those cases,²⁸ this is an important benefit. Less time spent on each case could translate to a number of valuable benefits for the radiologist. It could allow the radiologist to examine more cases in the same amount of time. It could also reduce some of the constant time pressure in the clinic, allowing the radiologist to spend more time on difficult cases. Finally, given the evidence that diagnostic accuracy in radiology decreases at the end of long work days,²⁹ having additional time might allow mammographers to have shorter work days, which may lead to better performance in the long run.

What causes the observed toggle benefit in reaching a decision despite imperfect coregistration of the images? The most likely cause is a change in oculomotor behavior. In SBS viewing, radiologists must make a continuous set of saccadic eye movements back and forth between the corresponding locations in the two images. Saccadic movements are fast and frequent—3 to 4/s—but in a comparison of two complex images, there can be a real-time savings in the ability to leave the eyes in one place while examining changes in that vicinity. Confirmation of this hypothesis would require eye tracking data on our observers which we do not have for this study. The benefit may also be at least partially driven by the amount of screen space devoted to each image in the toggle viewing mode. While traditional hanging protocols (SBS) divide the screen into one, two, and sometimes four images, the mode evaluated here displayed one image per screen. This allowed each image to be larger, which may have contributed to the observed time benefit. However, screen space is not the full explanation of the toggle benefit. In our previous evaluation of toggle versus SBS viewing modes with photographs and nonradiologist observers, we

controlled for this difference and still observed a substantial toggle benefit.²¹

Prior basic research from our group also demonstrated that the toggle benefit observed is strongly related to how similar the images are to one another.²¹ The current data suggest that despite misalignment, the past and present mammograms used in this study were similar enough to produce a sizable time benefit. Based on the data from nonradiologists viewing photographs, we suspect that the benefit would be larger with better coregistration of the mammograms. We hope to test this prediction in future research. This work could also be extended to other imaging tasks in which past and present are being compared; from chest radiographs to satellite images of potential nuclear sites. In all such tasks, we believe that a critical determinant of degree of success will be the accuracy of coregistration techniques.

Our current knowledge of the human visual system suggests that the SBS method of comparison may not be optimally suited for detection of the sort of subtle abnormalities that occur in mammography. The act of moving the eyes from one item or one screen to another disrupts perceptual processing.^{30,31} Indeed, in change-blindness research, a change made during an eye movement is just as hard to detect as a change between two images separated by a blank interval.^{32–34} This research suggests that minimizing eye movements should lead to better detection.

4.1 Limitations

There were several limitations in this study. As it was designed as a pilot study, our results are based on a small number of cases with relatively few readers. As noted above, the failure to find significant differences in diagnostic accuracy and AUC could be attributed to these limitations in power (though it is also possible there was no real difference in these measures). We hope to conduct a larger follow-up study to examine whether coregistration accuracy may mediate the observed toggle benefit. Further, there is some evidence that the benefits of CAD in mammography are most pronounced for junior radiologists and are much less helpful for experts.³⁵ It would be interesting to determine whether a similar pattern is evident with the toggle benefit. Our reader population was a small, relatively heterogeneous group not suited to answering that question. A larger sample would allow us to better address this open question. Finally, there was a substantial difference in how familiar our observers were with the toggle viewing methodology. While they had all spent years viewing mammograms in the SBS mode, they were given less than 5 min of practice viewing the images in toggle mode immediately prior to the study. It is, therefore, encouraging that we observed a reliable time-benefit despite this difference. However, in order to get a more accurate assessment of the benefit of the toggle mode, it may be necessary to either conduct a more extensive training session prior to the experiment, or specifically recruit radiologists who have already adopted the toggle viewing mode.

5 Conclusions

Our data are consistent with the idea that radiologists are capable of finding differences (or of being sure there was no difference) more quickly when the critical comparisons can be made by toggling the images rather than moving the eyes from one screen to another. While the results of this pilot study should be interpreted with caution, this is a promising first step to a simple

change in how mammograms are viewed that may lead to substantial benefits for the radiologist, and most importantly the patient.

Acknowledgments

The authors thank Merge Healthcare (previously DR Systems) for providing space to conduct this research. MAR and MT are shareholders and employees of Merge Healthcare, which provided technological support for this research. This work was partially supported by grants EY017001 from NEI and N0001410278 from ONR awarded to JMW.

References

1. J. S. Mandelblatt, "Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms," *Ann. Intern. Med.* **151**(10), 738 (2009).
2. A. Bleyer and H. G. Welch, "Effect of three decades of screening mammography on breast-cancer incidence," *N. Engl. J. Med.* **367**(21), 1998–2005 (2012).
3. M. G. Marmot et al., "The benefits and harms of breast cancer screening: an independent review," *Br. J. Cancer* **108**(11), 2205–2240 (2013).
4. L. Tabár et al., "Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades," *Radiology* **260**(3), 658–663 (2011).
5. E. D. Pisano et al., "Diagnostic performance of digital versus film mammography for breast-cancer screening," *N. Engl. J. Med.* **353**(17), 1773–1783 (2005).
6. J. M. Lewin et al., "Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4, 945 paired examinations," *Radiology* **218**(3), 873–880 (2001).
7. T. W. Freer and M. J. Ulissey, "Screening mammography with computer-aided detection: prospective study of 12, 860 patients in a community breast center," *Radiology* **220**(3), 781–786 (2001).
8. L. Philpotts, "Can computer-aided detection be detrimental to mammographic interpretation?," *Radiology* **253**(1), (2009).
9. J. J. Fenton, L. Abraham, and S. H. Taplin, "Effectiveness of computer-aided detection in community mammography practice," *J. Natl. Cancer Inst.* **103**(15), 1152–1161 (2011).
10. J. J. Fenton et al., "Influence of computer-aided detection on performance of screening mammography," *N. Engl. J. Med.* **356**(14), 1399–1409 (2007).
11. V. Rao et al., "How widely is computer-aided detection used in screening and diagnostic mammography?," *J. Am. Coll. Radiol.* **7**(10), 802–805 (2010).
12. P. Skaane et al., "Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program," *Radiology* **267**(1), 47–56 (2013).
13. M. L. Zuley et al., "Digital breast tomosynthesis versus supplemental diagnostic mammographic views for evaluation of noncalcified breast lesions," *Radiology* **266**(1), 89–95 (2013).
14. "CMS Establishes Breast Tomosynthesis Values in 2015 MPFS Final Rule," <http://www.acr.org/News-Publications/News/News-Articles/2014/Economics/20141105-CMS-Establishes-Values-for-Breast-Tomosynthesis-in-2015-Final-Rule> (5 August 2015).
15. D. Gur et al., "Dose reduction in digital breast tomosynthesis (DBT) screening using synthetically reconstructed projection images," *Acad. Radiol.* **19**(2), 166–171 (2012).
16. D. J. Simons and D. T. Levin, "Change blindness," *Trends Cognit. Sci.* **1**(7), 261–267 (1997).
17. D. J. Simons and R. A. Rensink, "Change blindness: past, present, and future," *Trends Cognit. Sci.* **9**(1), 16–20 (2005).
18. C. W. Tombaugh, "The search for the ninth planet, Pluto," *Astrono. Soc. Pac., Leaflet* **73**(5) (1946).
19. K. Scott-Brown, M. Baker, and H. Orbach, "Comparison blindness," *Visual Cognit.* **7**(1–3), 253–267 (2000).
20. J. K. O'Regan, R. A. Rensink, and J. J. Clark, "Change-blindness as a result of 'mudsplashes'," *Nature* **398**(6722), 34–34 (1999).
21. E. Josephs, T. Drew, and J. Wolfe, "Shuffling your way out of change blindness," *Psychon. Bull. Rev.* (2015).

22. A. Brady, R. Ó. Laoide, and P. McCarthy, "Discrepancy and error in radiology: concepts, causes and consequences," *Ulster Med.* **81**(1), 3–9 (2012).
23. L. Berlin, "Malpractice issues in radiology. Perceptual errors," *Am. J. Roentgenol.* **167**(3), 587–590 (1996).
24. S. E. Seltzer et al., "Spiral CT of the chest: comparison of cine and film-based viewing," *Radiology* **197**(1), 73–78 (1995).
25. A. Kristjánsson, O. Sigurjonsdóttir, and J. Driver, "Fortune and reversals of fortune in visual search: reward contingencies for pop-out targets affect search efficiency and target repetition effects," *Atten. Percept. Psychophys.* **72**(5), 1229–1236 (2010).
26. J. M. Wolfe, K. R. Cave, and S. L. Franzel, "Guided search: an alternative to the feature integration model for visual search," *J. Exp. Psychol. Hum. Percept. Perform.* **15**(3), 419–433 (1989).
27. F. Faul et al., "G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences," *Behav. Res. Methods* **39**(2), 175–191 (2007).
28. K. P. Andriole et al., "Optimizing analysis, visualization, and navigation of large image data sets: one 5000-section CT scan can ruin your whole day," *Radiology* **259**(2), 346–362 (2011).
29. E. A. Krupinski et al., "Long radiology workdays reduce detection and accommodation accuracy," *J. Am. Coll. Radiol.* **7**(9), 698–704 (2010).
30. S. Liversedge and J. M. Findlay, "Saccadic eye movements and cognition," *Trends Cogn. Sci.* **4**(1), 6–14 (2000).
31. D. Irwin, "Information integration across saccadic eye movements," *Cogn. Psychol.* **23**(3), 420–456 (1991).
32. B. Bridgeman, "Failure to detect displacement of the visual world during saccadic eye movements," *Vision Res.* **15**, 719–722 (1975).
33. J. Henderson and A. Hollingworth, "The role of fixation position in detecting scene changes across saccades," *Psychol. Sci.* **10**, 438–443 (1999).
34. J. Grimes, "On the failure to detect changes in scenes across saccades," in *Perception*, K. A. Akins, Ed., pp. 89–110, Oxford University Press, New York (1996).
35. C. Balleyguier et al., "Computer-aided detection (CAD) in mammography: does it help the junior or the senior radiologist?," *Eur. J. Radiol.* **54**(1), 90–96 (2005).

Trafton Drew is an assistant professor in the Psychology Department at the University of Utah. He received his BS degree from the UNC Chapel-Hill. He received his MS and PhD degrees in cognitive psychology from the University of Oregon. His lab's (<http://aval.psych.utah.edu/>) current interests include medical image perception, eye-tracking, and the neural substrates of working memory. He is a member of the Medical Image Perception Society, Vision Sciences Society, and the Psychonomic Society.

Biographies for the other authors are not available.