



The Radiologist's Gaze: Mapping Three-Dimensional Visual Search in Computed Tomography of the Abdomen and Pelvis

Linda C. Kelahan^{1,2} · Allan Fong³ · Joseph Blumenthal³ · Swaminathan Kandaswamy⁴ · Raj M. Ratwani³ · Ross W. Filice¹

Published online: 5 October 2018
© Society for Imaging Informatics in Medicine 2018

Abstract

A radiologist's search pattern can directly influence patient management. A missed finding is a missed opportunity for intervention. Multiple studies have attempted to describe and quantify search patterns but have mainly focused on chest radiographs and chest CTs. Here, we describe and quantify the visual search patterns of 17 radiologists as they scroll through 6 CTs of the abdomen and pelvis. Search pattern tracings varied among individuals and remained relatively consistent per individual between cases. Attendings and trainees had similar eye metric statistics with respect to time to first fixation (TTFF), number of fixations in the region of interest (ROI), fixation duration in ROI, mean saccadic amplitude, or total number of fixations. Attendings had fewer numbers of fixations per second versus trainees ($p < 0.001$), suggesting efficiency due to expertise. In those cases that were accurately interpreted, TTFF was shorter ($p = 0.04$), the number of fixations per second and number of fixations in ROI were higher ($p = 0.04$, $p = 0.02$, respectively), and fixation duration in ROI was increased ($p = 0.02$). We subsequently categorized radiologists as “scanners” or “drillers” by both qualitative and quantitative methods and found no differences in accuracy with most radiologists being categorized as “drillers.” This study describes visual search patterns of radiologists in interpretation of CTs of the abdomen and pelvis to better approach future endeavors in determining the effects of manipulations such as fatigue, interruptions, and computer-aided detection.

Keywords Visual search · Eye tracking · CT · Body imaging · Drillers · Scanners

Background

A radiologist's search pattern is generally shaped by experience, usually by learning of missed findings. When a radiologist does not see a finding, in other terms, makes an error of detection,

they are forced to confront a potential “blind spot” in their visual search pattern. Discovering blind spots through detection error is one way in which radiologists learn about their visual process; however, it is not entirely clear to what extent radiologists, at baseline, understand and reflect on their visual search pattern, nor is it clear how individual search patterns differ.

Eye tracking technology has been utilized to better understand the radiologist's search pattern, most robustly in chest radiography interpretation.[1–6] A commonly utilized eye tracking metric is that of the time to first fixation, or the time duration for the subject to focus on the abnormality for the first time, which has been shown to be shorter in experts versus novices.[7, 8] Intuitively, it follows that an expert would be quicker to identify an abnormality as compared to a novice. It is thought that experience fine tunes a search pattern, and while it has been reported that senior radiology residents have search patterns closely approximating their attendings', the attendings maintain greater diagnostic accuracy, likely due to differences in cognitive processing of image findings.[6]

There is growing interest in better evaluating eye movements when examining cross-sectional imaging. Most of this

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10278-018-0121-8>) contains supplementary material, which is available to authorized users.

✉ Linda C. Kelahan
LKELAHAN2@gmail.com

¹ MedStar Georgetown University Hospital, 3800 Reservoir Road NW, Washington, DC 20007, USA

² 300 Pasteur Drive Room H1307, Stanford, CA, USA

³ MedStar Institute for Innovation, 3007 Tilden St NW, Washington, DC 20008, USA

⁴ University of Massachusetts, 120H Maraton Hall, 160 Governors Dr, Amherst, MA 01003, USA

work has utilized chest CTs and almost exclusively as it relates to the detection of lung nodules.[9–11] This has resulted in better understanding of differing visual search patterns, and also the relative success rate of different search patterns. Drew et al. have described grouping radiologists into two categories of searching behavior: “scanners” and “drillers” in regard to interpreting chest CTs. “Drillers” were described as having relatively more restricted eye movements in the two-dimensional plane while scrolling through the image, while “scanners” had a wider two-dimensional search pattern, progressing more slowly through the depth of a CT chest. “Drillers” were found to have greater accuracy than “scanners”.[9] Our goal was to map the self-directed visual search patterns of practicing radiologists, at the attending and trainee level (fellows and residents), interpreting normal and abnormal CTs of the abdomen and pelvis. Prior studies have relied on a cine loop or single-direction scrolling of stack images to assess visual search.[12–14] Our study allows radiologists to scroll through axial CT images naturally on a viewer, to better recreate the real-world experience. We used 3D eye tracking data to create a map of each individual’s search pattern viewing multiple CT cases, both with and without pathology. We examined time to first fixation (TTFF), number of fixations, number of fixations per second, number of fixations in the region of interest (ROI), total fixation duration, total fixation duration in ROI, and mean saccadic amplitude with respect to accuracy and experience (attending vs. trainee). We also attempted to categorize radiologists’ search patterns as “scanners” or “drillers” as has previously been done when searching for lung nodules on chest CTs.[9]

Methods

Our study was approved by our institutional review board (IRB). Participants were recruited by email from our institution’s radiology department. Volunteers were given a written description of the experiment and given the opportunity to ask questions. Participants gave verbal consent to a member of our research team who supervised the participant during the study. The study was set in a small, private conference room with standard room illumination. All imaging was viewed on a laptop computer. Eye movements were tracked by a Tobii (Danderyd, Sweden) eye tracking device mounted on the bottom of the laptop display. Subject’s eye movements were calibrated to the eye tracking device. Subjects viewed a total of six cases categorized as “normal,” “easy,” and “hard” with respect to the absence or presence of pathology on the CT. The level of difficulty of the presented cases was determined by consensus between a senior radiology resident (PGY-4) and radiology attending. The first case presented was always a practice case (an “easy” case of acute, uncomplicated appendicitis) and was not included in analysis. The subsequent five

cases were randomized for each subject and included: two “normal” cases, a “hard” case of a right adrenal nodule, a “hard” case of a lipoma in a loop of the jejunum, and an “easy” case of diverticulitis. In those cases that were not normal, the site of pathology on CT was outlined as a volumetric ROI on the viewer for later analysis (i.e., a ROI square drawn around the jejunal lipoma). This region of interest marking was not visible to the subjects.

Eye movements were mapped to the CT images by synchronization with slice number by time stamp. Subjects were not given a time limit and recorded any and all findings on a worksheet. Participants were instructed to only record their findings once they were finished viewing the case to reduce head motion and eye deviation away from the screen. Once all cases had been viewed, the participants also completed a short questionnaire about their demographics and search pattern practices.

The subject’s responses were scored for accuracy in diagnosis. Essentially, the main diagnosis (diverticulitis, jejunal lipoma, or adrenal adenoma) was accepted as correct and, if it was not reported, this was considered incorrect, even if the radiologist had correctly noted another less salient finding that was present (such as an anatomic variant). Normal cases were scored as correct unless the radiologist describe a false positive finding—a finding that was not reported on the accompanying radiology report and was also not present on secondary review of the images by a senior radiology resident (PGY-5).

Once visual data was acquired, we developed logistic regression models with dependent variable as accuracy (two levels: correct, incorrect) or experience (two levels: attending, resident) and independent variables as TTFF, number of fixations, number of fixations per second, number of fixations in ROI, total fixation duration, total fixation duration in ROI, and mean saccadic amplitude. We used participants as a blocking factor to determine which of the independent variables were statistically significant. We defined fixations based on the I-VT algorithm as described by Salvucci and Goldberg and used a threshold of 20°/s.[15]

Drew et al. have described grouping radiologists into “scanners” and “drillers” in regard to their visual search pattern when interpreting chest CTs.⁹ We applied their classification techniques to our participants in two ways. The first was a modification of Drew et al.’s technique by a qualitative assessment: plotting eye movements on a depth versus time chart and analyzing the curve by consensus between two of the authors.[9] When examining depth versus time curves to subjectively classify a participant as a “scanner,” these curves demonstrated relatively less change in depth over time (Fig. 1) with curves resembling hills. Conversely, “drillers” would cover more depth in a shorter amount of time giving the appearance of a seismic trace (Fig. 2). The quantitative method we used was also modified from Drew et al.⁹ We plotted the cluster fixation per second versus the mean saccadic amplitude for each trial and ranked the aggregate eye movements.

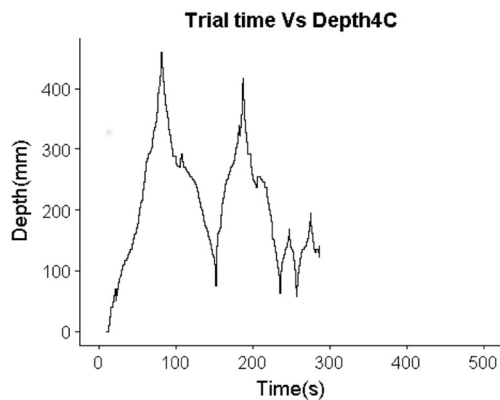


Fig. 1 “Scanner” depth vs. time curve. An example of a “Scanner” visual search pattern by qualitative assessment on a depth versus time curve. More time is spent at a given depth suggesting search focused in the x-y plane

Results

Of the 19 radiology trainee and attending participants, one was excluded from analysis after failing to record diagnoses for all cases and a second participant was excluded for inadvertently viewing a practice case that was different from the other subjects. In our analysis, fellow and resident data were combined as both were considered trainees. Nine attendings and 8 radiology trainees examined a total of 85 cases. Fifteen of these cases were excluded from analysis due to > 25% eye tracking data loss with 70 cases remaining for data analysis. The age of participants ranged from 29 to 61 years old. Subspecialties of the participating attendings included body imaging, women’s imaging, breast imaging, and neuroradiology. We found that participants’ overall accuracy was relatively consistent with the assigned level of difficulty of each case with fewer participants correctly diagnosing difficult cases (Table 1). There was no significant difference in accuracy between trainees and attendings ($p = 0.18$).

In analyzing eye movement data for all participants, we found the TTFF was shorter for correct responses 77.6 s vs. 131 s ($p =$

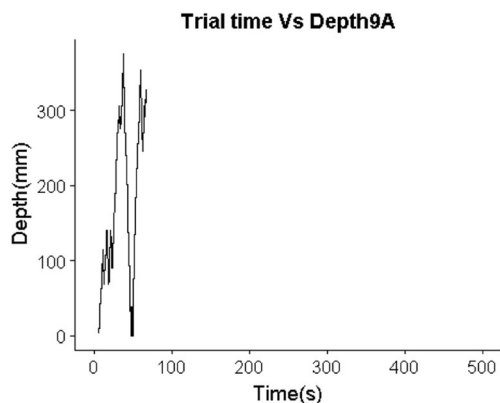


Fig. 2 “Driller” depth vs. time curve. An example of a “Driller” visual search pattern by qualitative assessment on a depth versus time curve. The changes in depth are more dramatic over a shorter time frame

Table 1 Overall accuracy by case and level of difficulty

Case	Level of difficulty	Participants reporting correctly
Appendicitis	Easy	Practice case, not scored
Diverticulitis	Easy	13 (76%)
Normal B	–	13 (76%)
Right adrenal nodule	Hard	4 (24%)
Normal F	–	11 (64%)
Jejunal lipoma	Hard	4 (24%)

0.04) and did not differ by experience level (attendings 92.1 s; residents 99.5 s; $p = 0.74$). The number of fixations per second was higher in correct responses (2.37 vs. 2.05; $p = 0.04$) with attendings having significantly fewer fixations per second than residents (1.89 vs. 2.59; $p < 0.001$). The number of fixations in ROI was higher in correct versus incorrect responses (6.15 vs. 0.9; $p = 0.02$) without significant differences between attendings (3.3) and residents (3.9) ($p = 0.76$). Fixation duration in ROI was likewise increased in correct responses (3.44 s vs. 0.78 s; $p = 0.02$) with no significant differences between attendings (2.10 s) and residents (2.16 s) ($p = 0.95$). There were no significant differences in accuracy or experience in measurement of mean saccadic amplitude (accuracy: $p = 0.60$; experience: $p = 0.24$) or total number of fixations (accuracy: $p = 0.83$; experience: $p = 0.329$).

We subsequently attempted to classify radiologists as “scanners” or “drillers” using both qualitative measures and quantitative measures.⁹ In qualitatively examining the depth versus time curves by consensus, we determined that 13 participants were “drillers” (participants 3, 5, 6, 8, 11–19) and 4 were “scanners” (participants 2, 4, 9, 10). Depth vs. time curves appeared similar between cases for each individual radiologist (supplementary material). Utilizing the quantitative method of sorting by plotting the cluster fixation per second versus the mean saccadic amplitude for each trial and ranking the aggregate eye movements, we found two radiologists were “scanners” [8, 9] and the remainder were “drillers.” Finally, when we performed a rank eye index, we found inconsistencies between our quantitative (Fig. 3) and qualitative assessments (Fig. 4). We found no significant differences in accuracy between using our quantitative sorting method (Fig. 5) ($p = 0.6312$ (Pearson’s chi-squared test)) and qualitative sorting method (Fig. 6) ($p = 1$ (Pearson’s chi-squared test)).

Discussion

Interestingly, the only significant difference in eye tracking metric between attendings and trainees was the number of fixations per second. This is likely explained by the increased efficiency of expertise and experience described by Manning et al. as decreased mean scrutiny time per film.[16] Experts have fewer

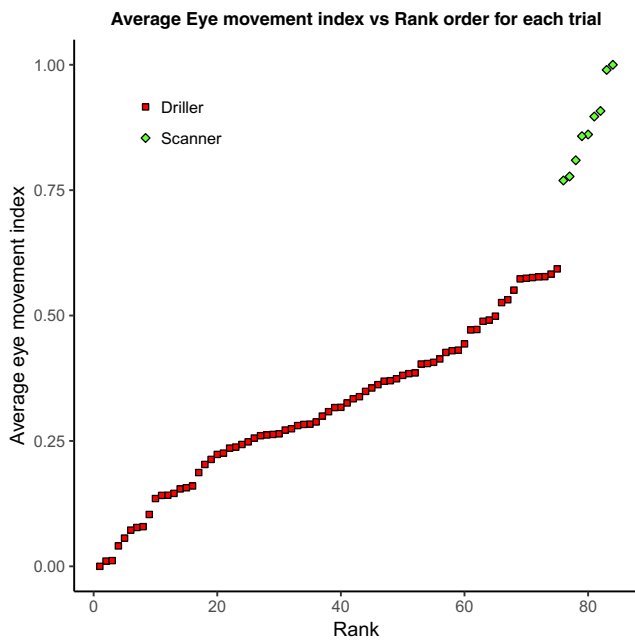


Fig. 3 Average eye movement index vs. rank order per trial based on quantitative assessment. The categorization by mean saccadic amplitude per trial was plotted demonstrating the separation between “scanners” (green diamonds) and “drillers” (red squares)

number of fixations overall when searching images and before encountering a region of interest, suggesting their search patterns are more focused than novices who presumably have a more broad and random search.[6, 16, 17] Otherwise, our lack of significant differences in eye tracking metrics between

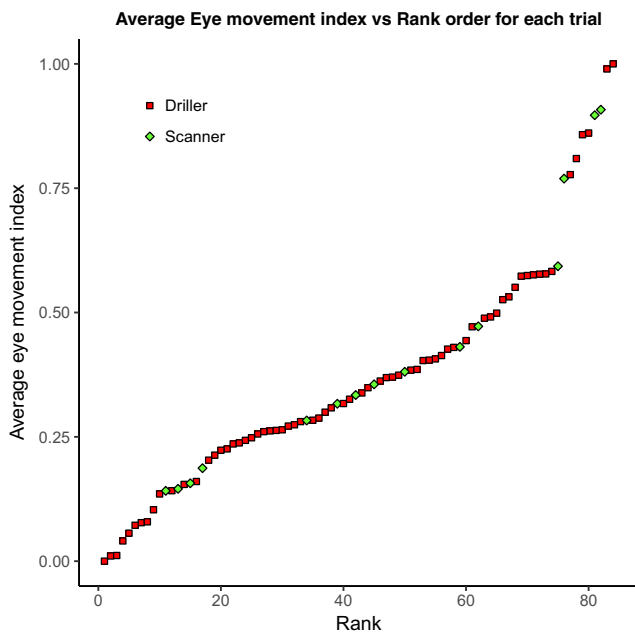


Fig. 4 Average eye movement index vs. rank order per trial based on qualitative assessment. The categorization of each depth vs. time curve per trial was plotted demonstrating significant overlap in “scanners” (green diamonds) versus “drillers” (red squares)

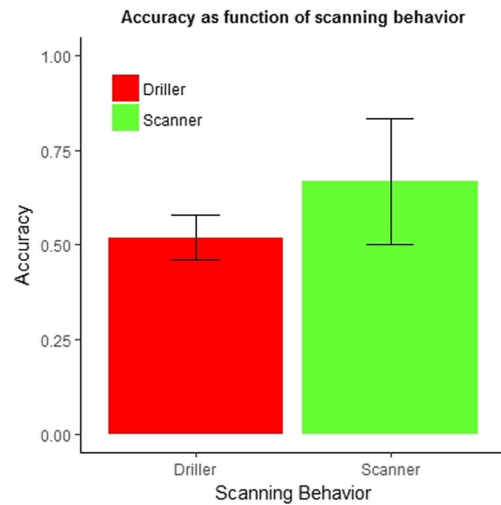


Fig. 5 Accuracy between “drillers and “scanners” based on quantitative assessment. There was no significant difference in accuracy between “scanners” and “drillers”

attendings and trainees with respect to TTFF, number of fixations in ROI, fixation duration, mean saccadic amplitude, and total number of fixations is in line with Kelly et al. who discovered no significant differences in eye tracking metrics between trainees and attendings.[6] This may also account for the similar rate of accuracy between trainees and attendings in our study. Potentially, differences in eye tracking metrics may have been more dramatic in naïve individuals without radiology experience, and this was likely a limitation in our study.

Our study found no difference in mean saccadic amplitude with respect to accuracy. Decreased saccadic amplitude is associated with “drilling” behavior and demonstrated elsewhere to result in greater accuracy in pulmonary nodule detection.[9] Bertram et al. found that experts (radiologists) had decreased saccadic amplitude when viewing CT images of the abdomen and pelvis as compared to non-experts (psychology students and

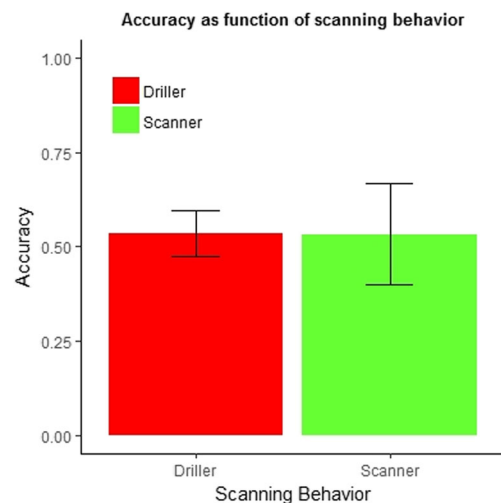


Fig. 6 Accuracy between “drillers and “scanners” based on qualitative assessment. There was no significant difference in accuracy between “scanners” and “drillers”

CT technologists). In fact, radiologists appeared to uniformly decrease their saccadic amplitude in the presence of abnormal imaging findings (in their case, abnormally enlarged lymph nodes), suggesting that experts need only examine a smaller area to come to the correct conclusion as compared to naïve participants [14]. Drew et al. similarly found that “scanners” tended to be radiologists that did not routinely read chest CTs in practice, suggesting that those with more experience reading chest CTs tended to be “drillers” either because they had found this style to be more successful in practice and had adapted it or because the more experienced radiologists tended to be “drillers” skewing the results in favor of experienced versus the predominantly more inexperienced “scanners” [9].

The number of fixations and fixation duration in the ROI were all increased in accurately interpreted case. This is likely explained by the fact that the participant found the ROI salient enough to report, and, therefore, warranted greater attention.

Using both quantitative and qualitative methods to sort radiologists into “scanners” or “drillers,” we found no significant correlation with accuracy, even after power analysis, with our quantitative method of sorting ($p = 0.6312$ (Pearson’s chi-squared test)) (Fig. 5) or with our qualitative method of sorting ($p = 0.1$) (Fig. 6).

While there is evidence to suggest that “drilling” is advantageous when distinguishing round nodules from linear pulmonary vessels given their different morphology and the tendency for spherical pulmonary nodules to “pop” out in the z-plane greater than the x-y plane [9], the lesions and pathology encountered in abdominal imaging are more varied, irregular in shape, and typically low contrast with respect to their background (i.e., soft tissue lesions in soft tissue organs versus soft tissue nodules in the air-filled lung). Also, there are more organs in the abdomen and pelvis than in the thorax, and a pathologic process affecting one organ can predispose to involvement of another, unlike a pulmonary nodule, which is less often related to other organs in the chest (such as the heart). Further, instead of the single goal of identifying pulmonary nodules, radiologists are often looking for multiple possible imaging findings to explain a patient’s clinical symptoms, and frequently balancing salient but unimportant findings (abdominal wall injection site granulomas) with more subtle but critical findings (iliac vein thrombus). These complicating factors favor a more creative approach to visual search representation. Our attempts to characterize radiologists as “scanners” versus “drillers” suggest that the majority of radiologists take a “drilling” approach to interpretation of CTs of the abdomen and pelvis. However, this may be an imperfect model to apply to CTs of the abdomen and pelvis given the inconsistencies in categorization between qualitative and quantitative approaches. This may be explained by differences in image analysis when calculating average cluster fixations: the axial plane needs to be sectioned into quadrants, which may not be meaningful given that the organs are not all sectioned into quadrants and they occupy a significant amount of space in the z-axis. Ideally,

examining the accuracy of detection of lesions in a single organ (such as the liver) using a “drilling” or “scanning” pattern may be a future direction for research.

By both our methods used to classify radiologists, the majority of participants were characterized as “drillers.” Identifying “scanners” proved to be more inconsistent as only one participant was characterized as a scanner by at least two methods (participant 9). Potentially, utilizing 3D gaze maps to illustrate visual search patterns can be more effective not only in describing individual visual search pattern, but also highlighting areas that were not examined (Fig. 7). These 3D gaze maps also illustrate the differences in eye movements between “drillers” and “scanners” more saliently. The “scanner” participant (Fig. 7, right panel) spends the most time moving horizontally through the individual slices, with much longer fixations than the “driller” (Fig. 7, left panel). The tracing resembles a meticulous slicing through the CT as opposed to the briefer, more chaotically spaced fixation points of the “driller.” In addition, the “driller” appears to have more space devoid of fixations in the center of the abdomen, whereas the “scanner” has covered the central abdomen more thoroughly.

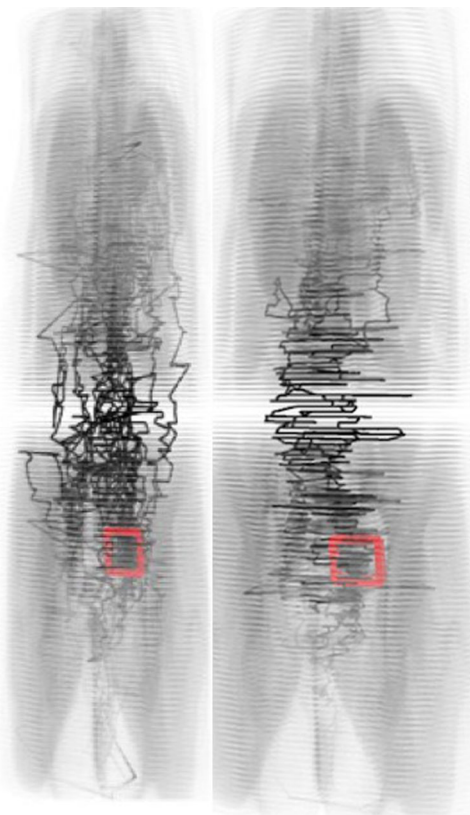


Fig. 7 Differences between search patterns of “drillers” vs. “scanners” in a case of acute diverticulitis. Left panel: 3D Gaze Map in three planes of a “Driller”. Participant’s eye movement tracing superimposed on a CT of the abdomen and pelvis with ROI in the left lower quadrant (red square). Right panel: Gaze Map in three planes of a “Scanner.” Participant’s eye movement tracing superimposed on a CT of the abdomen and pelvis with ROI in the left lower quadrant (red square)

Our study had several limitations, some of which we elicited from participant survey responses. Our viewer was designed to mimic a PACS system and provide user control over navigation through the images with a mouse scroll wheel, similar to a reading room experience, rather than relying on previous methods of cine clips [14, 18], navigating with keyboard arrows [9] or limited, unidirectional scrolling.[8] However, participants reported that the inability to window/level the images or view orthogonal reformatted images, as well as have to perform the task in ambient lighting brighter than their normal work environment, all detracted from a “realistic” reading room experience. We acknowledge that the aforementioned factors are potential limitations of our study in that they may have influenced the participant’s searching behavior; however, our viewer posed no greater limitations than those in other studies and was arguably more advanced in allowing navigation by self-directed scrolling.

The inherent intent of our study to observe radiologists’ search pattern could have contributed to a “Hawthorne effect”[19]. In other words, since the participants were aware that their eye movements were being tracked, they may have more closely adhered to their search patterns than when in practice. Our hope in presenting six CTs of the abdomen and pelvis was to introduce some level of fatigue by the fifth case to overcome such behavior. In response to the question “Did you adhere to your typical search pattern while interpreting these CTs?” 10 out of 17 respondents answered “yes for all” while the remaining participants answered “mostly.” In response to the question, “Will your search pattern change once you find an abnormality on a CT of the abdomen and pelvis?” one participant responded “always,” 6 out of 17 responded “most of the time,” 8 out of 17 responded “not very often,” and 2 out of 17 responded with “never.” In a follow-up question as to how their search pattern would change after an abnormality was encountered, most participants answered that they would evaluate the area of abnormality more closely and evaluate pertinent positives and negatives. Given the similarities of individual radiologists’ depth vs. time curves between cases (supplementary material), it appears that radiologists adhered to their search pattern regardless of whether an abnormality was present, as most claimed to on their questionnaire.

Several of the CT cases, in hindsight, had incidental findings beyond the key diagnosis and even in those cases labeled “normal.” For example, in the “hard” case of a subtle right adrenal nodule, there were multiple incidental findings that participants reported and that were subsequently verified, including: atherosclerosis, a uterine fibroid, diverticulosis, a retroaortic left renal vein, a distended gallbladder, a punctate liver lesion, mild distal small bowel wall thickening and submucosal fat deposition, left ovarian cyst, and trace fluid. Interestingly, there were also false positive findings reported in the “normal” cases (Table 2). The presence of additional true and false positive findings in “normal” cases highlights the need for better

Table 2 False positives reported in “normal” cases

Normal B	Pancreatic head malignancy, dilated small bowel loops in the left upper quadrant, mildly prominent pancreas size, asymmetric gastric antrum wall thickening, prominent vessels in the duodenum
Normal F	Small bowel wall thickening, small bowel dilatation, terminal ileitis, possible appendicitis, left adrenal nodule, pelvic lymphadenopathy

understanding into false perception and the potential value for double readers in radiology studies. In light of these multiple findings, compounded with an artificial viewing environment, it is not surprising that the majority of subjects missed such a subtle finding. Our questionnaire did specifically ask participants if they stopped searching the images after making a finding and all participants uniformly responded “No,” suggesting that participants searched the entire study as instructed.

Conclusion

Our data demonstrates that radiologists have increased number of fixations, and fixation duration in correctly interpreted cases, as well as shorter time to first fixation. We also found that attendings had fewer fixations per second than residents, highlighting the role of expertise. However, attendings and trainees were otherwise similar in most eye metric analysis.

We further discovered that previously defined search behavior in chest radiographs and chest CTs may not be adequate to describe the complexity of the visual search process in the abdomen and pelvis. Future avenues of research include establishing coverage or heat maps for individual radiologists to determine blind spots, as well as studying the role of interruptions, fatigue, and computer-aided detection (CAD) on visual search pattern adherence.

References

1. Samuel S, Kundel HL, Nodine CF, Toto LC: Mechanism of satisfaction of search: Eye position recordings in the reading of chest radiographs. *Radiology*. 194(3):895–902, 1995. <https://doi.org/10.1148/radiology.194.3.7862998>
2. Kundel HL, Nodine CF, Krupinski EA: Searching for lung nodules: Visual dwell indicates locations of false-positive and false-negative decisions. *Invest Radiol*. 24(6):472–478, 1989. <https://doi.org/10.1097/00004424-198906000-00012>
3. Berbaum KS, Franken EA, Dorfman DD et al.: Role of faulty visual search in the satisfaction of search effect in chest radiography. *Acad Radiol*. 5(1):9–19, 1998. <http://www.ncbi.nlm.nih.gov/pubmed/9442202>. Accessed March 21, 2017
4. Kundel, H. L., Nodine, C. F., & Carmody, D. Visual scanning, pattern recognition and decision-making in pulmonary nodule

- detection. *Investigative Radiology*, 13(3), 175–181, 1978. <http://doi.org/10.1097/00004424-197805000-00001>
5. Drew T, Evans K, Vö ML-H, Jacobson FL, Wolfe JM: Informatics in radiology: What can you see in a single glance and how might this guide visual search in medical images? *RadioGraphics*. 33(1): 263–274, 2013. <https://doi.org/10.1148/rg.331125023>
 6. Kelly BS, Rainford LA, Darcy SP, Kavanagh EC, Toomey RJ: The development of expertise in radiology: In chest radiograph interpretation, “expert” search pattern may predate “expert” levels of diagnostic accuracy for pneumothorax identification. *Radiology*. 280(1):252–260, 2016. <https://doi.org/10.1148/radiol.2016150409>
 7. Kundel HL, Nodine CF, Conant EF, Weinstein SP: Holistic component of image perception in mammogram interpretation: Gaze-tracking study. *Radiology*. 242(2):396–402, 2007. <https://doi.org/10.1148/radiol.2422051997>
 8. Cooper, L., Gale, A., Darker, I., Toms, A., & Saada, J. (2009). Radiology image perception and observer performance: How does expertise and clinical information alter interpretation? Stroke detection explored through eye-tracking (p. 72630K), 2009. <http://doi.org/10.1117/12.811098>.
 9. Drew T, Vo ML-H, Olwal A, Jacobson F, Seltzer SE, Wolfe JM: Scanners and drillers: Characterizing expert visual search through volumetric images. *J Vis*. 13(10), 2013. <https://doi.org/10.1167/13.10.3>
 10. Williams LH, Drew T: Distraction in diagnostic radiology: How is search through volumetric medical images affected by interruptions? *Cogn Res Princ Implic*. 2(1):12, 2017. <https://doi.org/10.1186/s41235-017-0050-y>
 11. Rubin GD, Roos JE, Tall M, Harrawood B, Bag S, Ly DL, Seaman DM, Hurwitz LM, Napel S, Roy Choudhury K: Characterizing search, recognition, and decision in the detection of lung nodules on CT scans: Elucidation with eye tracking. *Radiology*. 274(1): 276–286, 2015. <https://doi.org/10.1148/radiol.14132918>
 12. Helbren E, Halligan S, Phillips P, Boone D, Fanshawe TR, Taylor SA, Manning D, Gale A, Altman DG, Mallett S: Towards a framework for analysis of eye-tracking studies in the three dimensional environment: A study of visual search by experienced readers of endoluminal CT colonography. *Br J Radiol*. 87(1037):20130614, 2014. <https://doi.org/10.1259/bjr.20130614>
 13. Ratwani RM, Wang E, Fong A, Cooper CJ: A human factors approach to understanding the types and sources of interruptions in radiology reading rooms. *J Am Coll Radiol*. 13(9):1102–1105, 2016. <https://doi.org/10.1016/j.jacr.2016.02.017>
 14. Bertram R, Helle L, Kaakinen JK, Svedström E. The effect of expertise on eye movement behaviour in medical image perception. Paterson K, ed. *PLoS One*. 8(6):e66169, 2013. doi: <https://doi.org/10.1371/journal.pone.0066169>.
 15. Salvucci DD, Goldberg JH. Proceedings : Eye Tracking Research & Applications Symposium 2000 : Palm Beach Gardens, FL, November 6–8, 2000. In: Association for Computing Machinery; 2000:71–78. <https://dl.acm-org.ezproxy.galter.northwestern.edu/citation.cfm?id=355017>. Accessed August 15, 2018.
 16. Manning D, Ethell S, Donovan T, Crawford T: How do radiologists do it? The influence of experience and training on searching for chest nodules. *Radiography*. 12(2):134–142, 2006. <https://doi.org/10.1016/j.radi.2005.02.003>
 17. Giovinco NA, Sutton SM, Miller JD, Rankin TM, Gonzalez GW, Najafi B, Armstrong DG: A passing glance? Differences in eye tracking and gaze patterns between trainees and experts reading plain film bunion radiographs. *J Foot Ankle Surg*. 54(3):382–391, 2015. <https://doi.org/10.1053/j.jfas.2014.08.013>
 18. Ellis SM, Hu X, Dempere-Marco L, Yang GZ, Wells AU, Hansell DM: Thin-section CT of the lungs: Eye-tracking analysis of the visual approach to reading tiled and stacked display formats. *Eur J Radiol*. 59(2):257–264, 2006. <https://doi.org/10.1016/j.ejrad.2006.05.006>
 19. Mayo E. Hawthorne and the Western Electric Company First Phase-The Test Room. In: *The Social Problems of an Industrial*. ; 161–182, 1949. doi: <https://doi.org/10.1017/CBO9781107415324.004>.