

InformatICS in Radiology

What Can You See in a Single Glance and How Might This Guide Visual Search in Medical Images?¹

TEACHING POINTS

See last page

Trafton Drew, PhD • Karla Evans, PhD • Melissa L.-H. Võ, PhD • Francine L. Jacobson, MD, MPH • Jeremy M. Wolfe, PhD

Diagnostic accuracy for radiologists is above that expected by chance when they are exposed to a chest radiograph for only one-fifth of a second, a period too brief for more than a single voluntary eye movement. How do radiologists glean information from a first glance at an image? It is thought that this expert impression of the gestalt of an image is related to the everyday, immediate visual understanding of the gist of a scene. Several high-speed mechanisms guide our search of complex images. Guidance by basic features (such as color) requires no learning, whereas guidance by complex scene properties is learned. It is probable that both hardwired guidance by basic features and learned guidance by scene structure become part of radiologists' expertise. Search in scenes may be best explained by a two-pathway model: Object recognition is performed via a selective pathway in which candidate targets must be individually selected for recognition. A second, nonselective pathway extracts information from global or statistical information without selecting specific objects. An appreciation of the role of nonselective processing may be particularly useful for understanding what separates novice from expert radiologists and could help establish new methods of physician training based on medical image perception.

©RSNA, 2012 • radiographics.rsna.org

RadioGraphics 2013; 33:263–274 • Published online 10.1148/rg.331125023 • Content Codes: **ED** **IN**

¹From the Visual Attention Laboratory, Department of Surgery, Brigham and Women's Hospital, 64 Sidney St, Suite 170, Cambridge, MA 02139-4170 (T.D., K.E., M.L.H.V., F.L.J., J.M.W.); and Harvard Medical School, Boston, Mass (T.D., K.E., M.L.H.V., J.M.W.). Presented as an education exhibit at the 2011 RSNA Annual Meeting. Received February 29, 2012; revision requested May 7; final revision received August 17; accepted August 21. Supported in part by research funding from Toshiba. J.M.W. has disclosed financial relationships (see p 000); all other authors have no financial relationships to disclose. Address correspondence to T.D. (e-mail: traftondrew@gmail.com).

Funding: The work was supported in part by the National Institutes of Health [grant numbers 1F32EB011959-01, 1F32EY019819-01, ONR MURI N000141010278].

Introduction

Searching images for significant clinical findings is a common task for radiologists. Given the difficulty of this task, it is not surprising that mistakes are made even though errors can be very costly. If we think casually about the process of searching for something, whether a pencil or a malignant tumor, we probably imagine a deliberate process of selecting one location after another for analysis until the target of the search is found or the searcher decides the target is not present. However, expert radiologists will often report the sensation of “knowing” that a particular image contains a lesion before they are able to locate it. The assumption is that there is information in the first look at the image, and that the “gestalt” impression (1) is a useful part of the expert search process.

Kundel and Nodine (2) began the experimental analysis of this phenomenon in radiology by measuring the diagnostic accuracy of radiologists who were allowed only a 200-msec glimpse of a chest radiograph. They found that radiologists were able to correctly detect 70% of lesions in chest radiographs that they viewed for only one-fifth of a second. Of course, nobody is suggesting that radiologists should make decisions based on a single glimpse. However, the fact that diagnostic accuracy far exceeded that expected by chance confirmed that radiologists are capable of extracting valuable information from an image without having enough time to carefully examine it.

We consider this ability to be a specialized use of global visual processes available to all visually normal humans. Radiologists are expert medical image searchers who spend thousands of hours refining their search abilities. In a related sense, given our lifelong experience moving about the world from one visual scene to another, we are all expert scene processors. A growing body of research has shown that, just as radiologists may be able to categorize a chest radiograph in the blink of an eye, humans in general can categorize real scenes (eg, as natural or man-made) after exposures as short as 0.025 seconds (3–7).

Of course, while there may be information in a single glimpse of a scene, be it a vacation snapshot or a radiograph, that glimpse is generally not sufficient for us to find or identify specific targets. Our group has recently suggested that visual search and visual awareness of complex images may best be explained by a two-pathway process (8) (Fig 1).

Converging evidence suggests that a *nonselective* pathway is the driving force behind the ability to successfully perform a variety of visual tasks after brief exposure to a scene. This pathway appears to extract global or statistical information from across the visual field, allowing the observer to quickly extract the mean size (9), orientation (10), or direction of motion (11) of an entire image. The nonselective pathway does not support a fine differential diagnosis nor indicate if the scissors are in the kitchen drawer, but its analysis of the visual statistics of the world will rapidly provide an assessment of the scene before you. For a novice, analysis of these statistics would identify an image as a radiograph after a fraction of a second's exposure. For an expert, analysis of these statistics might support a fairly accurate (above that expected by chance) assessment of whether a patient requires further evaluation.

More specifically, object identification relies on a *selective* path, so named because candidate objects must be individually selected and processed to be identified. It is simply not possible, from a computational point of view, to recognize all the objects in the visual world at the same time (12). Under attentional control, the selective pathway processes object after object at a maximum rate of approximately 20–40 objects per second. For less readily identifiable “objects,” such as ambiguous findings on a radiograph, the time per object will be longer, but identification of those objects will still be accomplished by the selective pathway.

The fields of medical image perception (13) and basic vision science seek to advance our understanding of the mechanisms that underlie our perception of visual information. In this article, we briefly summarize research from both literatures, highlighting the role that nonselective

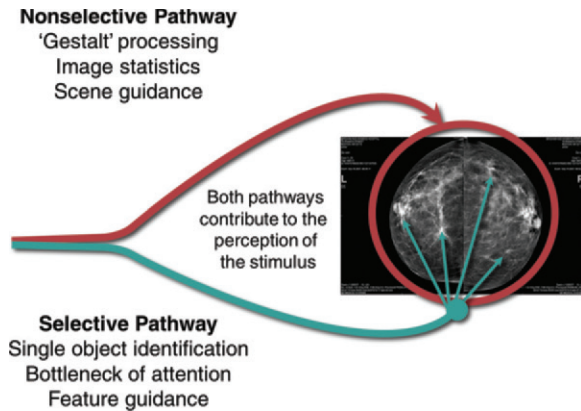


Figure 1. Two-pathway architecture for visual processing. Diagram and medical image show the nonselective (red arrow, red circle) and selective (blue line, blue arrows) pathways of visual search. The selective pathway can combine features and recognize objects. However, it is capacity limited; it is difficult for it to process more than one item or location at a time. We posit that searches in everyday scenes and medical images are guided by the nonselective pathway, which can extract statistics from the entire scene or image. This global information can be used to categorize scenes (eg, as natural or man-made) and their contents (eg, an animal or a tumor). Nonselective processing does not support precise object recognition.

tive mechanisms play in the understanding of medical images. Specific topics discussed are the nature of guidance in visual search, forms of guidance in visual search, scene guidance in radiologic search, eye movements in medical image perception, development of expertise in nonselective search, and training for expertise in nonselective search.

An important implication of the existence of a nonselective pathway is that even the briefest of glances at an image may contain valuable information that might be exploited in the development of teaching tools or more effective computer-aided detection algorithms. Established radiologists may also find it useful to better appreciate the role of nonselective processing in their interpretation of images.

Nature of Guidance in Visual Search

Visual search is the subject of a vast, ever-expanding literature in cognitive psychology (14,15). In the laboratory, visual search tasks typically involve searching some well-specified area for a target. Almost all of this enormous body of work involves nonexpert observers (typically, undergraduate psychology students) searching a computer screen for a simple target among simple nontargets or distractors. For example, an observer might be asked to search for a single *T* target among a number of *L* distractors. The target would typically be present in 50% of the trials.

Much of the visual search literature is concerned with understanding the factors that influence the ease and speed with which targets are found in these types of experiments. For instance, it is easy to find a red *T* among green *L*s because the unique color of the target draws attention to it. On the other hand, presenting *T*s and *L*s of the same color in arbitrary orientations increases the difficulty of the task. Experimenters typically take great pains to remove any influence of observers' experience with and knowledge about the display. The goal is to place all observers on an equal footing so researchers can evaluate universal basic building blocks that underlie our ability to find things in the real world.

While this body of research is the bedrock of the visual search literature in visual science, it neglects important aspects of the searches that radiologists undertake every day. These aspects include the years of training that radiologists receive and the dramatically increased complexity and ambiguity of medical images. Moreover, until recently, the basic search literature had typically used arrays of randomly scattered items for stimuli. This stimulus design ignored the fact that real-world scenes and medical images are structured. It has become increasingly clear that any account of visual search in scenes, be they natural scenes or medical images, will require an account of how the scene structure information can guide the search (Fig 2).

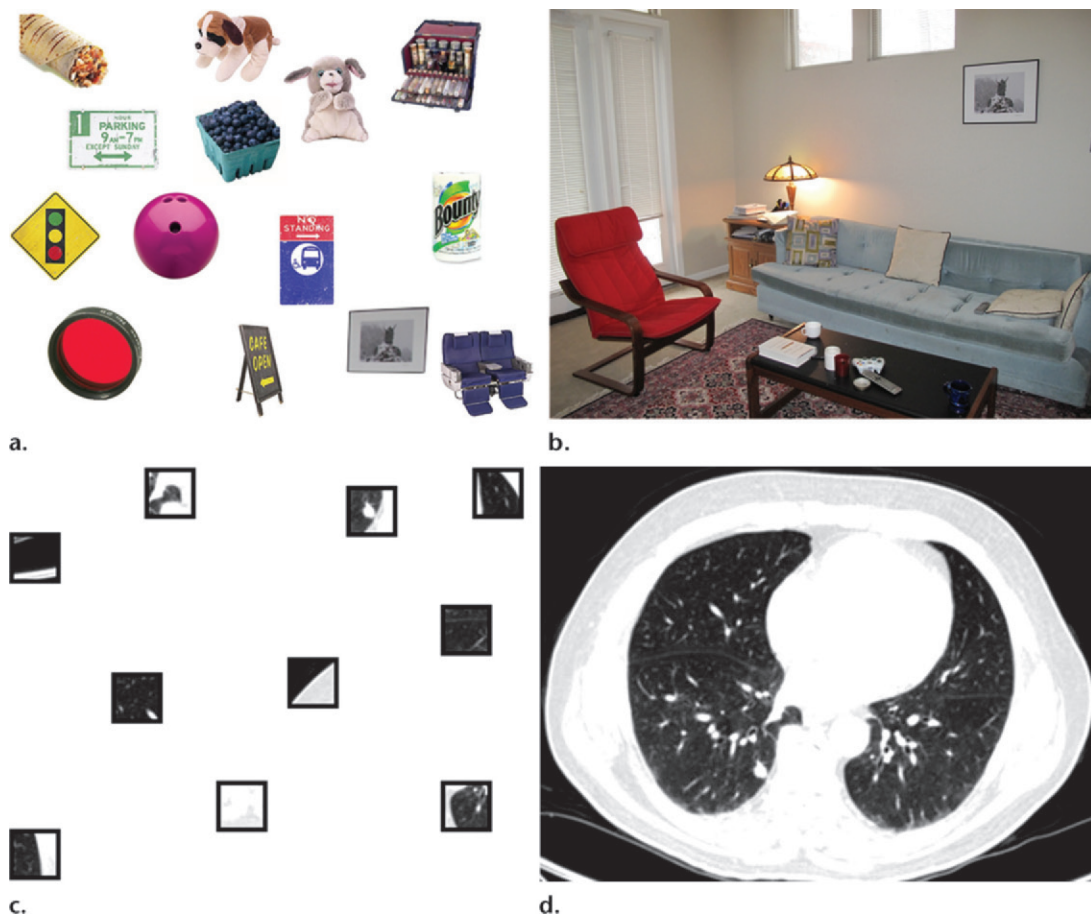


Figure 2. Visual search of random arrays versus real scenes. **(a, b)** Search for a framed picture. **(a)** In classic visual search experiments, random arrays of items are presented, deliberately eliminating the possibility of guidance of search by scene context or scene structure. **(b)** In real scenes, context and structure drastically constrain where we search. In this case, search will start on the walls, where pictures most commonly hang. **(c, d)** Search for a lung nodule. **(c)** If the background and spatial layout of a medical image are removed, search becomes much more difficult. **(d)** With a real medical image, nonselective scene processing will help expert radiologists find lung nodules quickly and with fewer extraneous eye movements.

In the remainder of this article, we outline evidence that supports the importance of what we refer to as nonselective, scene-based information during medical image search.

Forms of Guidance in Visual Search

Before becoming board certified, a radiologist undergoes years of intensive training that involves reading many thousands of images. One important outcome of this training is to refine where radiologists look for targets. Recall that, in most laborato-

ry visual search experiments, the observers would have no idea where in the image to look for their target (Fig 2a). The situation is not so extreme for novice radiologists, as they begin their training with an understanding of the basic anatomy in most radiologic images. **However, an important consequence of training is to teach radiologist trainees that some areas are more likely to contain a lesion than others. Thus, while eye movement recordings suggest that novice radiologists search in a relatively haphazard fashion when looking for lung nodules, experienced radiologists tend to exhibit more concise eye movements, with fewer fixations needed to extract more information (16).**

Teaching Point

Evidence from a laboratory paradigm known as “contextual cueing” (17) suggests that aspects of this sort of expertise can develop outside of the searcher’s awareness. Chun and Jiang (17) had naive observers repeatedly search through artificially created displays similar to that in Figure 2a. Unbeknownst to the observers, some of the displays were repeated several times. Observers responded faster to the repeated displays, even though, when tested later, they were unable to categorize displays as repeated or not. In other words, observers are able to find targets more quickly in scenes that they have searched previously.

Of course, radiologic expertise is more complex, but it seems likely that some learning of the regularities of medical “scenes” will occur outside of the radiologist’s explicit awareness. Furthermore, an understanding of these regularities could help explain why experts are able to find more targets and tend to find targets more quickly (18,19).

In addition to learning where items of interest might be in a scene, experts also develop superior encoding of large-scale visual patterns. In the basic literature, expert chess players provide a useful example of this sort of superior perceptual encoding. Their understanding of the game allows them to grasp and remember chess positions far more rapidly and effectively than novice players can (20). However, if the pieces are placed randomly, in an analog of the random search display of Figure 2a, chess experts are less able than before to rapidly encode the “scene” into memory (21). (For a review and discussion of the relationship of this work to radiologic expertise, see reference 22.)

From this perspective, the search model of Kundel et al (23) is similar to the theory of Chase and Simon (24) on skilled performance in chess. In both cases, experts are thought to possess an enhanced vocabulary in their domain of expertise, which allows them to efficiently evaluate groups of features rather than focusing on individual features, resulting in more information being encoded in less time. It therefore follows that the benefits of expertise in medical imaging should be domain specific.

Whether radiologists are explicitly aware of the details or not, training has taught them to guide their attention to likely items of interest more efficiently. As with the chess experts, training also

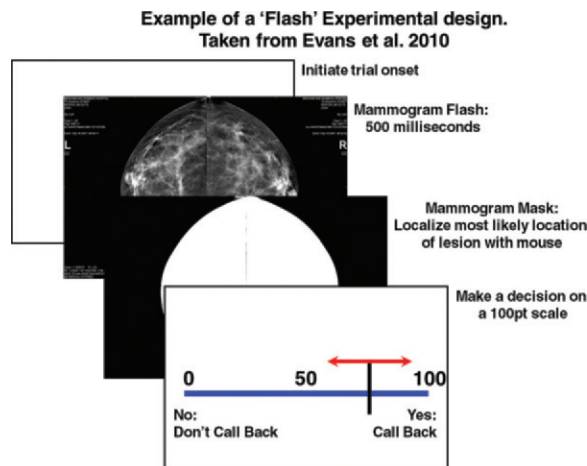
allows them to identify what they have found once attention is given; however, our focus in this article is on the initial guidance of their search. It is useful to distinguish between *feature* guidance and *scene* guidance. In feature guidance, visual features of the item of interest attract attention. Thus, to return to a previous example, if the observer knows that the target, a *T*, will be red (if present), attention will be guided to red items in the display (25).

Not all possible features will guide attention. Instead, there appears to be a set of one or two dozen that will serve this role (26). In medical image perception, the concept of Swenson (27) of a two-stage detection model was a precursor to ideas about feature guidance. Briefly, this model postulates two serial processing stages: an initial stage of global processing that functions as an attentional filter, followed by a second stage in which explicit attention is focused on those areas that were tagged as interesting in the first stage.

In scene guidance, properties of the broader scene constrain the likely locations of items. Pictures are typically on the wall. Lung nodules are in the lung and more likely in some regions than others. The global-focal search model of Kundel et al (28) captures the idea that “holistic” processing of the image constrains subsequent focal search. In that model, the holistic stage takes place before most eye movements while informing the second stage, which is termed the search-to-find stage. This stage involves shifting the eye gaze to different suspicious locations. Thus, experts’ knowledge of what does and does not warrant additional visual examination should allow them to quickly find suspicious regions.

We envision this global or holistic processing a bit differently, as illustrated in Figure 1. Without any firm commitment to a neural substrate, Wolfe et al (8) describe two pathways whose outputs lead to our experience of the visual world. The selective pathway allows observers to identify specific objects, but due to capacity limits in the nervous system, it is limited to one or a few objects at the same time. In this view, your precise recognition and awareness of any object, including for example *these words*, is based on the output of the selective pathway. That said, even if your attention is directed to *these words*, you are not blind to

Figure 3. Experimental procedure in the study of Evans et al (36). Diagnostic performance was reliably above that expected by chance when radiologists classified mammograms after a very brief display. However, the ability to localize the lesion was not significantly above that expected by chance. The red arrows reflect the fact that the radiologists were able to move the selection bar to reflect confidence in their decision. pt = point.



objects elsewhere in the field. Your nonselective pathway provides some information across the entire visual field even as recognition is restricted to the current object of attention.

Scene guidance, which is roughly equivalent to the global or holistic aspects of the Kundel et al (23) model, is guidance of attention on the basis of outputs of nonselective processing. Nonselective processing goes by various names in various literatures: Kundel et al (23,28) refer to it as global or holistic processing; the vision science literature refers to the gist of a scene; and the vision science and radiology literatures refer to the gestalt of a scene. Neuroimaging studies provide some evidence for this two-pathway proposal. Some regions in the occipitotemporal cortex selectively respond to individual objects, while other regions respond to global properties such as spatial layout, texture, or image statistics (29–32).

Scene Guidance in Radiologic Search

Some of the earliest evidence for scene guidance in medical image perception comes from the previously discussed study in which trained radiologists were asked to detect abnormalities on chest radiographs after viewing each image for just 200 msec (2). Performance was well above that expected by chance (~70% correct) and increased to nearly perfect with unlimited viewing. The basic finding that radiologists can quickly extract diagnostic information from medical images has been replicated a number of times with

lung nodules of varying conspicuity (33,34) as well as mammography (35). Unsurprisingly, in all of these flash studies, performance was much improved when radiologists were given more time to search the images. However, together the results of these flash studies strongly suggest that radiologists are able to extract important information in a single glance, before any eye movements and without foveal vision.

One interesting caveat to this conclusion is that although detection accuracy is consistently above that expected by chance, localization accuracy is not. Evans et al (36) asked mammographers to categorize and then localize malignancies on mammograms that were each flashed for 500 msec (Fig 3). Performance equivalent to that of chance in the localization portion of the tasks supports the idea that performance of these very fast normal-versus-abnormal categorization tasks may be based on a global signal, with poor spatial resolution that is adequate to establish the likelihood of an abnormality without providing information about its location. This finding represents a puzzle that warrants further investigation.

On the one hand, if very fast detection of abnormalities is not accompanied by localization of the item of interest, how would nonselective processing guide the deployment of attention and the movements of the eyes? One possibility is that the scene structural information that is guiding the eyes takes a bit longer to develop. According to Nodine and Mello-Thoms (37), the holistic stage might last a second or so. Another possibility is that the scene structural information guides attention to places that could contain items of in-

Teaching
Point

Teaching
Point

terest (eg, walls could hold pictures), but that this aspect of nonselective guidance is distinct from a global or statistical process that can be used to roughly categorize images as, for example, normal or abnormal.

The ability to categorize a medical image as normal or abnormal after a brief exposure may be a trained specialization of the ability of nonexperts to categorize real scenes. To study ultrafast scene processing, researchers typically display a “pattern mask” of irrelevant, high-contrast information immediately after a briefly flashed image. Pattern masks are thought to prevent further perceptual processing of the image. Research has shown that humans can categorize scenes at levels above that expected by chance after an exposure of 20–100 msec followed by display of a pattern mask.

Under these conditions, observers can monitor a stream of images for a single type of scene (eg, a picnic scene) (5,6). They can categorize scenes as natural or urban (7,38,39) and as indoor or outdoor (3). They can even decide if a scene contains an animal or a vehicle (40,41). This rapid categorization of scenes on the basis of the presence of a target remains successful even when the images are presented at low contrast (42), when they are rotated or inverted (43,44), or when chromatic information is removed from them (45,46).

Eye Movements in Medical Image Perception

Humans make three or four voluntary, ballistic eye movements (“saccades”) per second as they examine the visual world. How are those eye movements organized? If you are currently fixated on one spot in an image, what sort of visual information and what sort of decision processes determine the next fixation?

We need to move the eyes because we have high-resolution vision only for images that fall on the fovea of the eye. The choice of the next fixation point must be based on the lower-resolution input from the extrafoveal periphery of the visual field. In the brief exposure experiments described earlier, durations of one-fourth of a second or less were used because, in that period, one would not have time to make a voluntary eye movement. Consequently, the input to the visual system would consist of a single snapshot of the scene.

The development of eye-tracking technology allowed researchers to look at the sequences of saccades as an observer examined an image. Kundel and La Follette (16) and Kundel et al (47,48) were pioneers in the use of eye tracking to study medical image perception. While it is difficult to measure the deployment of covert attention as radiologists view medical images, it is possible to monitor eye movements during free viewing of those images. **The general finding is that experts tend to fixate the malignancy or other items of interest soon after first seeing the image, long before their attention would be expected to reach those sites if the first glance of the image did not provide any location information (23).** When this methodology was used in one study, more than one-half of the lesions fixated by the participating fellows and residents were first fixated within the first 1.1 seconds of the trial (23).

According to the model of Kundel et al (23), the holistic processing stage uses information from the first glance at the image to identify anomalous areas that are categorized as “perturbations.” Gaze is then directed at the perturbations for closer inspection. Another study by the same group analyzed eye-tracking data by using a mixture model. By analyzing how long it took to first fixate cancerous lesions, they found that approximately two-thirds of the lesions were initially fixated within the 1st second of the trial. The remaining one-third of cancer locations were fixated later in the trial. This result suggests that most lesions detected on mammograms are located with the aid of guidance that is available within that 1st second. If that guidance fails, a slower, perhaps more systematic search yields the remaining one-third of the lesions that are eventually found.

Development of Expertise in Nonselective Search

It seems clear that guidance in medical image search must be learned. The most obvious aspect of this learning is the acquisition of explicit knowledge about anatomy, disease states, and so on. Hand in hand with explicit learning is implicit learning of search strategies, as seen, for example, in the evidence that expert radiologists tend to have much more efficient eye movement

Teaching
Point

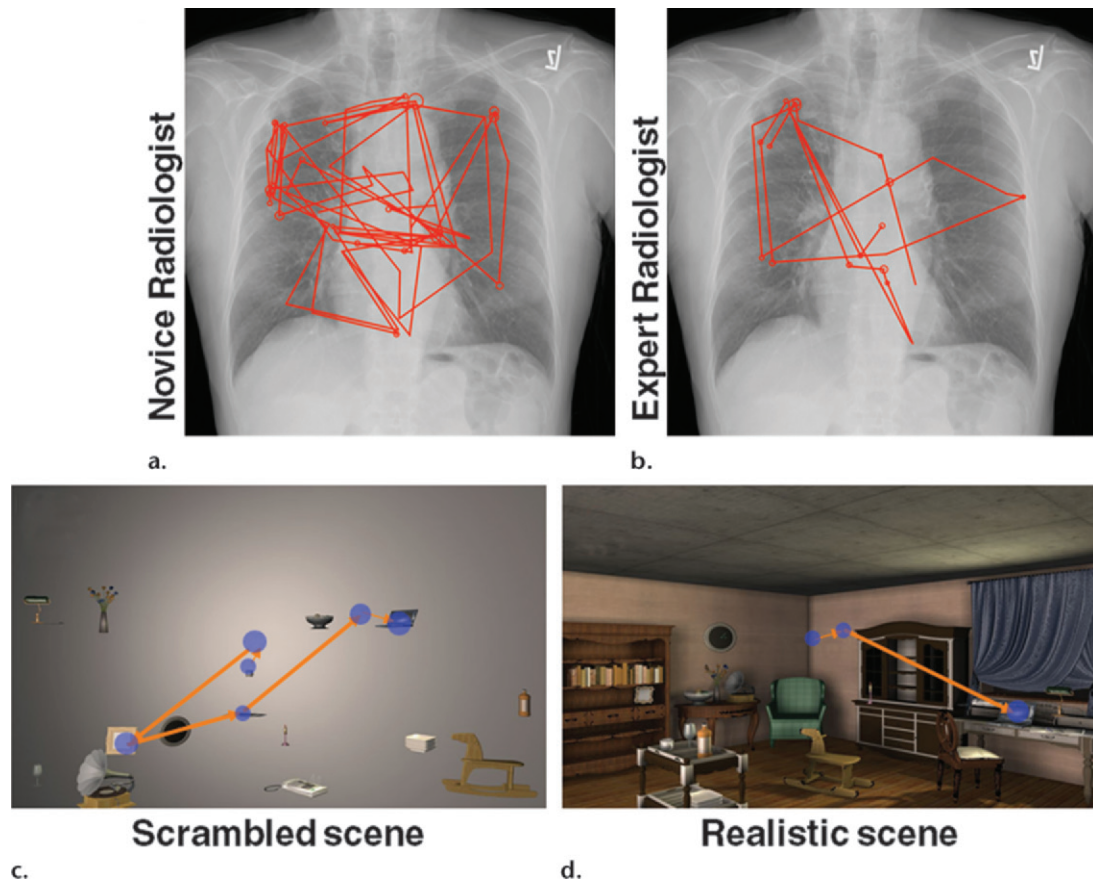


Figure 4. Scan paths in visual search. **(a, b)** Scan paths (red lines) for a first-year resident **(a)** and an expert radiologist **(b)** while searching a chest radiograph for lung nodules. **(c, d)** Analogous scan paths for naive observers searching artificial scene stimuli. Scene guidance allows the target to be quickly found in a realistic scene **(d)**. However, when the background is removed and the object positions are scrambled, more eye movements and more time are necessary to find the target **(c)**. (Figs 4a and 4b courtesy of Elizabeth Krupinski, PhD, University of Arizona College of Medicine, Tucson, Ariz.)

scan paths than novices (Fig 4a, 4b). Typically, people are not aware of the details of their eye movements, and in most cases the scan paths of radiologists, while learned, are probably not explicitly trained. Nevertheless, they spend less time looking at empty locations and initially fixate the lesions much more quickly than their less-experienced peers (16,18). Relative to experts, novices tend to show more fixations and saccades and more coverage of the total image and to arrive at the abnormality later (16,49,50).

What makes the scan paths of expert radiologists more efficient? Presumably, experts are more familiar with the class of medical images they are examining. This familiarity allows the eye movements of expert radiologists to be drawn to the sites of lesions, with less time spent on irrelevant sites. A similar pattern of efficient eye

movements can be observed when humans search real scenes for a prespecified target (51–54). Told to search for a toaster, a person would not bother to search the bathroom and would be unlikely to search on the floor of the kitchen. Similarly, during a search for chest nodules, an expert radiologist will learn to search areas where the target is mostly likely to occur.

We can see how the role of scene structure emerges on a very short time scale by using a method known as the flash-preview moving-window paradigm (54,55). In this paradigm, a scene is briefly presented to observers. This provides the observer with the gestalt of the scene. Once the scene is gone, observers are told what to search for: a lamp, for instance. Finally, they are allowed to search the space where the scene was presented using a gaze-contingent window, which reveals only a small region around the current point of fixation while leaving the remainder of

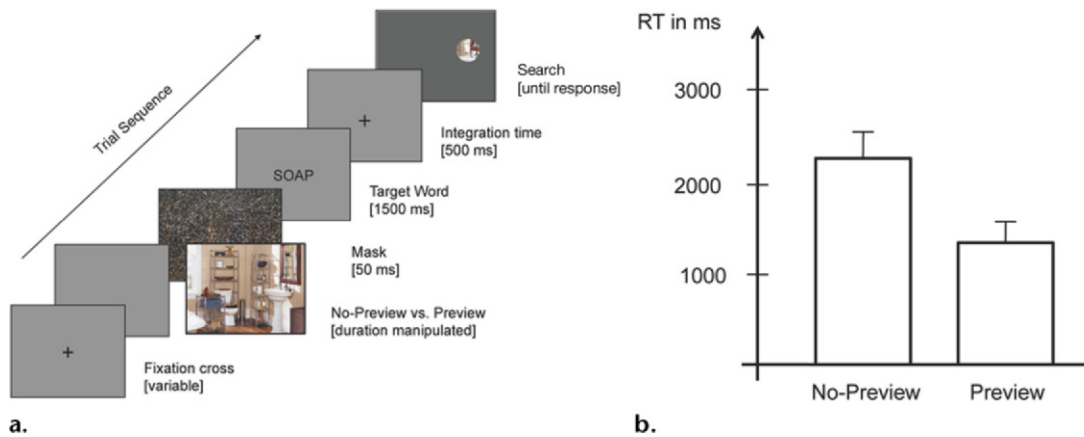


Figure 5. Experimental procedure (a) and results (b) of the study of Vö and Henderson (54). A 50-msec preview of a scene was found to significantly decrease the time needed to find the target item in subsequent viewing of the same scene by using a gaze-contingent window. *RT* in b = response time.

the scene obscured. The experience is like shining a flashlight around a darkened room (Fig 5). In this experiment, the spotlight of the observer's fixation is guided by the gestalt encoded during the brief flash.

By comparing performance in the presence and absence of a brief initial presentation of the entire scene, Vö and Henderson (54) concluded that an initial presentation of just 50 msec was enough to establish a global scene representation that made the search more efficient. In that experiment, the target was never present in the initial preview of the scene. Thus, guidance was not based on a lucky discovery of the target location in the preview, nor was it based on feature guidance (eg, noticing where the red thing was located). This was scene guidance, in which a rapid understanding of some aspects of the scene shaped the deployment of attention and the eyes. Because we are all expert scene processors, trained by our many years of getting around in the world, the first glance at a scene is enough to guide the subsequent search to logical locations (eg, walls to find a picture).

In the medical perception literature, Kundel and colleagues (56) used a similar gaze-contingent window to assess the importance of central and peripheral vision in search through a chest radiograph. Unlike the previously discussed studies in scene processing, the observers in this study were never given a preview of the entire image. Rather, the size of the gaze-contingent window was systematically varied across trials. Nodule detection rates decreased below full-field performance only when the window was less than 5° of visual angle in diameter. As these observers were never given a chance to process the medical

scene in a nonselective fashion, it would be interesting to see whether preview of the entire image before the search would improve performance in the presence of a small gaze-contingent window.

In the model of medical image perception of Kundel et al (23), expert radiologists are able to very quickly detect perturbations, or deviance from perceptual expectation, in the holistic perception of a medical image. This may be why the eye movements of experts searching through medical images tend to be much more refined than those of their less-experienced counterparts (49) (Fig 4). Scrambling scenes destroys our ability to use our experience with scene structure, making perturbations more difficult to detect and ultimately resulting in less-efficient search. In that case, as in the unstructured search of Figure 2a, scene guidance is not possible, and the scan paths look similar to those of a radiologist in training (Fig 4a) (57).

One might wonder whether expert radiologists are naturally better at searching for all manner of targets than are nonradiologists. This possibility was convincingly refuted in a study in which radiologists and nonradiologists were asked to perform a number of nonmedical search tasks, including finding "where's Waldo." The authors found no difference in performance (58). More recently, Evans et al (59) asked mammographers and naive observers to memorize sets of scenes and medical images. While the mammographers were much better at recalling mammograms, radiologists and naive observers had equivalent memory performance when recalling scenes.

Teaching Point

Thus, it appears that expertise in medical image perception is domain specific and dependent on the extensive training that radiologists receive in that domain. As described earlier, a wide gulf separates expert radiologists from trainees in terms of both diagnostic accuracy and speed. Perhaps, with a better understanding of the specialized processing that separates expert radiologists from trainees, we can teach trainees to perform more like experts.

Training for Expertise in Nonselective Search

More than 40 years of research has shown that the first glance at a medical image can convey a great deal of information to an expert and that one of the most dramatic differences between experts and trainees may be the amount of information extracted during the first moments of perception. While this evidence has been accumulating, radiology has been changing as a field. The chest radiographs used in the landmark studies in the 1970s are rapidly being replaced by chest CT acquisitions composed of hundreds of sections that can be reformatted in any number of ways. Similar changes have occurred in other subspecialties. In almost every corner of radiology, radiologists are being asked to process an ever-increasing number of images in the same amount of time (60). It is now more important than ever to understand how to transform a novice into an expert who can process this flood of images efficiently and effectively.

Although a number of longitudinal studies have started to quantify the evolution of search as radiologists acquire more experience, it is not yet clear what determines how much training is necessary for a trainee to approach expert performance (37,61). Clearly, extensive practice in reading medical images is important, but are some types of practice more effective than others? There is an extensive literature on perceptual learning, in which a naive observer learns to improve performance on a specific task through extensive training (62), although this training is still very brief by the standards of medical education.

A more limited body of research has used medical images. For example, Sowden et al (63) showed that less than 1000 trials of practice improved both sensitivity to contrast in radiographs and microcalcification detection in mammograms for naive observers. An important feature of these studies is that they included in-

stant, reliable feedback after each trial and the presentation of many trials in a short time. In fact, in some cases, perceptual learning does not occur without feedback (64).

Using the model of Kundel et al (23), we can speculate that there may be an explicit role for nonselective processing in the training of radiologists. For example, suppose that trainees in mammography were given practice, with feedback, in performing normal versus abnormal categorization of images. Would that increase the speed or effectiveness of the more-standard training required to become an expert in mammography? Many testable hypotheses of this sort lie at the intersection of radiology and the science of rapid scene perception.

Conclusions

Expert radiologists often speak of knowing or feeling that a case contained an abnormality before being able to find it. Skeptics may hear such statements and assume that this is due to the availability heuristic, in which memory of correct hunches is better than that of incorrect hunches (65). However, a great deal of information about a scene is available in the proverbial blink of an eye. Humans can identify the category of a scene after viewing a picture for a fraction of a second.

Trained radiologists can perform analogous tasks with medical images. Converging evidence suggests that in both cases this ability is due to extensive experience with specific types of scenes. Experience seems to enhance the ability of the observer to use nonselective processing to quickly extract valuable information from scenes, even if the scene in question is a complex medical image.

It is important to note that the search for lesions is only one step in the process of successfully interpreting a medical image. Once an anomaly has been detected, there is still much work to do in terms of identification, diagnosis, and ultimately the final recommendation for the patient. While all of these steps must be performed correctly for the correct diagnosis to be made, the sequential nature of this decision process ensures that, if an error is made in the search process, it is unlikely that the correct diagnosis will be made.

In our view, to become an expert radiologist, one has to attain a level of expertise in both explicit medical knowledge and more implicit perceptual knowledge. This article focused on the role of fast, nonselective processing in contributing to perceptual knowledge. At present, little is known about how educators might foster the development of expert perceptual knowl-

edge in the medical domain. However, given the evidence that nonselective processing plays an important role in diagnostic radiology, we hope that educators will find ways to encourage the development of nonselective perceptual expertise in radiologists.

Disclosures of Conflicts of Interest.—**J.M.W.:** *Related financial activities:* none. *Other financial activities:* consultant for Procter & Gamble and LongShortWay; speaker at the Visualization Tools in Medical Education and Expertise meeting; royalties from Sinauer.

References

- Palmer SE. Modern theories of gestalt perception. *Mind Lang* 1990;5(4):289–323.
- Kundel HL, Nodine CF. Interpreting chest radiographs without visual search. *Radiology* 1975;116(3):527–532.
- Fei-Fei L, Iyer A, Koch C, Perona P. What do we perceive in a glance of a real-world scene? *J Vis* 2007;7(1):10.
- Greene MR, Oliva A. The briefest of glances: the time course of natural scene understanding. *Psychol Sci* 2009;20(4):464–472.
- Intraub H. Rapid conceptual identification of sequentially presented pictures. *J Exp Psychol Hum Percept Perform* 1981;7(3):604–610.
- Potter MC. Short-term conceptual memory for pictures. *J Exp Psychol Hum Learn* 1976;2(5):509–522.
- Rousselet GA, Joubert OR, Fabre-Thorpe M. How long to get to the “gist” of real-world natural scenes? *Vis Cogn* 2005;12(6):852–877.
- Wolfe JM, Võ ML, Evans KK, Greene MR. Visual search in scenes involves selective and nonselective pathways. *Trends Cogn Sci* 2011;15(2):77–84.
- Chong SC, Treisman A. Representation of statistical properties. *Vision Res* 2003;43(4):393–404.
- Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci* 2001;4(7):739–744.
- Williams DW, Sekuler R. Coherent global motion percepts from stochastic local motions. *Vision Res* 1984;24(1):55–62.
- Tsotsos JK. *A computational perspective on visual attention*. Cambridge, Mass: MIT Press, 2011.
- Krupinski EA. Future of medical image perception. In: Samei E, Krupinski EA, eds. *The handbook of medical image perception and techniques*. Cambridge, England: Cambridge University Press, 2010; 413–416.
- Evans KK, Horowitz TS, Wolfe JM. When categories collide: accumulation of information about multiple categories in rapid scene perception. *Psychol Sci* 2011;22(6):739–746.
- Wolfe JM. What can 1 million trials tell us about visual search? *Psychol Sci* 1998;9(1):33–39.
- Kundel HL, La Follette PS Jr. Visual search patterns and experience with radiological images. *Radiology* 1972;103(3):523–528.
- Chun MM, Jiang Y. Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognit Psychol* 1998;36(1):28–71.
- Krupinski EA. Visual scanning patterns of radiologists searching mammograms. *Acad Radiol* 1996;3(2):137–144.
- Krupinski EA, Tillack AA, Richter L, et al. Eye-movement study and human performance using telepathology virtual slides: implications for medical education and differences with experience. *Hum Pathol* 2006;37(12):1543–1556.
- Reingold EM, Charness N, Pomplun M, Stampe DM. Visual span in expert chess players: evidence from eye movements. *Psychol Sci* 2001;12(1):48–55.
- Gobet F, Simon HA. Templates in chess memory: a mechanism for recalling several boards. *Cognit Psychol* 1996;31(1):1–40.
- Reingold EM, Sheridan H. Eye movements and visual expertise in chess and medicine. In: Liversedge SP, Gilchrist ID, Everling S, eds. *Oxford handbook on eye movement*. Oxford, England: Oxford University Press, 2011; 528–550.
- Kundel HL, Nodine CF, Conant EF, Weinstein SP. Holistic component of image perception in mammogram interpretation: gaze-tracking study. *Radiology* 2007;242(2):396–402.
- Chase WG, Simon HA. Perception in chess. *Cognit Psychol* 1973;4(1):55–81.
- Egeth HE, Virzi RA, Garbart H. Searching for conjunctively defined targets. *J Exp Psychol Hum Percept Perform* 1984;10(1):32–39.
- Wolfe JM, Horowitz TS. What attributes guide the deployment of visual attention and how do they do it? *Nat Rev Neurosci* 2004;5(6):495–501.
- Swenson RG. A 2-stage detection model applied to skilled visual-search by radiologists. *Percept Psychophys* 1980;27(1):11–16.
- Kundel HL, Nodine CF, Thiekman D, Toto L. Searching for lung nodules: a comparison of human performance with random and systematic scanning models. *Invest Radiol* 1987;22(5):417–422.
- Aguirre GK, Zarahn E, D’Esposito M. Neural components of topographical representation. *Proc Natl Acad Sci U S A* 1998;95(3):839–846.
- Epstein R, Graham KS, Downing PE. Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 2003;37(5):865–876.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 2001;293(5539):2425–2430.
- Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature* 1998;392(6676):598–601.
- Carmody DP, Nodine CF, Kundel HL. Finding lung nodules with and without comparative visual scanning. *Percept Psychophys* 1981;29(6):594–598.
- Oestmann JW, Greene R, Kushner DC, Bourgoin PM, Linetsky L, Llewellyn HJ. Lung lesions: correlation between viewing time and detection. *Radiology* 1988;166(2):451–453.
- Mugglestone MD, Gale AG, Cowley HC, Wilson AR. Diagnostic performance on briefly presented mammographic images. *Proc SPIE* 1995;2436:106–115.

36. Evans KK, Georgian-Smith D, Birdwell RL, Wolfe JM. Discrimination and localization of abnormalities in mammograms from a global signal [abstr]. In: Radiological Society of North America scientific assembly and annual meeting program. Oak Brook, Ill: Radiological Society of North America, 2010; 438.
37. Nodine CF, Mello-Thoms C. The role of expertise in radiologic image interpretation. In: Samei E, Krupinski EA, eds. *The handbook of medical image perception and techniques*. Cambridge, England: Cambridge University Press, 2010; 139–156.
38. Greene MR, Oliva A. Recognition of natural scenes from global properties: seeing the forest without representing the trees. *Cognit Psychol* 2009;58(2): 137–176.
39. Joubert OR, Rousselet GA, Fize D, Fabre-Thorpe M. Processing scene context: fast categorization and object interference. *Vision Res* 2007;47(26): 3286–3297.
40. Evans KK, Treisman A. Perception of objects in natural scenes: is it really attention free? *J Exp Psychol Hum Percept Perform* 2005;31(6):1476–1492.
41. Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. *Nature* 1996;381(6582): 520–522.
42. Macé MJ, Delorme A, Richard G, Fabre-Thorpe M. Spotting animals in natural scenes: efficiency of humans and monkeys at very low contrasts. *Anim Cogn* 2010;13(3):405–418.
43. Guyonneau R, Kirchner H, Thorpe SJ. Animals roll around the clock: the rotation invariance of ultrarapid visual processing. *J Vis* 2006;6(10):1008–1017.
44. Rousselet GA, Macé MJ, Fabre-Thorpe M. Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *J Vis* 2003;3(6): 440–455.
45. Delorme A, Richard G, Fabre-Thorpe M. Ultrarapid categorisation of natural scenes does not rely on colour cues: a study in monkeys and humans. *Vision Res* 2000;40(16):2187–2200.
46. Thorpe SJ, Gegenfurtner KR, Fabre-Thorpe M, Bülthoff HH. Detection of animals in natural images using far peripheral vision. *Eur J Neurosci* 2001; 14(5):869–876.
47. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol* 1978;13(3):175–181.
48. Kundel HL, Nodine CF, Krupinski EA. Searching for lung nodules: visual dwell indicates locations of false-positive and false-negative decisions. *Invest Radiol* 1989;24(6):472–478.
49. Krupinski EA, Lund PJ. Scanning differences in film versus monitor viewing of radiographs [abstr]. *Radiology* 1996;201(P):789.
50. Manning D, Barker-Mill SC, Donovan T, Crawford T. Time-dependent observer errors in pulmonary nodule detection. *Br J Radiol* 2006;79(940): 342–346.
51. Castelhana MS, Heaven C. The relative contribution of scene context and target features to visual search in scenes. *Atten Percept Psychophys* 2010;72(5):1283–1297.
52. Võ ML, Wolfe JM. When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. *J Exp Psychol Hum Percept Perform* 2012;38(1):23–41.
53. Torralba A, Oliva A, Castelhana MS, Henderson JM. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychol Rev* 2006;113(4): 766–786.
54. Võ ML, Henderson JM. The time course of initial scene processing for eye movement guidance in natural scene search. *J Vis* 2010;10(3):14.1–13.
55. Castelhana MS, Henderson JM. Initial scene representations facilitate eye movement guidance in visual search. *J Exp Psychol Hum Percept Perform* 2007;33(4):753–763.
56. Kundel HL, Nodine CF, Toto L. Searching for lung nodules: the guidance of visual scanning. *Invest Radiol* 1991;26(9):777–781.
57. Foulsham T, Alan R, Kingstone A. Scrambled eyes? Disrupting scene structure impedes focal processing and increases bottom-up guidance. *Atten Percept Psychophys* 2011;73(7):2008–2025.
58. Nodine CF, Krupinski EA. Perceptual skill, radiology expertise, and visual test performance with NINA and WALDO. *Acad Radiol* 1998;5(9): 603–612.
59. Evans KK, Cohen MA, Tambouret R, Horowitz T, Kreindel E, Wolfe JM. Does visual expertise improve visual recognition memory? *Atten Percept Psychophys* 2011;73(1):30–35.
60. Andriole KP, Wolfe JM, Khorasani R, et al. Optimizing analysis, visualization, and navigation of large image data sets: one 5000-section CT scan can ruin your whole day. *Radiology* 2011;259(2):346–362.
61. Nodine CF, Kundel HL, Mello-Thoms C, et al. How experience and training influence mammography expertise. *Acad Radiol* 1999;6(10):575–585.
62. Sagi D. Perceptual learning in vision research. *Vision Res* 2011;51(13):1552–1566.
63. Sowden PT, Davies IR, Roling P. Perceptual learning of the detection of features in x-ray images: a functional role for improvements in adults' visual sensitivity? *J Exp Psychol Hum Percept Perform* 2000;26(1):379–390.
64. Pauli R, Sowden PT. The role of feedback in learning screening mammography. *Proc SPIE* 1997;3036: 205–211.
65. Tversky A, Kahneman D. Availability: a heuristic for judging frequency and probability. *Cognit Psychol* 1973;5(2):207–232.

Informatics in Radiology

What Can You See in a Single Glance and How Might This Guide Visual Search in Medical Images?

Trafton Drew, PhD • Karla Evans, PhD • Melissa L.-H. Võ, PhD • Francine L. Jacobson, MD, MPH • Jeremy M. Wolfe, PhD

RadioGraphics 2013; 33:263–274 • Published online 10.1148/rg.331125023 • Content Codes: ED IN

Page 266

However, an important consequence of training is to teach radiologist trainees that some areas are more likely to contain a lesion than others. Thus, while eye movement recordings suggest that novice radiologists search in a relatively haphazard fashion when looking for lung nodules, experienced radiologists tend to exhibit more concise eye movements, with fewer fixations needed to extract more information.

Page 268

Some of the earliest evidence for scene guidance in medical image perception comes from the previously discussed study in which trained radiologists were asked to detect abnormalities on chest radiographs after viewing each image for just 200 msec. Performance was well above that expected by chance (~70% correct) and increased to nearly perfect with unlimited viewing.

Page 268

One interesting caveat to this conclusion is that although detection accuracy is consistently above that expected by chance, localization accuracy is not.

Page 269

The general finding is that experts tend to fixate the malignancy or other items of interest soon after first seeing the image, long before their attention would be expected to reach those sites if the first glance of the image did not provide any location information.

Page 271

In the model of medical image perception of Kundel et al, expert radiologists are able to very quickly detect perturbations, or deviance from perceptual expectation, in the holistic perception of a medical image. This may be why the eye movements of experts searching through medical images tend to be much more refined than those of their less-experienced counterparts.