Emerging Technologies in Healthcare and Medicine, Vol. 116, 2023, 306–316
https://doi.org/10.54941/ahfe1004384

Differences in Eye Movements in Chest X-Ray Diagnosis and Exploration of Effective Diagnostic Strategies: A Study in Annual Medical Checkup Conditions

Yijia Wang[1], Hirotaka Aoki[1], Koji Morishita[2], Marie Takahashi[2], Rea Machida[1], Atsushi Kudoh[2], Mitsuhiro Kishino[2], and Tsuyoshi Shirai[2]

[1]Tokyo Institute of Technology, Tokyo, Japan
[2]Tokyo Medical and Dental University, Tokyo, Japan

ABSTRACT

During the process of medical interpretation and diagnosis in medical image, doctors' attention allocations are various depending on individuals and cases. It is difficult for doctors to explicitly explain their strategies because the process is often tacit. This study aims to uncover the doctor's attention allocation and transition patterns in reading chest X-ray image, elicit diagnostic strategies based on doctor's eye movements and interviews, and find the differences of diagnostic strategies between expert and novice doctors. Finally, prospective suggestions for leading novice doctors to an effective diagnostic strategy in reading X-ray image can be presented. We simulate the scenario of annual medical check-up condition and recruit participants with diverse medical experiences and specialties in Tokyo Medical and Dental University Hospital to compare the differences of attention allocations between doctors. Eye-tracking and semi-structured interview are used to study doctors' diagnostic process. Both qualitative and quantitative analysis are used to describe doctors' diagnostic strategies and compare the differences between expert and novice doctors. We find that each doctor has personal characteristics when diagnosing. And four typical patterns of change of diagnostic strategies in timeline are found. Furthermore, the differences are found between expert and novice doctors in attention allocation and the use of historical image. The potential effective diagnostic strategy we propose is that performing the inspection routine of the current image separate with comparison with the historical image to avoid distracting and missing information. The comparison should focus on important areas and suspicious areas rather than the whole image. The suggested important areas are the lung apex, mediastinum, heart, left lung hilar and the lower lung field. The proposed effective strategies could be included in the medical education and new doctor training to improve novice doctors' ability to diagnose by multiple images.

Keywords: Eye tracking, Visual perception, Diagnostic strategies, Chest X-ray image, Annual medical checkup condition

INTRODUCTION

Medical imaging such as X-ray imaging, CT, MRI and so on is an important non-intrusive method to detect diseases. Chest X-rays which are the most common film taken in medicine are usually the first step to detect conditions affecting the chest and nearby structures. Currently, the diagnosis of lung diseases does not rely solely on 2D X-rays to make the final diagnosis, but will be followed by more sophisticated examinations, such as CT which is an advanced 3D tomography method. While, despite the advance of 3D tomography, 2D chest X-rays are still widely used in many medical facilities and hospitals for early detection of lung abnormalities, especially new and small lung tumors, due to the low cost, high resolution and lower radiation dosages with 2D technique compared to CT.

In chest X-rays, some structures are overlapping such as lung fields and scapulae because of the 2D visualization (Joo *et al.*, 2016). The detection of abnormalities in chest X-ray images is based on medial knowledge. Doctors process visual information in chest X-ray images to find abnormalities and accurately diagnose abnormalities as lesions, which is a complex cognitive process. The methods of reading chest X-ray images vary depending on the hospital, the supervising doctor, and the individual experience. The methods related to the environment and individual experience is tacit knowledge (Earle, 1959). It is difficult or impossible to explicitly express or extract to other people even by experts (Chugh, 2015). Effective transfer of tacit knowledge generally requires extensive personal contact, regular interaction (Goffin and Koners, 2011), and trust, which is usually time consuming. Therefore, revealing doctors' diagnostic strategies in reading chest X-ray images and finding the gap between expert and novice doctors to convert the tacit knowledge especially of expert doctors to explicitly knowledge is necessary for the faster improvement of novice doctors and medical students.

Studies over years have shown the involvement of various perceptual and cognitive processes during medical image interpretation and diagnosis (Kundel and La Follette, 1972). The process involves complex interplay between visual search and pattern matching, hypothesis generation and testing, and reasoning and problem-solving. The methods to investigate the cognitive process during diagnosing medical image have included think-aloud protocols, knowledge and memory probes, practical exercises, and tracking physicians' interface navigation behaviour while they inspect visual images (e.g., radiographs, histology slides) (Brunyé *et al.*, 2019). Within these years, eye tracking technology is increasingly used in medical domain, which can provide more detailed qualitative and quantitative assessments of how and where the eyes move during diagnosing (John *et al.*, 2018).

Two typical models of visual search patterns have been proposed to capture the complexity of diagnosing X-ray images so far. The earlier model was first proposed by Nodine and Kundel in 1978. The 3-phase theory defined three phases in the visual search, initial overall pattern recognition, focal attention for image detail, and final decision making (Nodine and Kundel, 1987). More recent model discerns only two parts of visual search called a global-focal search model. The global-focal search model contains: (1) a relatively fast

global impression that signals possible abnormalities and (2) a slower, more detailed focal search for recognition and evaluation of abnormalities. And the model also defined as a pre-attentive filter and cognitive evaluation stage (Swensson, 1980); or nonselective and selective pathways running in parallel (Drew *et al.*, 2013). The previous studied has demonstrated the feasibility of eye tracking approach for studying the doctors' diagnostic process, but the phases and interpretation of the diagnostic process have not yet been academically reached a uniform conclusion.

In the real work condition, doctors not only diagnose based on a single image, but more often based on comparison between multiple images. In this condition, the process of diagnosing involves complex interplay between visual perception and effective information acquisition strategies coupled with medical knowledge. However, most of these studies have been conducted based on the diagnostic situation of a single X-ray image. Therefore in this study we simulate one of the most common scenarios of annual medical check-up condition that present two X-ray images taken in different time of the same patient to study doctors' diagnostic strategies.

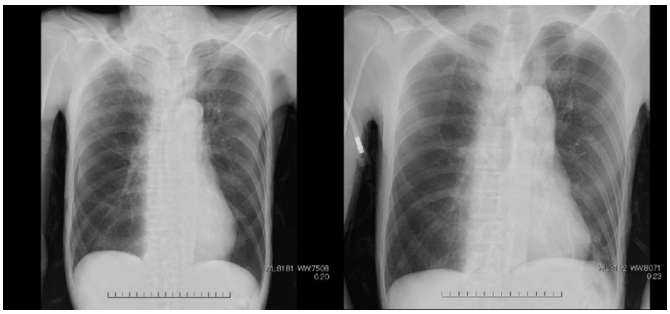

Figure 1: X-ray image presented to simulate annual medical checkup condition.

RESEARCH OBJECTIVES

In this study, we aim to explore both expert and novice doctors' diagnostic strategies in annual medical check-up condition and find the gap between them, hoping to propose effective suggestions to improve doctors training system and medical education in X-ray diagnosing. We aim to explore doctors' diagnostic strategies and find the gap between expert and novice doctors. Therefore, there are three objectives in this thesis. First, we uncover the doctor's attention allocation and transition patterns in reading chest X-ray image. Second, elicit diagnostic strategies based on doctor's eye movements and interviews. Third, find the differences of diagnostic strategies between expert and novice doctors. At the end of this thesis, prospective suggestions for leading novice doctors to an effective diagnostic strategy in reading X-ray image can be presented.

EXPERIMENT DESIGN AND DATA COLLECTION

We selected four cases which include both gender, and both normal and abnormal case (see Table 1) and simulated annual medical checkup condition by presenting two chest X-ray images of the same patient (see Figure 1).

Table 1. Description of cases.

Case	Age	Gender	Time interval of taking two images	Lesion part	Diagnosis
Case 1	76	Male	7 months	Nodular shadow in the upper lobe of the left lung	Lung cancer
Case 2	27	Male	1 week	Nodular shadow in the right 10th rib	Normal
Case 3	63	Female	1 month	Infiltrative shadow in the left lung apex	Lung cancer
Case 4	62	Male	1 year and 5 months	Speckled shadow in the right lung apex	Lung cancer

The whole experiment contains three steps (see Figure 2).

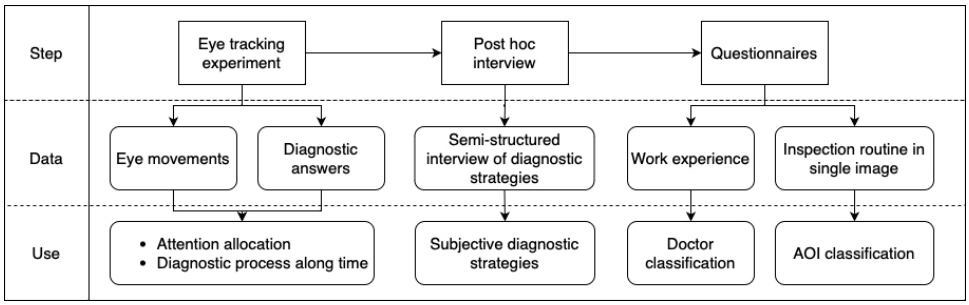

Figure 2: Data collection procedure.

In the first step of eye tracking experiment, the eye movements were recorded by Tobii Pro X3-120, which can record human eye movements via an eye tracker attached to a monitoring screen. Doctors' eye movements, their diagnostic answers and subjective evaluations of each case are collected, which are used to describe the attention allocation and diagnostic process of each doctor in each case. The participants switched to the next page by clicking the mouse themselves during the process. The contents shown on each page are, in order, basic information of the case, the X-ray image and evaluations. Each case followed the same procedure and there is no chance to go back to previous page until finishing all four cases. The basic information shows the patient gender, age and the images taken time interval. After checking the basic information, the participant can click to go to the next page that the current and historical chest X-ray images were presented. The participant was asked to state where he/she found is abnormal and give a

final diagnosis verbally when he/she finished examining the X-ray images. And after diagnosing, the participant could click to the next page and answer the questions shown on this page. The questions are: 1) the first impression of the X-ray image: normal, abnormal, or not sure; 2) the usefulness of the historical image: 1 to 5 scale for very useless to very useful; 3) the difficulty of finding the lesion part: 1 to 5 scale for very easy to very difficult; 4) the difficulty of making a diagnostic decision: 1 to 5 scale for very easy to very difficult; 5) the confidence of the diagnosis: 0 to 100%.

After finishing the eye tracking experiment, doctors were invited to an interview session. A semi-structured interview was conducted to elicit doctors' subjective diagnostic strategies. The interview columns see in Table 2. And during this session, the eye movements were replayed as cues. The interview records doctors' statements of their diagnostic strategies, which are used to transcribe doctors' subjective diagnostic strategies. And the subjective strategies can match with the diagnostic process from eye movements to support each other better revealing doctors' diagnostic strategies.

Table 2. Semi-structured interview columns.

No.	Columns
1	How to identify the lesion
2	The reason for first impression and whether change the reading pattern according to first impression
3	Is there any inspection routine
4	The usefulness of the historical image and how to make use of the historical image
5	The timing of finding the lesion and whether change the reading method according to the timing of find the lesion
6	Where to first to look at
7	Where are important parts in reading
8	Whether change the strategy based on patient information (i.e. gender, age)
9	Whether change the strategy based on the image taken time
10	How to make the final diagnosis
11	Own thoughts of the eye movements
12	The differences between work condition and experimental condition
13	What is different between reading a single X-ray image and reading two X-ray images
14	The influence of image taken condition
15	Differences between examining and treatment

In the last step, participants answered two kinds of questionnaire. One is about their work experience. Doctors need to report how long and what type of hospital they have worked for, and how long they have got Medical License. The other one asked doctors to report their inspection routine of a single chest X-ray image by lines with arrows and circle up to five areas that they think it's important in chest X-ray examine. The work experience is used to classify doctors into different groups for further comparative analysis. The important areas when diagnosing is used to divide the whole chest X-ray image into several areas of interest (AOIs), and the inspection routine of single

chest X-ray image is used to decide the order of these AOIs, which is helpful to study the diagnostic process along time.

TASK ANALYSIS

The analysis procedure follows the principle of task analysis, which contains processing raw data, task description, and analysis. The eye movements data exported from Tobii system are processed by the number of fixation and fixation duration. The number of fixation calculate the number of times the eye fixates in a particular area of interest, which is interpreted as the informational value of the area (Henderson, Malcolm and Schandl, 2009). Fixation duration is how long the eye fixates on a region prior to a saccade. The total fixation duration of an area is related to how much information available in this area. The two metrics were calculated based on areas of interest (AOIs) that classified by lung anatomy and doctors' subjective statement of important parts (see Figure 3). A and B are right and left apex. C and D are right and left lung hilum. I and J are right and left upper lung field. K and L are right and left middle lung field. M and N are right and left lower lung field. G and H are right and left diaphragm. E is upper mediastinum. O is lower mediastinum. F is heart shade.

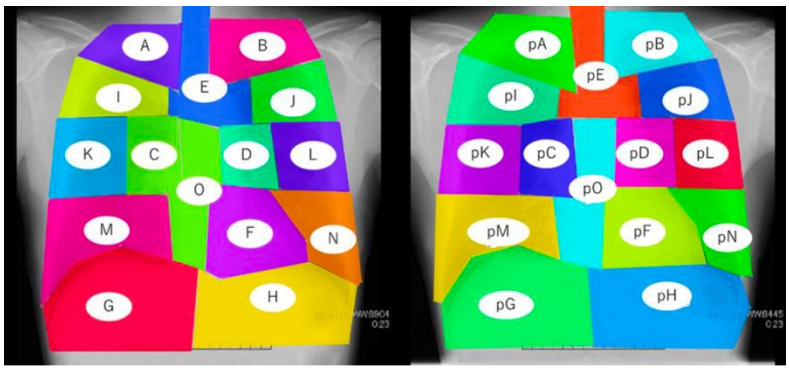

Figure 3: AOI classification.

There are eighteen doctors participate in the experiment, however, due to the technical problem, some participants eye-tracking data was not recorded successfully. Therefore, the data used in the analysis is fourteen doctors. Then fixation duration in each AOI sort from the highest to lowest following ABC analysis to select the AOIs that occupy 80% of total fixation duration. Those AOIs which are occupied the top 80% of fixation duration are seen as important AOIs, for example in Table 3.

To find whether there is any universal difference between cases and doctors. ANOVA is applied. However, the results show that both recognitive difficulty and diagnosing difficulty are not significant different between cases and doctor groups. And other case attributes, completing time and confidence, are also compared, however, it neither show any significant difference between cases and doctor groups. Though the sample size is limited, it still

can be speculated that doctor's cognitive difficulty towards different case is person to person. As for the task completing time of each case, the difference didn't show between doctor groups, which may be attributed to having both fast decision makers and slow decision makers within each group. Then we compared doctors' attention in the current and historical image by adding the type of image as a within subject factor. The results show the number of fixation o and the number of AOI in the top 80% of fixation duration are significant different between the current and historical image. The number of fixation of the current image is larger than of the historical image ($F = 55.453$, $p < 0.001$). The number of AOI in the top 80% of total fixation duration of the current image is larger than the historical image ($F = 236.959$, $p < 0.001$). It indicates that when diagnosing doctors usually process more information of the current image. It can be expected that no significant differences can be found without considering attention allocation in which AOI and attention transition along time. Therefore, further analysis consider AOI and time is performed.

Table 3. ABC analysis of doctor 1 case 1 (AOIs in top 80%).

AOI	The number of fixation	Fixation duration (ms)	Ratio of total fixation duration
A (Right lung apex)*	89	11154	18.4%
E (Upper mediastinum)*	66	8466	14.0%
pA (Right lung apex)	63	7001	11.6%
J (left upper lung field)	36	4575	7.6%
F (Heart shade)*	30	4544	7.5%
B (left lung apex)*	30	4077	6.7%
pI (Right upper lung field)	26	3350	5.5%
D (left lung hilum)*	14	2576	4.3%
I (Right upper lung field)	17	2241	3.7%
pE (Upper mediastinum)	19	2167	3.6%

We calculated the number of times that the AOI ranks in the top 80% of total fixation duration. The value ranges from 0 to 4. Take the number of time more than half as important AOIs (see Figure 4). Both experts and novices view D (left lung hilum), E (upper mediastinum), F (heart shade), O (lower mediastinum) as important areas.

When considering the attention transition along time, the diagnostic process can be divided into different phases. The attention transition is visualized by the fixation transition between AOIs in line chart (see Figure 5). The principle of diagnostic process division is based on typica patterns of eye movements that are correspond to the diagnostic strategies that doctors reported in the interview. The typical pattern are scanning the whole image for abnormalities, conducting personal inspection routine of a single image, comparing between two images or two sides of a single image.

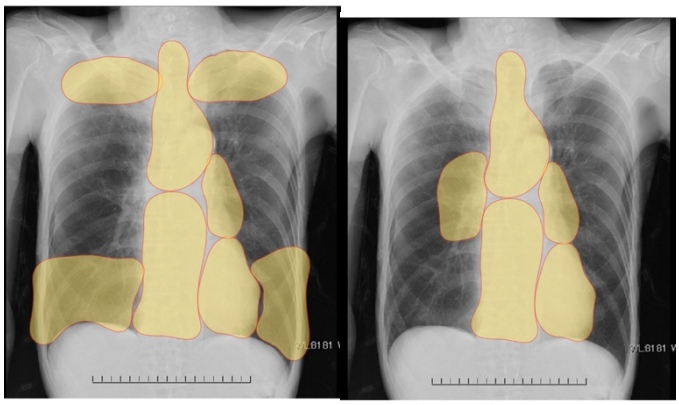

Figure 4: Important areas when diagnosing (left is expert, right is novice).

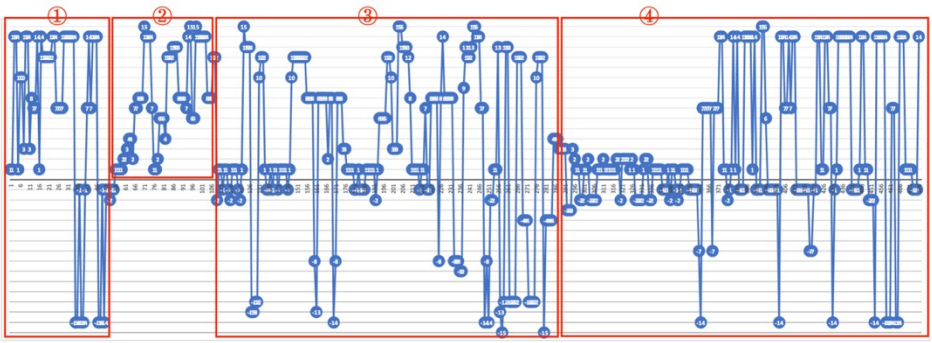

Figure 5: Example of diagnostic process division.

RESULT

The results of important areas when diagnosing show that these areas are where physiological structures overlap. Experts view the lung apex and the lower lung field as the important areas while novice don't. The lower lung field contains the border of lung anatomy and the lung apex is where overlap other physiological structures. Novices view the right lung hilum as an important area while experts don't. Experts just view the left lung hilum as an important area. It may because the left side is where heart is, which means the hilum areas is not symmetrical.

When we consider the fixation transition, we find that doctors' diagnostic strategies in the whole process can be summarized in four patterns. Pattern 1is scanning for abnormalities first, and regardless of whether the abnormality is found, full inspection routine of the current image performed. After the inspection routine of the current image, important areas and suspicious areas are compared. And there are some cases start with the inspection routine of the current image also include in pattern 1 because the absence of scanning phase may due to the time of scanning is too short to find. Pattern 2 is after scanning for abnormalities which is also possible absent, the doctor performs the inspection routine of both current and historical images

and compare after it. Pattern 3 is after scanning for abnormalities, the doctor directly compares the abnormalities. After compare, the inspection routine is performed to avoid missing information. Pattern 4 is comparison through the whole process. The biggest difference between novice and expert doctors in diagnostic strategy pattern is that all the experts have clear phases when diagnosing (Pattern 1, 2 or 3). Though the diagnostic strategies are not the same in expert group, all experts have performed the inspection routine of the current image in all cases. However, the most common diagnostic pattern of novice doctors is pattern 4, which does not contain the clear inspection routine.

Compared to novices, expert doctors spare time to examine the current image separately and the comparison is just cover the areas that the doctor is concerned. It indicates that expert doctor's diagnosis is based on the finding in the current image, and the historical image is used to provide additional information to confirm of the concerned areas. On the other hand, most of novice doctors don't spare time to focus on the single current image for full examining. The inspection routine of the whole current image is integrated in the comparison, which means after they read a part of the current image, they would refer to the same part of the historical image until they finish comparing all areas in their minds. This pattern interrupts the inspection routine, which lead to higher possibility to miss important information. It indicates novice doctors rely on the historical image more. In other word, novice doctors are easier distracted by the historical information.

DISCUSSION

From the comparison between expert and novice doctors, the biggest difference is that novice doctors rely on the historical chest X-ray image too much, which distracts them to make full inspection routine of the current chest X-ray image. Therefore, to improve novice doctors' performance in diagnosing chest X-ray images, we should reduce the distraction of the historical image from the current image diagnosis.

Suggestions for improving the performance of novice doctors in reading chest X-ray images in the annual medical check-up condition are presented here. Firstly, in order to reduce the distraction of historical images from the current image reading process, the single image reading process needs to be enhanced. And where expert doctors view as important areas should be presented to novice doctors. At the same time, in order to be more adaptable to the medical check-up condition which uses multiple X-ray images, a section on multi-image diagnostic strategies should be included in medical education and training, emphasizing methods of focusing on single-image information and utilizing historical images.

All doctors participating in the experiment reported that if there is a historical chest X-ray image, they will read it. And evaluation results of the usefulness of historical images shows that all doctors participating in the experiment think the historical image is help for them to make diagnosis of the current image (more than 3 in the 1~5 scale). It indicates that the comparative diagnosis in medical check-up condition has the rationality.

CONCLUSION

This study simulates the annual medical check-up condition to enable doctors give diagnosis based on two chest X-ray images, aiming to reveal doctors' diagnostic strategies. First, doctors' attention allocation and transition patterns are described by fixations. They indicate that each doctor has their own characteristics when diagnosing. The attention allocation shows that doctors tend to pay more attention to areas that they consider suspicious and areas where physiological structure overlap. Besides, in case 2 which is female patient, doctors pay more attention to the lower lung field where the bread is overlap, which indicates that doctors tend to consider the gender when diagnosing. The fixation transition patterns show the whole diagnostic process can be divided into some phases. The typical phases include in the diagnostic process are scanning the whole image for abnormalities, conducting personal inspection routine of a single image, comparing between two images or two sides of a single image. And four diagnostic strategy change patterns are found based on fixation transition and interview.

Second, the differences of diagnostic strategies between expert and novice doctors are found. All experts will perform inspection routine on the current chest X-ray image, while novices like mixing the inspection routine with the comparison. As a result, novice doctors don't like reading the current chest X-ray image separately. They are easier distracted by the information in the historical chest X-ray image than expert doctors, and their diagnosis rely on the historical image too much. And in the diagnostic process, expert doctors cover more areas as important areas than novice doctors. Expert doctors view the lung apex, mediastinum, heart, left lung hilar and the lower lung field as the important areas, while novice doctors view mediastinum, heart, left and right lung hilar as important areas.

Finally, the effective diagnostic strategies are proposed. When diagnosing in the annual medical checkup condition, the full inspection routine of the single current image should be performed, and the compare between different images should be focused to important areas and suspicious areas. The suggested important areas are the lung apex, mediastinum, heart, left lung hilar and the lower lung field. The proposed effective strategies could be included in the medical education and new doctor training to improve novice doctors' ability to diagnose by multiple images.

ACKNOWLEDGMENT

This research was partly supported by the Grant-in-Aid for Scientific Research (C) (No. 21K04528), The Japan Society for the Promotion of Science. We would like to acknowledge Dr. Kei-ichi Akahoshi for his great effort to arrange the experimental environment.

REFERENCES

Brunyé, T. T. et al. (2019) 'A review of eye tracking for understanding and improving diagnostic interpretation', Cognitive Research: Principles and Implications, 4(1), p. 7. Available at: https://doi.org/10.1186/s41235-019-0159-2.

Chugh, R. (2015) 'Do Australian Universities Encourage Tacit Knowledge Transfer?', in Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. 7th International Conference on Knowledge Management and Information Sharing, Lisbon, Portugal: SCITEPRESS - Science and and Technology Publications, pp. 128–135. Available at: https://doi.org/10.5220/0005585901280135.

Drew, T. et al. (2013) 'Informatics in radiology: what can you see in a single glance and how might this guide visual search in medical images?', Radiographics: A Review Publication of the Radiological Society of North America, Inc, 33(1), pp. 263–274. Available at: https://doi.org/10.1148/rg.331125023.

Earle, W. (1959) 'Personal Knowledge. Towards a postcritical philosophy. Michael Polanyi. University of Chicago Press, Chicago, Ill., 1958. xiv + 428 pp. $6.75.', Science, 129(3352), pp. 831–832. Available at: https://doi.org/10.1126/science.129.3352.831.

Goffin, K. and Koners, U. (2011) 'Tacit Knowledge, Lessons Learnt, and New Product Development', Journal of Product Innovation Management, 28(2), pp. 300–318. Available at: https://doi.org/10.1111/j.1540-5885.2010.00798.x.

Henderson, J. M., Malcolm, G. L. and Schandl, C. (2009) 'Searching in the dark: Cognitive relevance drives attention in real-world scenes', Psychonomic Bulletin & Review, 16(5), pp. 850–856. Available at: https://doi.org/10.3758/PBR.16.5.850.

John, K. K. et al. (2018) 'Emerging applications of eye-tracking technology in dermatology.', Journal of dermatological science, pp. S0923-1811(18)30156–7. Available at: https://doi.org/10.1016/j.jdermsci.2018.04.002.

Joo, Y. C. et al. (2016) 'Comparison with Overlap Area of Scapular and Lung Field according to Position on Chest Radiography', Indian Journal of Science and Technology, 9(25). Available at: https://doi.org/10.17485/ijst/2016/v9i25/97229.

Kundel, H. L. and La Follette, P. S. (1972) 'Visual Search Patterns and Experience with Radiological Images', Radiology, 103(3), pp. 523–528. Available at: https://doi.org/10.1148/103.3.523.

Nodine, C. F. and Kundel, H. L. (1987) 'Using eye movements to study visual search and to improve tumor detection', Radiographics: A Review Publication of the Radiological Society of North America, Inc, 7(6), pp. 1241–1250. Available at: https://doi.org/10.1148/radiographics.7.6.3423330.

Swensson, R. G. (1980) 'A two-stage detection model applied to skilled visual search by radiologists', Perception & Psychophysics, 27(1), pp. 11–16. Available at: https://doi.org/10.3758/BF03199899.