Journal of Pathology Informatics 13 (2022) 100104

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

An analysis of pathologists' viewing processes as they diagnose whole slide digital images

Fatemeh Ghezloo [a,*], Pin-Chieh Wang [b], Kathleen F. Kerr [c], Tad T. Brunyé [d], Trafton Drew [e], Oliver H. Chang [f], Lisa M. Reisch [c], Linda G. Shapiro [a], Joann G. Elmore [b]

[a] Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA
[b] Department of Medicine, University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, CA, USA
[c] Department of Biostatistics, University of Washington, Seattle, WA, USA
[d] Center for Applied Brain and Cognitive Sciences, Tufts University, Medford, MA, USA
[e] Department of Psychology, University of Utah, Salt Lake City, UT, USA
[f] Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA

ARTICLE INFO

Keywords:
Digital imaging
Digital pathology
Whole slide imaging
Statistical analysis
Computational pathology

ABSTRACT

Although pathologists have their own viewing habits while diagnosing, viewing behaviors leading to the most accurate diagnoses are under-investigated. Digital whole slide imaging has enabled investigators to analyze pathologists' visual interpretation of histopathological features using mouse and viewport tracking techniques. In this study, we provide definitions for basic viewing behavior variables and investigate the association of pathologists' characteristics and viewing behaviors, and how they relate to diagnostic accuracy when interpreting whole slide images. We use recordings of 32 pathologists' actions while interpreting a set of 36 digital whole slide skin biopsy images (5 sets of 36 cases; 180 cases total). These viewport tracking data include the coordinates of a viewport scene on pathologists' screens, the magnification level at which that viewport was viewed, as well as a timestamp. We define a set of variables to quantify pathologists' viewing behaviors such as zooming, panning, and interacting with a consensus reference panel's selected region of interest (ROI). We examine the association of these viewing behaviors with pathologists' demographics, clinical characteristics, and diagnostic accuracy using cross-classified multilevel models. Viewing behaviors differ based on clinical experience of the pathologists. Pathologists with a higher caseload of melanocytic skin biopsy cases and pathologists with board certification and/or fellowship training in dermatopathology have lower average zoom and lower variance of zoom levels. Viewing behaviors associated with higher diagnostic accuracy include higher average and variance of zoom levels, a lower magnification percentage (a measure of consecutive zooming behavior), higher total interpretation time, and higher amount of time spent viewing ROIs. Scanning behavior, which refers to panning with a fixed zoom level, has marginally significant positive association with accuracy. Pathologists' training, clinical experience, and their exposure to a range of cases are associated with their viewing behaviors, which may contribute to their diagnostic accuracy. Research in computational pathology integrating digital imaging and clinical informatics opens up new avenues for leveraging viewing behaviors in medical education and training, potentially improving patient care and the effectiveness of clinical workflow.

1. Introduction

Diagnostic approaches in pathology have recently expanded from glass slides and traditional microscopes to digital pathology.[1] With the introduction of Whole Slide Imaging (WSI) in 1999, it became possible to represent tissue on a glass slide using a high-resolution digital image. Over the following 2 decades, the technology for acquiring virtual slides and their applications in pathology, such as diagnosis, education, and research, has grown exponentially.[2–4] In 2017, digital whole slide imaging for primary diagnosis in pathology was approved by the US Food and Drug Administration (FDA), paving the way for widespread adoption of this technology in everyday practice.[5]

As the field of pathology embraces digital imaging, more cases will be interpreted in the digital format. The interpretation workflow when a pathologist reviews digital whole slide images (WSIs) is different from the traditional use of a microscope.[6] Surprisingly little is known about how pathologists' viewing behaviors while diagnosing relate to their diagnostic assessment. The advent of digital pathology and whole slide imaging has

* Corresponding author at: Paul G. Allen School of Computer Science and Engineering, University of Washington, Box 352350, Seattle, WA 98195, USA.
E-mail addresses: fghezloo@uw.edu (F. Ghezloo), jasonw@mednet.ucla.edu (P.-C. Wang), katiek@uw.edu (K.F. Kerr), tbruny01@tufts.edu (T.T. Brunyé), trafton.drew@psych.utah.edu (T. Drew), ochang@uw.edu (O.H. Chang), lreisch@uw.edu (L.M. Reisch), shapiro@cs.washington.edu (L.G. Shapiro), jelmore@mednet.ucla.edu (J.G. Elmore).

http://dx.doi.org/10.1016/j.jpi.2022.100104
Received 31 January 2022; Received in revised form 7 April 2022; Accepted 9 April 2022
Available online 21 May 2022

made it possible to record and study pathologists' interpretive behaviors. As a pathologist examines a case, the visual search process involves zooming and panning, as well as allocating visual attention to image characteristics. The former can be measured by tracking viewport coordinates of a WSI viewing tool, whereas the latter can be measured using eye-tracking devices. Although eye-tracking techniques are the most direct method of recording viewing behaviors, these techniques are expensive and complex and require in-person data collection. Moreover, processing and analyzing eye-tracking data is time-consuming and requires specialized expertise. In contrast, viewport-tracking data can be obtained remotely, on a large scale, and at a lower cost. Viewport-tracking data provide information on pathologists' movement, zooming behavior, and total interpretation time, offering insight into pathologists' attention.

An early study by Raghunath et al.[6] showed that viewport-tracking data might be used to investigate pathologists' diagnostic accuracy and efficiency when using WSIs on breast tissue. Various studies have been conducted using eye-tracking and viewport-tracking analysis in the fields of radiology and pathology. These studies have identified factors related to the visual search process that are associated with diagnostic accuracy. For instance, adopting a drilling strategy in volumetric images in radiology,[7,8] having fewer eye fixations overall,[9] and spending more time fixating eyes in ROIs[10,11] have been associated with higher diagnostic accuracy. Previous research in radiology[7] and breast pathology[12] demonstrates that 1 of the 2 search strategies is used by physicians when evaluating medical images: drilling or scanning. A pathologist with a drilling strategy focuses on a specific area and uses magnification settings to zoom in and out at various locations, while a pathologist with a scanning strategy utilizes a fixed zoom level while searching and panning over a wide area of interest. While the search strategy used was not a predictor of diagnostic accuracy in breast pathology, some association among pathologists' characteristics and their search strategy was found.[12] Also, using a drilling strategy in searching volumetric images in radiology was associated with correctly localizing more lung nodules.[7] No previous study has investigated the use of these 2 search strategies in dermatopathology. Because of microanatomical differences between organ systems, as well as the pathologies affecting them, it is important to investigate whether or not viewing methodology is associated with diagnostic accuracy.

In this study we outline the types of data that can be gathered to describe pathologists' viewing behavior using viewport tracking data. We then investigate how these behaviors are associated with pathologists' demographics and clinical characteristics. Moreover, we examine the associations among pathologists' viewing behaviors and diagnostic accuracy.

2. Methods

2.1. Study overview

The study design and development of skin biopsy cases have previously been described in detail.[13,14] To briefly summarize, we investigated zooming and panning behaviors of 32 pathologists who each viewed 1of the 5sets of 36 digital melanocytic skin cases (180 total cases). Pathologists viewed and diagnosed their assigned set of cases using a web-based viewer. The recording of these viewing sessions resulted in a total of 1073 interpretations.[1] In addition, for each digital case, we have a consensus reference diagnosis and an ROI representing important features for the diagnosis identified by our reference panel of experienced pathologists. For each interpretation, we used viewport tracking data, which included the location of the viewports, the zoom level used to view the viewports and the timestamps to define several viewing behavior variables. These variables measure and quantify pathologists' interactions with the digital slides such as zooming and panning patterns, total interpretation time, and attending to the consensus ROI selected by the reference panel. We used these variables to investigate the association of the pathologists' viewing behaviors with their characteristics and diagnostic accuracy in melanocytic skin lesions on whole slide images. Our main questions are:

- Are pathologists' demographics and clinical characteristics associated with viewing behaviors?
- Are specific viewing behaviors associated with diagnostic accuracy?

The following is an overview of the case and pathologist selection, the viewport tracking data collection, and the experts' consensus diagnosis.

2.2. Case selection

Skin biopsy specimens of melanocytic lesions (N = 240) were randomly selected from Dermatopathology Northwest in Bellevue, Washington, with stratification based on patient's age and the original diagnosis. To generate digital whole slide images, each glass slide was scanned at 40x magnification with a Hamamatsu NanoZoomer 2.0-RS digital slide scanner.[15] These cases were classified into 5 MPATH-Dx classes (example diagnostic terms): class 1 (nevus/mild atypia), class 2 (moderate atypia/dysplasia), class 3 (severe dysplasia/melanoma in situ), class 4 (stage pT1a invasive melanoma), and class 5 (stage pT1b or higher invasive melanoma). Detailed information about the MPATH-Dx classification can be found elsewhere.[13] For this study, a subset of 180 cases was chosen and divided into 5 different sets of 36 cases, each of which included the entire range of the 5 MPATH-Dx classes. The distribution of these 180 cases among the 5 classes is as follows: 8.3% class 1, 16.7% class 2, 25.0% class 3, 25.0% class 4, and 25.0% class 5.

2.3. Participants

Pathologists were recruited from 10 US states (California, Connecticut, Hawaii, Iowa, Kentucky, Louisiana, New Jersey, New Mexico, Utah, and Washington) to participate in the M-Path study. Pathologists were eligible if they had completed residency and/or fellowship training, had interpreted skin specimens in their clinical practices in the preceding year, and planned to do so for the next 2 years. Pathologists were invited to participate in a substudy of interpreting digital WSIs, and we used data from 32 pathologists from whom we had full tracking data in this investigation.

2.4. Data collection

All participants completed a baseline survey before the study to assess their demographics and clinical practice characteristics (Table 1). A custom web-based digital slide viewer was developed by our research team to meet the needs of this research project using HD View SL, Microsoft's open-source Silverlight platform. The viewer was loaded in Microsoft Internet Explorer (Redmond, WA, USA) and allowed pathologists to pan around the image (using mouse clicks or dragging) and zoom in and out up to x60 magnification (using mouse wheel or buttons), similar to most industry-developed WSI viewers. In addition to zooming and panning, the viewer provided tools for measuring lesion size and counting mitotic figures.[15] Pathologists viewed and interpreted each of the 36 cases one at a time using the online digital slide viewer. As pathologists viewed each slide, the viewer automatically logged a series of viewports in the order they were viewed and exported this information to a data file. A viewport is a rectangular area of the image that is visible on the pathologist's computer screen at any time during their interpretation. The location of a viewport is described with the coordinates (x and y positions) of the upper left corner of the viewport, as well as its dimensions (width and height). In addition to the location of the viewport, the zoom level used at that region and a timestamp were logged. After each interpretation, pathologists were asked to provide their diagnoses using an online histology form. A total number of 1073 interpretations (pathologist and case pairs) were logged and are analyzed in this study.

[1] Viewport tracking data on 79 interpretations were not available.

Table 1
Distribution of pathologists' demographics and characteristics.

Pathologists' demographics and characteristics	Count (Percentage%)
Gender	
Male	13 (41%)
Female	19 (59%)
Age (years)	
20–49	12 (37%)
50–64	20 (63%)
Board certification/ Dermatopathology Fellowship training	
Yes	10 (31%)
No	22 (69%)
Experience with interpreting melanocytic skin lesions (years)	
<5	3 (10%)
5–9	10 (31%)
10–19	9 (28%)
>20	10 (31%)
Caseload of melanocytic skin lesions (%)	
<10	14 (44%)
10–24	13 (41%)
25–49	5 (15%)
Ratings on difficulty level of interpreting melanocytic skin lesions	
1 (Very easy)	–
2	1 (3%)
3	10 (31%)
4	18 (56%)
5	3 (10%)
6 (Very challenging)	–
Ratings on confidence level of interpreting melanocytic skin lesions	
1 (Not at all confident)	–
2	2 (6%)
3	4 (12%)
4	7 (22%)
5	16 (50%)
6 (Extremely confident)	3 (10%)

2.5. Consensus diagnosis and ROI

Three dermatopathologists with expertise in cutaneous melanocytic lesions (our M-Path Study reference panel) independently interpreted the full set of cases in glass slide format and then participated in a series of review meetings. They agreed on a consensus diagnosis for each case using a modified Delphi approach[13] as well as agreeing on a consensus region for each case as the Region of Interest (ROI) which supported their diagnosis and best represented the critical features on the slide. Note that these consensus ROIs were never visible to the participants in this study; they were used only to define critical image regions for data analysis.

2.6. Viewing behaviors

We use the information in our viewport tracking dataset to define variables that quantify pathologists' viewing behaviors. Using the time stamps, we calculated the duration of each viewport being viewed as well as the total interpretation time. Viewports associated with a duration of more than 1 minute (frozen at one location without any activity for more than 1 min) were excluded due to the assumption that the pathologist was not actively interpreting during that time. Variables used to summarize pathologists' viewing behavior are defined in Table 2.

3. Analysis

Both case and pathologist contribute to the variation of the outcome of our models. Due to this crossed-level structure of cases and pathologists in our dataset, we used the cross-classified multilevel model[16] to address our study's questions. To investigate possible associations between pathologists' demographics and clinical characteristics and their viewing behaviors, we used a cross-classified multilevel model. For each model, we used one of the pathologists' characteristics shown in Table 1 as the explanatory variable and each of the viewing behaviors defined in Table 2 as the outcome. The notation of the model is defined in Expression 1; y_i denotes the viewing behavior variable of interpretation i, x_i denotes the pathologist's demographic/clinical characteristic, $u_{pathologist(i)}$ and $u_{case(i)}$ indicate the pathologist and case random effects, and e_i denotes the interpretation-level residual error.

$$y_i = \beta_0 + \beta_1 x_i + u_{pathologist(i)} + u_{case(i)} + e_i \quad (1)$$

To investigate associations between pathologists' viewing behavior and diagnostic accuracy, we used an analogous generalized linear mixed model with logit link. We define diagnostic accuracy as the binary agreement of a pathologist's diagnosis with the consensus reference diagnosis. For each univariate model, we used one of the viewing behaviors defined in Table 2 as the explanatory variable of interest and diagnostic accuracy as the outcome. To control for pathologist experience or expertise, all models also included pathologists' years of experience with melanocytic skin lesions (categorical covariate with 4 levels) and having board certification and/or fellowship training (binary variable). In Expression 2; P_i denotes the probability of an accurate diagnosis for interpretation i, logit(p) = log (p/(1-p)), and x_i denotes the viewing behavior, E_2, E_3, and E_4 are indicators of the second, third, and fourth levels of pathologists' years of experience, and F indicates board certification and/or fellowship training.

$$logit(P_i) = \beta_0 + \beta_1 x_i + \beta_2 E_2 + \beta_3 E_3 + \beta_4 E_4 + \beta_5 F + u_{pathologist(i)} + u_{case(i)} + e_i \quad (2)$$

To further analyze the association of these viewing behaviors with diagnostic accuracy in the presence of each other, a subset of variables was chosen as explanatory variables to study using a multivariate model. Scanning percentage was chosen for the panning behavior, zoom variance for the zooming behavior, ROI time percentage for the interaction with consensus ROI behavior, and total interpretation time to address the diagnostic efficiency. The multivariate model included the same covariates as the univariate models to study accuracy. SAS version 9.4 (SAS Institute, NC) was used to perform all the statistical analyses in this study.

4. Results

4.1. Pathologists' viewing behaviors and characteristics

We hypothesized that pathologists with different characteristics might demonstrate different viewing behavior. To investigate this, we modeled the association between pathologist characteristics (Table 1) and viewing behaviors (Table 2). Table 3 presents all analyses performed and, in this section, we highlight those results with a P-value < 0.1. Pathologists with a board certification and/or fellowship training and those with a higher caseload of melanocytic skin lesions have lower average, maximum, and variance of zoom levels. In addition, pathologists reporting higher confidence in interpreting melanocytic skin lesions have lower average, maximum, and variance of zoom levels. Lastly, older pathologists have higher maximum and variance of zoom levels compared to younger pathologists. No other statistically significant associations were found between viewing behaviors and pathologists' characteristics.

4.2. Diagnostic accuracy

To study associations among viewing behaviors and diagnostic accuracy, we used a series of cross-classified multilevel models. Seven separate models were generated for each of the defined viewing behavior variables in Table 2. The Odds Ratio (OR) and the P-value of each model are shown in Table 4. All viewing behaviors show a statistically significant association with diagnostic accuracy (P-value < 0.05), except for scanning percentage which was marginally significant (0.05 < P-value < 0.1). Except for magnification percentage, each viewing behavior was positively associated with accuracy (adjusted OR >1), meaning that interpretations exhibiting more of the behavior were more likely to yield an accurate diagnosis (Table 4).

Table 2
Pathologist viewing behaviors' definitions.

Viewing behavior	Definition	Equation
Total interpretation time	Using the time stamp (TS) of each viewport (v_i), we calculated the duration (d) of each viewport being viewed. Total interpretation time (T) is calculated by summing the durations of all the viewports.	$d(v_i) = TS(v_{i+1}) - TS(v_i)$ $T = \sum_{v_i=1}^{v_n} d(v_i)$
Average zoom level	The average zoom level (avg) used during an interpretation.	$avg = \frac{\sum_{v_i=1}^{v_n} zoom(v_i)}{n}$
Maximum zoom level	The maximum zoom level (mx) used during an interpretation.	$mx = \max\{zoom(v_i): i = 1..n\}$
Zoom level variance	The zoom level variance (var) during an interpretation.	$var = \frac{\sum_{v_i=1}^{v_n} (zoom(v_i) - avg)^2}{n-1}$
Magnification percentage	Magnification percentage (MP) is calculated based on the number of times a pathologist zooms in consecutively. This variable captures how deeply and frequently a pathologist zooms while interpreting a case. We count the number of viewports that are associated with consecutive zoom-in behavior (m). A consecutive zoom-in is a sequence of viewports where the zoom level of each viewport is greater than its previous viewport in the sequence. Magnification percentage calculates the proportion of viewports associated with this behavior.	$m(v_i) = \begin{cases} 1, & zoom(v_{i+1}) \geq zoom(v_i) \\ 0, & otherwise \end{cases}$ $MP = \frac{\sum_{v_i=1}^{v_n} m(v_i)}{n}$
ROI time percentage	ROI time percentage (RTP) measures the amount of time a pathologist spends viewing regions that experts marked as ROI. When a pathologist's viewport intersects with the consensus ROI by 40% or more, we consider that the pathologist is viewing the consensus ROI (r). However, to ensure that the pathologist is actually attending to the ROI, we apply a size constraint. We exclude cases where the ratio of the ROI area to the viewport area is smaller than 10%. This way we make sure that a viewport intersects with a large area of the ROI and this intersection covers the most parts of the viewport. ROI time percentage calculates the proportion of interpretation time spent viewing such regions. Pathologists were not informed about the consensus ROI at the time of interpretation, spending more time on such regions means that they independently identified the region as important.	$r(v_i) = \begin{cases} 1, & \frac{area(intersect(v_i\ ROI))}{\text{area(ROI)}} > 0.4 \\ & And \\ & \frac{area(ROI)}{area(v_i)} > 0.1 \\ 0, & otherwise \end{cases}$ $RTP = \frac{\sum_{v_i=1}^{v_n} r(v_i)}{n}$
Scanning percentage	Scanning percentage (SP) is defined similar to the E. Mercan et. al[12] study for digital breast pathology. This variable provides information regarding zoom level changes in consecutive log entries, regardless of zoom level itself. We count the number of viewports that are associated with consecutive zoom in behavior (s). Scanning percentage calculates the proportion of viewports associated with panning around the image with a fixed zoom level.	$s(v_i) = \begin{cases} 1, & zoom(v_{i+1}) == zoom(v_i) \\ 0, & otherwise \end{cases}$ $SP = \frac{\sum_{v_i=1}^{v_n} s(v_i)}{n}$

Interpretations with a larger magnification percentage were less likely to yield an accurate diagnosis (adjusted OR <1).

To further investigate the associations between viewing behavior and accuracy in the presence of other confounding factors, we modeled our data using a multivariate cross-classified multilevel model. We selected a subset of predictor variables, including one variable for each of the zooming (zoom variance), panning (scanning percentage), interacting with ROI behaviors (ROI time percentage), and interpretation efficiency (total time), based on the relative strength of odds ratios shown in Table 4. The results from the multivariate model are shown in Table 5. Total interpretation time and ROI time percentage are significantly associated with diagnostic accuracy in this multivariate model (P-value < 0.05), whereas zoom variance and scanning percentage are marginally significant (0.05 < P-value < 0.1).

5. Discussion

This study leveraged WSI viewing behavior data to reveal associations between viewing behavior and pathologist characteristics and diagnostic accuracy. When exploring the former association, we showed that average, maximum and variance of zoom level were negatively associated with pathologists' caseload of melanocytic skin lesions, having board certification and/or fellowship training in dermatopathology, and their confidence level in interpreting melanocytic skin lesions. This means pathologists with these characteristics on average used a lower and limited range of zoom levels. In addition, we found a positive association between pathologists' age and maximum and variance of zoom level.

When investigating the associations among viewing behaviors and diagnostic accuracy, we showed average, maximum, and variance of zoom levels, total interpretation time, and the proportion of interpretation time spent viewing consensus ROIs have positive associations with diagnostic accuracy. Magnification percentage, which measures consecutive zoom-in behaviors, was seen to have a negative association with diagnostic accuracy. In other words, pathologists who performed many consecutive zoom-ins on various image locations were less likely to reach a correct diagnosis. Scanning percentage, which measures the proportion of time spent panning with a fixed zoom level, has a marginally significant positive association

Table 3
Pathologist's characteristics, clinical experience and ratings of difficulty and confidence on melanocytic skin lesions as predictor variables and average zoom, maximum zoom, and zoom variance as outcome variables. Contrast specifies the difference in the mean outcome among a predictor variable's categories. For example, 0.39 indicates that mean of maximum zoom for pathologists in the older age group (50-64 years) was 0.39 magnification level higher than for pathologists in the younger age group (40-49 years). As reflected in Table 1, experience, caseload, difficulty, and confidence rating were analyzed as ordinal variables to investigate trends and thus contrast represents the difference in the mean outcome comparing groups one apart on the ordinal scale.

Pathologists' demographics, clinical characteristics and ratings of difficulty and confidence on melanocytic skin lesions	Average zoom		Maximum zoom		Zoom variance	
	Contrast	P-value	Contrast	P-value	Contrast	P-value
Pathologists' demographics						
Gender (Female vs. Male)	0.03	0.878	0.05	0.825	0.22	0.290
Age (50–64 vs. 20–49)	0.29	0.192	0.39	0.068	0.41	0.038
Clinical experience level						
Board certification or Fellowship training (Yes vs. No)	−0.62	0.003	−0.45	0.037	−0.45	0.030
Experience with melanocytic skin lesions	0.12	0.286	0.12	0.266	0.16	0.100
Caseload of melanocytic skin lesions[a]	−0.35	0.015	−0.29	0.039	−0.31	0.017
Ratings on melanocytic skin lesions						
Difficulty level[a]	−0.05	0.765	−0.03	0.849	−0.09	0.534
Confidence level[a]	−0.21	0.044	−0.19	0.059	−0.17	0.075

[a] Ordinal variable, summarized in Table 1.

Table 4
Each row represents one model with a viewing behavior as the predictor variable and diagnostic accuracy as the outcome. Each model was adjusted for pathologists' years of experience in interpreting melanocytic skin lesions and having board certification and/or fellowship training as covariates. OR stands for Odds Ratio.

Predictor variable	Adjusted OR (95% CI)	P-value
Total interpretation time	1.33 (1.09, 1.62)	0.005
Average zoom	1.26 (1.03, 1.54)	0.023
Maximum zoom	1.24 (1.03, 1.50)	0.026
Zoom variance	1.37 (1.11, 1.68)	0.003
Magnification percentage	0.76 (0.63, 0.92)	0.006
ROI time percentage	1.35 (1.07, 1.69)	0.011
Scanning percentage	1.21 (1.00, 1.47)	0.054

Table 5
Multivariate model with four viewing behaviors as predictor variables, and diagnostic accuracy as the outcome. Each model was adjusted for pathologists' years of experience with interpreting melanocytic skin lesions and having board certification and/ or fellowship training in dermatopathology as covariates. OR stands for Odds Ratio.

Predictor variables	Adjusted OR (95% CI)	P-value
Total interpretation time	1.25 (1.01, 1.54)	0.036
Zoom variance	1.22 (0.98, 1.53)	0.079
ROI time percentage	1.38 (1.10, 1.73)	0.006
Scanning percentage	1.20 (0.98, 1.47)	0.072

with accuracy. Following prior work in radiology which introduced 2 viewing strategies: scanning and drilling[7], we suggest that a pathologist with a higher scanning percentage has adopted a scanning strategy. On the other hand, a pathologist with a lower scanning percentage is most likely adopting a relative zooming (drilling) strategy in their interpretation. However, more recent evidence from digital pathology suggests that these 2 strategies are not mutually exclusive.[17]

The association between time spent viewing the consensus ROI and diagnostic accuracy highlights the importance of detecting critical image regions, deeming them worthy of interrogation, and gaining high-power views of histopathological features in these regions. As digital WSI and computer-aided diagnostic (CAD) tools continue to pervade training and clinical practice, we believe this result can be leveraged in future research and development. For example, given the relatively strong association between time spent examining the ROI and diagnostic accuracy, these specific regions can be used to train CAD and artificial intelligence (AI) algorithms on the histopathological features critical to enabling accurate diagnoses. A few computer models have been previously developed based on pathologists' viewing behavior while diagnosing breast histopathology images.[18,19] Future adaptive tutoring systems can also monitor trainee viewing behavior and adaptively guide novice pathologists towards these features, helping them discover the most critical image regions for deriving an accurate diagnosis.

With FDA approval, using digital pathology is becoming an essential part of daily practice of pathology. As a result, digital whole slide imaging has the potential to alter practically every area of the clinical workflow, teaching and education, and research. We hypothesize that the rich depth of new data becoming available since the advent of WSI will open up future studies that might use this information in teaching and evaluation. In this paper we outlined methods of collecting data remotely on pathologists' viewing behaviors. Given the wide range of pathologists' interpretations and diagnoses of complex melanocytic lesions, studying pathologists' viewing behaviors and interpretive strategies might be beneficial in many areas, and specifically in education. The WSI has caused a significant shift in thinking about education in histology and pathology, allowing for the introduction of previously impossible activities and skills.[20] It has been demonstrated that students are comfortable using WSI because they have prior computer experience. As a result, learners become accustomed to the tool quickly and can focus on the histological aspects of the slides without having to learn how to use the microscope.[20] Moreover, a whole slide image can be accessed and used in a more flexible setting from any device or location, requiring only a computer with internet access.

Despite the fact that using WSIs in digital pathology has many benefits, there are various challenges in obtaining such technology and procedures. Each step of high-quality pathology slide preparation, including embedding, cutting, staining, and scanning, is critical to the successful adaptation of whole-slide images in digital pathology.[21] The methods and operational quality controls must be standardized to reduce system mistakes and random errors, because a single noise in huge data might cause misclassification of the case. Therefore, acquisition of high-quality scanners and staff to manage the complete WSI system is costly. High-capacity servers are needed for storage and distribution purposes. Moreover, numerous technological and ethical issues must be resolved when allowing clinical teams to share and analyze imaging data and patient information across a larger platform. Besides the technical challenges, the small experimental sample sizes in pathology studies may limit generalizability and introduce challenges to the statistical analysis.

6. Conclusions

Diagnosis of pathology slides is a complex task, and pathologists go through years of training to be able to view image data and make a diagnosis. Even experienced pathologists are prone to uncertainty and errors when confronted with massive amounts of information in cases. To gain a better understanding of how pathologists view the complex image data and reach a diagnosis, it is essential to understand their interpretive strategies and viewing behaviors. The advent of digital pathology and whole slide imaging has made it possible to record and study pathologists' interpretive behaviors during their medical decision-making process in a novel way. In this study we examined various summaries of pathologists' viewing behaviors such as zooming, panning, and interacting with the consensus ROI. We investigated the association of these viewing behaviors with pathologists' characteristics and diagnostic accuracy. The results of such behavioral studies can be beneficial in multiple areas,[22] such as improving the training and education of younger pathologists, determining the reasons for diagnostic errors to enhance pathologists' performance, and assisting with the development of computer-aided and AI tools for diagnosis purposes.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Research reported in this study was supported by the National Cancer Institute under Awards No. R01 CA15130, R01 CA225585, and R01 CA201376 and the Office of the Assistant Secretary of Defense for Health Affairs through the Melanoma Research Program under Awards No. W81XWH-20-1-0797 and W81XWH-20-1-0798. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the funders. We thank the study participants for their commitment to improving clinical care in dermatopathology, as well as the University of Washington CME Office.

References

1. Rigel DS, Russak J, Friedman R. The evolution of melanoma diagnosis: 25 years beyond the ABCDs. CA Cancer J Clin 2010;60(5):301–316. https://doi.org/10.3322/caac.20074.
2. Jahn SW, Plass M, Moinfar F Digital pathology: advantages, limitations and emerging perspectives. J Clin Med 2020;9(11):3697. https://doi.org/10.3390/jcm9113697.

3. Kumar N, Gupta R, Gupta S Whole slide imaging (WSI) in pathology: current perspectives and future directions. J Digit Imag 2020;33(4):1034–1040. https://doi.org/10.1007/s10278-020-00351-z.
4. Pantanowitz L, Sharma A, Carter AB, Kurc T, Sussman ASaltz J A, Saltz J. Twenty years of digital pathology: an overview of the road travelled, what is on the horizon, and the emergence of vendor-neutral archives. J Pathol Inform 2018;9(1):40. https://doi.org/10.4103/jpi.jpi_69_18.
5. Boyce B. An update on the validation of whole slide imaging systems following FDA approval of a system for a routine pathology diagnostic service in the United States. Biotechn Histochem 2017;92(6):381–389. https://doi.org/10.1080/10520295.2017.1355476.
6. Raghunath V, Braxton MO, Gagnon SA, et al. Mouse cursor movement and eye tracking data as an indicator of pathologists' attention when viewing digital whole slide images. J Pathol Inform 2012;3(1):43. https://doi.org/10.4103/2153-3539.104905.
7. Drew T, Vo ML-H, Olwal A, et al Scanners and drillers: characterizing expert visual search through volumetric images. J Vision 2013;13(10):3. https://doi.org/10.1167/13.10.3.
8. Wen G, Aizenman A, Drew T, et al Computational assessment of visual search strategies in volumetric medical images. J Med Imag 2016;3(1), 015501. https://doi.org/10.1117/1.JMI.3.1.015501.
9. Krupinski EA, Graham A, Weinstein RS Characterizing the development of visual search expertise in pathology residents viewing whole slide images. Human Pathol 2013;44(3):357–364. https://doi.org/10.1016/j.humpath.2012.05.024.
10. Brunye TT, Carney PA, Allison KH, Shapiro LG, Weaver DL, Elmore JG. Eye movements as an index of pathologist visual expertise: a pilot study. PLoS One 2014;9(8), e103447. https://doi.org/10.1371/journal.pone.0103447.
11. 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. Human Pathol 2006;37(12):1543–1556. https://doi.org/10.1016/j.humpath.2006.08.024.
12. Mercan E, Shapiro LG, Brunyé TT, Weaver LD, Elmore JG Characterizing diagnostic search patterns in digital breast pathology: scanners and drillers. J Digit Imag 2018;31(1):32–41. https://doi.org/10.1007/s10278-017-9990-5.
13. Carney PA, Reisch LM, Piepkorn MW, et al. Achieving consensus for the histopathologic diagnosis of melanocytic lesions: use of the modified Delphi method. J Cutan Pathol 2016;43(10):830–837. https://doi.org/10.1111/cup.12751.
14. Piepkorn MW, Barnhill RL, Elder DE, et al. The MPATH-Dx reporting schema for melanocytic proliferations and melanoma. J Am Acad Dermatol 2014;70(1):131–141. https://doi.org/10.1016/j.jaad.2013.07.027.
15. Onega T, Barnhill RL, Piepkorn MW, et al. Accuracy of digital pathologic analysis vs traditional microscopy in the interpretation of melanocytic lesions. JAMA Dermatol 2018;154(10):1159–1166. https://doi.org/10.1001/jamadermatol.2018.2388.
16. Browne WJ, Goldstein H, Rasbash J Multiple membership multiple classification (MMMC) models. Stat Model 2001;1(2):103–124. https://doi.org/10.1177/1471082X0100100202.
17. Drew T, Lavelle M, Kerr KF, et al. More scanning, but not zooming, is associated with diagnostic accuracy in evaluating digital breast pathology slides. J Vision 2021;21(11):7. https://doi.org/10.1167/jov.21.11.7.
18. Mercan E, Aksoy S, Shapiro LG, Weaver DL, Brunye T, Elmore JG. Localization of diagnostically relevant regions of interest in whole slide images. 2014 22nd International Conference on Pattern Recognition. IEEE; 2014.
19. Mercan C, Mercan E, Aksoy S, Shapiro LG, Weaver DL, Elmore JG Multi-instance multi-label learning for whole slide breast histopathology. Medical Imaging 2016: Digital Pathology. SPIE; 2016.
20. Saco A, Bombi JA, Garcia A, Ramírez J, Ordi J Current status of whole-slide imaging in education. Pathobiology 2016;83(2–3):79–88. https://doi.org/10.1159/000442391.
21. Cui M, Zhang DY. Artificial intelligence and computational pathology. Lab Investig 2021;101(4):412–422. https://doi.org/10.1038/s41374-020-00514-0.
22. Brunyé TT, Drew T, Weaver DL, Elmore JG A review of eye tracking for understanding and improving diagnostic interpretation. Cognit Res Princ Implicat 2019;4(1):1-16. https://doi.org/10.1186/s41235-019-0159-2.