Title

Are Pathologists Self-Aware of Their Diagnostic Accuracy? Metacognition and the Diagnostic Process in Pathology

Journal

Medical Decision Making, 43(2)

ISSN

0272-989X

Authors

Clayton, Dayna A
Eguchi, Megan M
Kerr, Kathleen F
et al.

Publication Date

2023-02-01

DOI

10.1177/0272989x221126528

Peer reviewed

HHS Public Access

Author manuscript

Med Decis Making. Author manuscript; available in PMC 2023 February 01.

Published in final edited form as:

Med Decis Making. 2023 February ; 43(2): 164–174. doi:10.1177/0272989X221126528.

Are Pathologists Self-Aware of their Diagnostic Accuracy? Metacognition and the Diagnostic Process in Pathology

Dayna A. Clayton[1], **Megan M. Eguchi**[1], **Kathleen F. Kerr**[2], **Kiyofumi Miyoshi**[3], **Tad T. Brunyé**[4], **Trafton Drew**[5], **Donald L. Weaver**[6], **Joann G. Elmore**[1,*]

[1]Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, United States of America

[2]Department of Biostatistics, University of Washington, Seattle, WA, United States of America

[3]Department of Psychology, University of California, Los Angeles, CA, United States of America

[4]Center for Applied Brain and Cognitive Sciences, Tufts University, Medford, MA, United States of America

[5]Department of Psychology, University of Utah, Salt Lake City, UT, United States of America

[6]Department of Pathology & Laboratory Medicine, University of Vermont Larner College of Medicine, Burlington, VT

Abstract

Background: Metacognition is a cognitive process that involves self-awareness of thinking, understanding, and performance. This study assesses pathologists' metacognition by examining the association between their diagnostic accuracy and self-reported confidence levels while interpreting skin and breast biopsies.

Design: We studied 187 pathologists from the Melanoma Pathology Study (M-Path) and 115 pathologists from the Breast Pathology Study (B-Path). We measured pathologists' metacognitive ability by examining the AUC, the area under each pathologist's receiver operating characteristic (ROC) curve summarizing the association between confidence and diagnostic accuracy. We investigated possible relationships between this AUC measure, referred to as metacognitive sensitivity, and pathologist attributes. We also assessed whether higher metacognitive sensitivity affected the association between diagnostic accuracy and a secondary diagnostic action such as requesting a second opinion.

Results: We found no significant associations between pathologist clinical attributes and metacognitive AUC. However, we found that pathologists with higher AUC showed a stronger trend to request secondary diagnostic action for inaccurate diagnoses and not for accurate diagnoses compared to pathologists with lower AUC.

Limitations: Pathologists reported confidence in specific diagnostic terms, rather than the broader classes into which the diagnostic terms were later grouped to determine accuracy.

[*]Corresponding author: **Joann G. Elmore, MD, MPH**, Department of General Internal Medicine, David Geffen School of Medicine, University of California, Los Angeles, 1100 Glendon Ave., Suite 900, Los Angeles, CA, 90024, jelmore@mednet.ucla.edu.

Additionally, while there is no “gold standard” for the correct diagnosis to determine accuracy of pathologists’ interpretations, our studies achieved a high-quality reference diagnosis by using the consensus diagnosis of three experienced pathologists.

Conclusions: Metacognition can affect clinical decisions. If pathologists have self-awareness that their diagnosis may be inaccurate, they can request additional tests or second opinions, providing the opportunity to correct inaccurate diagnoses.

Keywords

metacognitive sensitivity; diagnostic accuracy; cognitive science; dermatopathology; breast pathology; secondary diagnostic actions; awareness

INTRODUCTION

Physicians routinely experience uncertainty during diagnostic decision-making, and some physicians might be more sensitive to the subjective experience of uncertainty than others [1]. The cognitive function allowing one to think about their own decisions in this manner is termed *metacognition* [2]. One important aspect of metacognition is the ability to rate correct judgments with higher confidence than incorrect judgments. The influence of metacognition on memory, decision-making, and learning has profound implications for the diagnostic process. In diagnostic pathology, pathologists match the histopathological attributes of a case to a diagnostic category. Efforts to improve diagnostic decision-making require a comprehensive understanding of successful and unsuccessful decision-making [3].

Davidson et al. outlined four ways in which metacognition contributes to problem solving: identifying the problem, mentally representing the problem, planning how to proceed, and evaluating one’s performance [4]. This study investigated the fourth process: evaluating one’s performance. This fourth process has been called “metacognitive sensitivity”, the extent to which confidence is associated with accuracy [5].

Physicians’ correctness and their self-reported confidence levels tend to be moderately associated [6]; that is, physicians tend to provide higher confidence ratings to cases that they correctly judge, and lower confidence ratings to cases they incorrectly judge. However, metacognitive abilities vary widely across individuals [7–9]. To our knowledge, no research has specifically examined whether metacognitive sensitivity is related to pathologists’ experience level and their tendency to seek additional information to help disambiguate a medical decision. To address this gap, this study measures physicians’ metacognitive sensitivity by assessing the association between pathologists’ self-reported confidence and their diagnostic accuracy, and investigates how it relates to pathologists’ characteristics and secondary diagnostic requests. We implemented the absolute metacognitive sensitivity methodology [10] to measure metacognitive sensitivity for each participating pathologist, using confidence ratings to measure pathologists’ feeling-of-knowing in relation to their actual performance diagnosing cases [5].

We hypothesized that pathologist attributes such as age, expertise, caseload, and years of experience would be positively associated with their metacognitive sensitivity [7, 11–

13]. We further evaluated associations between metacognitive sensitivity and pathologists' self-reported secondary diagnostic behaviors: asking for a second opinion to improve diagnostic accuracy, asking for special stains or ancillary tests, or considering a diagnosis to be borderline between two diagnostic categories. Given that metacognitive sensitivity helps people recognize differences in the quality or quantity of information available during diagnostic decision-making and prompts information-seeking behavior[14, 15], we hypothesized that pathologists would be more likely to utilize secondary diagnostic actions for inaccurate interpretations than accurate interpretations, and that the association would be stronger for pathologists with higher metacognitive sensitivity.

METHODS

Data Sources

We used data from the Melanoma Pathology Study (M-Path) and Breast Pathology Study (B-Path), which we analyzed separately throughout [16, 17]. In both studies, each participant interpreted one slide set of 48 cases (M-Path) or 60 cases (B-Path). Slide sets were mailed to participants sequentially, and participants returned slide sets after they interpreted all cases using their own microscopes. Participants based their diagnosis off one hematoxylin and eosin-stained slide per case; additional slides or other stains were not available during the study. Participants entered their diagnostic interpretations and other case assessments into an online histology form.

The research team mapped diagnoses into one of four diagnostic classes using the MPATH-Dx schema [18] for the M-Path study, and the BPATH-Dx schema [19] for the B-Path study. Although MPATH-Dx is a 5-class schema, we merged classes I and II due to prior observations of low accuracy differentiating between these classes and little clinical difference between these benign classes [16]. Each case interpretation by a study participant was assessed as accurate or inaccurate based on whether it mapped to the same diagnostic class as the expert consensus diagnosis (described below).

M-Path Study—To reach expert consensus for cases used in the M-Path study, three dermatopathologists with recognized expertise in melanocytic lesions served as the reference panel. They first independently assessed cases, then held a series of consensus meetings and used a modified Delphi method to reach a consensus reference diagnosis on 240 melanocytic skin lesions [16, 20, 21]. The 240 cases were arranged into five slide sets of 48 cases, balanced by MPATH-Dx class.

The 187 M-Path study participants were from ten geographically diverse US states. Participants were eligible if they had completed pathology training (residency and/or fellowship), interpreted melanocytic skin biopsies within the previous year, and planned to continue interpreting cutaneous melanocytic lesions for at least two years. Participating pathologists completed a baseline survey regarding their demographic and practice characteristics, training and experience, and then proceeded to case interpretations. Participants could select from more than 50 diagnoses for each case. The study team later mapped each diagnosis to the appropriate MPath-Dx diagnostic class, as described above, for assessment of accuracy.

B-Path Study—In the B-Path study, three breast pathologists with recognized expertise followed methods similar to the M-Path study to reach consensus on 240 breast biopsy cases [17]. Invasive breast cancer, ductal carcinoma in situ (DCIS), atypical hyperplasia, and benign cases without atypia were included. The 240 cases were randomly assigned to four sets of 60 cases stratified by the consensus reference diagnosis, difficulty rating, breast density, and patient age.

The 115 B-Path participants were from eight geographically diverse US States. Eligibility criteria included completion of residency, experience interpreting breast lesions for at least 1 year prior, and planning to interpret breast lesions for the following two years. As with M-Path, participating pathologists completed a baseline survey and then proceeded to case interpretations.

Outcome Measures

Metacognitive sensitivity—Participants provided a confidence rating in their assessment of each case on a six-point scale from 1 (Not at all confident) to 6 (Very confident). Applying established methods [5], we measured metacognitive sensitivity by assessing the relationship between confidence ratings and diagnostic accuracy. To do so, we plotted receiver operating characteristic (ROC) curves for each participant: at each possible cut-off on the 6-point confidence scale, the hit rate and the false alarm rate across a participant's diagnostic interpretations were plotted. The hit rate measures how often the pathologist reports high confidence among their diagnostically accurate interpretations, whereas the false alarm rate measures how often the pathologist reports high confidence among their inaccurate interpretations. AUC, the area under the ROC curve, is our estimate of the participant's metacognitive sensitivity. AUC values range from 0 to 1, with higher values indicating greater metacognitive sensitivity. AUC equal to 0.5 indicates metacognitive sensitivity that is no different from chance. This method of estimating metacognitive sensitivity is not affected by differences in participants' use of the confidence scale (e.g., some participants tend to be more or less confident) and does not rely on any distributional assumptions about the data [5].

Clinical impact of metacognitive sensitivity—Participants were asked at the time of each interpretation whether they would have pursued the following secondary diagnostic actions for each case: obtain second opinions, request special stains or ancillary tests (M-Path study only), or report their diagnosis as borderline between two diagnoses. Of note, participants were not actually able to obtain second opinions or additional stains or tests to assist with their interpretations during the study. We assessed associations between diagnostic accuracy and participants' indications that they would pursue these secondary diagnostic actions, specifically investigating whether associations were stronger for participants with higher AUCs.

Data Analysis

Differences in metacognitive sensitivity—We provide descriptive statistics for AUCs across participants in each study. To gain insight on whether AUCs differed from the distribution of AUCs with no metacognitive sensitivity, we developed a reference

distribution of AUCs by creating 2000 permuted datasets. In each permuted dataset, we randomly permuted each pathologist's confidence ratings across all his/her interpretations. We calculated the AUC for each pathologist using these permuted data, then calculated the mean AUC across pathologists. This simulated a null distribution of 2000 mean AUC values when all pathologists have null metacognitive sensitivity.

Pathologist characteristics—The pathologist characteristics included age, gender, years of experience, percent of caseload spent interpreting the specific case type (melanocytic skin lesions for the M-Path study and breast pathology cases for the B-Path study), and expertise in the subspecialty. For M-Path participants, expertise was defined as having board certification or fellowship training in dermatopathology. For B-Path participants, expertise was defined as having fellowship training in breast pathology, or self-report that peers considered them to be experts in breast pathology. Gender was self-reported with the options of male or female, and was studied for a possible association with metacognitive sensitivity because of a difference in average confidence between male and female medical students noted in previous research.[22, 23] Beyond years of experience, chronologic age itself may be associated with metacognition and was included in the analysis.[24]

Analysis Plan—Individual pathologists were the unit of analysis. We treated the AUC measure of metacognitive sensitivity as a continuous variable and performed linear regression to investigate associations with different factors, controlling for case test set.

To assess the relationship between metacognitive sensitivity and utilization of secondary diagnostic actions, we used a logistic regression model with a secondary diagnostic action requested on a case as the outcome. Variables in the model were accuracy of the interpretation (binary), AUC as the measure of absolute metacognitive sensitivity (continuous), and the two-way interaction between accuracy and AUC. We used cluster robust standard error estimates to account for interpretations by the same pathologist. We analyzed each secondary diagnostic action separately. Although AUC was analyzed as a continuous measure, to help describe results we present the model fit for example values of AUC (AUC=0.5, 0.65, and 0.8).

SAS 9.4 (SAS Institute, Cary, NC) was used for analysis and statistical significance was evaluated at a threshold of $p<0.05$.

RESULTS

Characteristics of pathologists are presented in Table 1. Of the 187 M-Path pathologists, 74 (40%) were experts in dermatopathology. Of the 115 B-Path pathologists, 27 (23%) were experts in breast pathology. Most participants had been interpreting the relevant type of cases for 10 or more years (60% in M-Path and 61% in B-Path).

Example ROC curves are plotted in Figure 1 for two M-Path study participants, showing one pathologist with high metacognitive sensitivity (AUC 0.83) and a second pathologist with near-chance metacognitive sensitivity (AUC 0.48), i.e. who was similarly confident in accurate and inaccurate diagnoses.

In the M-Path study, the mean AUC across all 187 pathologists was 0.64 (range 0.37 to 0.87). In the B-Path study, the mean AUC across all 115 pathologists was 0.66 (range 0.47 to 0.85). We found significant evidence from both studies that pathologists have positive metacognitive sensitivity on average ($p<0.005$). Figure 2 compares observed AUC values with the distribution expected under a null condition. Mean AUC was similar across subgroups of pathologists (Table 2). For example, mean AUC was similar across demographic subgroups and subgroups defined by pathologists' overall self-reported confidence in their diagnostic assessments. Notably, there was no evidence of an association between metacognitive sensitivity and self-reported baseline survey data on clinical expertise or level of overall confidence in diagnosing these types of biopsies, as detailed in Table 2.

Figure 3 displays the results of the analyses examining interactions between AUC and diagnostic accuracy when the pathologist considered the diagnosis to be borderline, or in the use of secondary diagnostic actions such as requesting second opinions. The odds ratios (OR) comparing the utilization of each action for inaccurate versus accurate interpretations are shown at three selected values of AUC: 0.5, 0.65, and 0.8, representing null, average, and high AUC values in our study. In the M-Path study, interaction terms between AUC and accuracy were significant for asking for a second opinion, ordering special stains or tests, and considering the diagnosis to be borderline between two different diagnoses. Participants with relatively high metacognitive sensitivity were much more likely to request a secondary diagnostic action for inaccurate diagnoses compared to accurate diagnoses. For example, according to the fitted model, M-Path participants with AUC=0.8 have 4.59-fold higher odds to ask for a second opinion for an inaccurate diagnosis than an accurate diagnosis (95% CI: 3.78–5.56) and 2.47-fold higher odds to order special stains or ancillary tests (95% CI: 2.00–3.05) for an inaccurate diagnosis than an accurate diagnosis. In contrast, participants with null metacognitive sensitivity (AUC=0.5) were not significantly more likely to take these actions for inaccurate vs. accurate diagnoses, with odds ratios 1.18 (95% CI: 0.97–1.45) and 1.08 (95% CI: 0.89–1.32). Odds ratios at AUC=0.65 were between those found at AUC=0.5 and AUC=0.8.

Results were very similar in the B-Path Study. Highly metacognitive sensitive participants with AUC 0.8 are estimated to have 4.11-fold higher odds of asking for a second opinion for an inaccurate diagnosis than an accurate diagnosis (95% CI: 3.19–5.32). Participants with AUC=0.50 are also estimated to be more likely to ask for a second opinion for inaccurate vs. accurate diagnosis, but the association between accuracy and this action is much weaker (OR 1.57, 95% CI: 1.25–1.97). Again, odds ratios at AUC=0.65 were between those for AUC=0.5 and AUC=0.8.

Finally, in both studies participants with high metacognitive sensitivity more often considered inaccurate diagnoses to be borderline between two diagnoses than accurate diagnoses, with a stronger association between borderline determinations and accuracy among highly metacognitive-sensitive participants (M-Path study OR 5.20, 95% CI: 4.19–6.46; B-Path study OR 5.17, 95% CI: 4.31–6.21) than low metacognitively sensitive participants (M-Path study OR 1.45, 95% CI: 1.18–1.80; B-Path study OR 1.35, 95% CI: 1.03–1.77).

DISCUSSION

This study investigated pathologists' metacognitive sensitivity using data from two independent studies in pathology, one in skin pathology and one in breast pathology. These tissues have very different histologic features, yet findings are strikingly similar in the two studies. We found that metacognitive sensitivity differs across pathologists, and was not demonstrably better than chance for some pathologists. No significant relationships were observed between AUC and pathologist characteristics such as age, gender, and level of clinical experience.

In both skin and breast pathology, we found that participants with higher AUCs tended to request secondary diagnostic actions for cases where they were inaccurate and not for cases where they were accurate, and this tendency was weaker or null for participants with lower AUCs. To assess the possibility that these results are confounded by accuracy, we verified that there was essentially no correlation between accuracy and AUC in each dataset (Online Resource Figure 1). This result extends research in developmental psychology and perceptual decision-making [14, 15] to the high-stakes domain of diagnostic medicine, showing that subjective confidence can be a valuable predictor of information-seeking behavior in pathology.

The finding that pathologists with higher AUC values showed a stronger association between secondary diagnostic actions and accuracy is consistent with research on the relationship between physician confidence and the correctness of diagnoses. In a prior study, clinicians' confidence in their diagnosis was defined as the probability of seeking assistance at the time of generating a differential diagnosis. Friedman et al. found that when clinicians reported low confidence, they were likely to be incorrect [6]. Thus, self-awareness of their uncertainty meant they were more likely to seek assistance, which in turn can provide opportunities to scrutinize decisions, and learn from mistakes [25]. Similarly, our study found evidence that metacognitive sensitivity can affect diagnostic outcomes and patient care in pathology. If pathologists can suspect when they are inaccurate, they can pursue actions that may improve diagnostic accuracy, such as seeking second-opinions [26–29].

The lack of association between AUC and expertise suggests that metacognitive sensitivity in these situations was not a function of training or experience. AUC is not simply a reflection of accuracy, instead it measures how well an individual can self-evaluate transient states of uncertainty during the interpretive process. Current pathology clinical training and practice does not seem to enhance metacognitive sensitivity. This raises a potential issue surrounding pathologists receiving feedback on their diagnostic accuracy in a timely manner. Previous research demonstrates the important role of feedback, especially immediate feedback, in improving metacognitive judgements in the context of both everyday decisions [30–34] and medical decision-making [35–38]. However, pathologists are often the final diagnostician in clinical practice, and therefore might receive little to no feedback compared to physicians in other fields, and any feedback based on patient outcome might be delayed by months or years.

The standard signal detection theory approach pertains to a binary choice task distinguishing between two alternatives on a continuous scale. In contrast, this study used four clinically meaningful diagnostic classes defined by the MPATH-Dx and BPATH-Dx diagnostic schema. We consider this a study strength in that we did not oversimplify into a binary classification problem. However, we note the deviation from the standard framework. This research also has several limitations. First, the measurement of pathologists' confidence ratings on the histology form inquired about their confidence in the specific diagnostic term they selected (e.g., type of melanoma in situ such as "lentiginous") and not in the overall diagnostic class (e.g., Class III or Class IV), which was assigned later using the MPATH-Dx or BPATH-Dx schema. Pathologists might not be confident of a very specific diagnosis, yet be highly confident the case belongs to a particular class of diagnoses (e.g., M-PATH-Dx Class IV). Third, there is no "gold standard" reference diagnosis of skin and breast cases. The two studies each used a panel of three experienced pathologists who agreed upon a single consensus reference diagnosis, which in turn defined participants' diagnostic accuracy. While the underlying biology and patient outcomes (e.g., death from cancer; recurrence) may be more ideal gold standards, they also have limitations: for example, the treatment for disease, including surgical removal of tissue (including the initial biopsy), alters the clinical course. Finally, we were limited to studying pathologist attributes that were included in the studies' baseline surveys; additional research could examine other clinical attributes, and even pathologist personality or psychological traits, such as self-regulation and executive control techniques [39–41].

This study found evidence that metacognitive sensitivity differs across pathologists and can affect clinical decisions. No significant associations were found between metacognitive sensitivity and pathologist characteristics, including training and expertise. While current clinical training and practice improves accuracy, [42] it does not appear to improve metacognitive sensitivity. Although there is some evidence that metacognition can be improved through deliberately directed training, this prior work [43] was in a student population and it is therefore an open question whether similar training techniques would be effective with clinicians. Pathologists might benefit from receiving immediate feedback on their level of accuracy (such as in CME), since feedback on diagnostic accuracy in clinical practice in pathology is often completely absent or substantially delayed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support for this study was provided in part by grants from The National Institute of Health (NIH) Diversity Supplement Grant 5R01CA225585-03 and the National Cancer Institute (R01 CA151306, R01 CA140560). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

References

[1]. Bhise V, Rajan SS, Sittig DF, Morgan RO, Chaudhary P, Singh H. Defining and Measuring Diagnostic Uncertainty in Medicine: A Systematic Review. J Gen Intern Med. 2018; 33(1):103–15. [PubMed: 28936618]

[2]. Nelson TO, Stuart RB, Howard C, Crowley M. Metacognition and clinical psychology: a preliminary framework for research and practice. Clinical Psychology & Psychotherapy. 1999; 6(2):73–9.

[3]. Foucar E. Diagnostic Decision-Making in Anatomic Pathology. Pathology Patterns Reviews. 2001; 116(suppl_1):S21–S33. [PubMed: 11993700]

[4]. Davidson JE, Deuser R, Sternberg RJ. The role of metacognition in problem solving. In: Metcalfe J, Shimamura AP, eds. Metacognition: Knowing about knowing: The MIT Press, Cambridge, MA 1994:207–26, Chapter xiii, 334 Pages.

[5]. Fleming SM, Lau HC. How to measure metacognition. Frontiers in Human Neuroscience. 2014; 8(443).

[6]. Friedman CP, Gatti GG, Franz TM, Murphy GC, Wolf FM, Heckerling PS, et al. Do physicians know when their diagnoses are correct? Journal of General Internal Medicine. 2005; 20(4):334–9. [PubMed: 15857490]

[7]. Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The Causes of Errors in Clinical Reasoning: Cognitive Biases, Knowledge Deficits, and Dual Process Thinking. Acad Med. 2017; 92(1):23–30. [PubMed: 27782919]

[8]. Nelson TO. Metamemory: A Theoretical Framework and New Findings. In: Bower GH, ed. Psychology of Learning and Motivation: Academic Press 1990:125–73.

[9]. Metcalfe J, Shimamura AP. Metacognition : knowing about knowing. Cambridge, Mass.: MIT Press 1994.

[10]. Maniscalco B, Lau H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Consciousness and Cognition. 2012; 21(1):422–30. [PubMed: 22071269]

[11]. Sherbino J, Dore KL, Wood TJ, Young ME, Gaissmaier W, Kreuger S, et al. The relationship between response time and diagnostic accuracy. Acad Med. 2012; 87(6):785–91. [PubMed: 22534592]

[12]. Weber EU, Böckenholt U, Hilton DJ, Wallace B. Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. J Exp Psychol Learn Mem Cogn. 1993; 19(5):1151–64. [PubMed: 8409853]

[13]. Groves M, O'Rourke P, Alexander H. Clinical reasoning: the relative contribution of identification, interpretation and hypothesis errors to misdiagnosis. Med Teach. 2003; 25(6):621–5. [PubMed: 15369910]

[14]. Desender K, Boldt A, Yeung N. Subjective Confidence Predicts Information Seeking in Decision Making. Psychol Sci. 2018; 29(5):761–78. [PubMed: 29608411]

[15]. Nussenbaum K, Cohen AO, Davis ZJ, Halpern DJ, Gureckis TM, Hartley CA. Causal Information-Seeking Strategies Change Across Childhood and Adolescence. Cogn Sci. 2020; 44(9):e12888. [PubMed: 32882077]

[16]. Elmore JG, Barnhill RL, Elder DE, Longton GM, Pepe MS, Reisch LM, et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. Bmj. 2017; 357:j2813. [PubMed: 28659278]

[17]. Elmore JG, Longton GM, Carney PA, Geller BM, Onega T, Tosteson AN, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA. 2015; 313(11):1122–32. [PubMed: 25781441]

[18]. Piepkorn MW, Barnhill RL, Elder DE, Knezevich SR, Carney PA, Reisch LM, et al. The MPATH-Dx reporting schema for melanocytic proliferations and melanoma. Journal of the American Academy of Dermatology. 2014; 70(1):131–41. [PubMed: 24176521]

[19]. Allison KH, Reisch LM, Carney PA, Weaver DL, Schnitt SJ, O'Malley FP, et al. Understanding diagnostic variability in breast pathology: lessons learned from an expert consensus review panel. Histopathology. 2014; 65(2):240–51. [PubMed: 24511905]

[20]. Dalkey N, Brown B, Cochran S. The Delphi method, III: Use of self-ratings to improve group estimates: RAND CORP SANTA MONICA CALIF; 1969.

[21]. Carney PA, Reisch LM, Piepkorn MW, Barnhill RL, Elder DE, Knezevich S, et al. Achieving consensus for the histopathologic diagnosis of melanocytic lesions: use of the modified Delphi method. Journal of Cutaneous Pathology. 2016; 43(10):830–7. [PubMed: 27247109]

[22]. Blanch DC, Hall JA, Roter DL, Frankel RM. Medical student gender and issues of confidence. Patient Educ Couns. 2008; 72(3):374–81. [PubMed: 18656322]

[23]. Blanch-Hartigan D. Medical students' self-assessment of performance: results from three meta-analyses. Patient Educ Couns. 2011; 84(1):3–9. [PubMed: 20708898]

[24]. Metcalfe J, Eich TS, Castel AD. Metacognition of agency across the lifespan. Cognition. 2010; 116(2):267–82. [PubMed: 20570251]

[25]. Cutting MF, Saks NS. Twelve tips for utilizing principles of learning to support medical education. Medical Teacher. 2012; 34(1):20–4. [PubMed: 21777127]

[26]. Elmore JG, Tosteson ANA, Pepe MS, Longton GM, Nelson HD, Geller B, et al. Evaluation of 12 strategies for obtaining second opinions to improve interpretation of breast histopathology: simulation study. BMJ. 2016; 353:i3069. [PubMed: 27334105]

[27]. Middleton LP, Feeley TW, Albright HW, Walters R, Hamilton SH. Second-Opinion Pathologic Review Is a Patient Safety Mechanism That Helps Reduce Error and Decrease Waste. Journal of Oncology Practice. 2014; 10(4):275–80. [PubMed: 24695900]

[28]. Piepkorn MW, Longton GM, Reisch LM, Elder DE, Pepe MS, Kerr KF, et al. Assessment of Second-Opinion Strategies for Diagnoses of Cutaneous Melanocytic Lesions. JAMA Network Open. 2019; 2(10):e1912597–e. [PubMed: 31603483]

[29]. Tosteson ANA, Yang Q, Nelson HD, Longton G, Soneji SS, Pepe M, et al. Second opinion strategies in breast pathology: a decision analysis addressing over-treatment, under-treatment, and care costs. Breast Cancer Research and Treatment. 2018; 167(1):195–203. [PubMed: 28879558]

[30]. Butler AC, Karpicke JD, Roediger HL. Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. J Exp Psychol Learn Mem Cogn. 2008; 34(4):918–28. [PubMed: 18605878]

[31]. Miller TM, Geraci L. Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. Metacognition and Learning. 2011; 6(3):303–14.

[32]. Callender AA, Franco-Watkins AM, Roberts AS. Improving metacognition in the classroom through instruction, training, and feedback. Metacognition and Learning. 2016; 11(2):215–35.

[33]. El Saadawi GM, Azevedo R, Castine M, Payne V, Medvedeva O, Tseytlin E, et al. Factors affecting feeling-of-knowing in a medical intelligent tutoring system: the role of immediate feedback as a metacognitive scaffold. Adv Health Sci Educ Theory Pract. 2010; 15(1):9–30. [PubMed: 19434508]

[34]. Kulik JA, Kulik C-lC. Timing of feedback and verbal learning. Review of Educational Research. 1988; 58(1):79–97.

[35]. Gonullu I, Artar M. Metacognition in medical education. Education for Health. 2014; 27(2):225. [PubMed: 25420992]

[36]. Medina MS, Castleberry AN, Persky AM. Strategies for Improving Learner Metacognition in Health Professional Education. Am J Pharm Educ. 2017; 81(4):78-. [PubMed: 28630519]

[37]. Quirk M. Intuition and metacognition in medical education: keys to developing expertise: Springer publishing company 2006.

[38]. Turan S, Demirel O, Sayek I. Metacognitive awareness and self-regulated learning skills of medical students in different medical curricula. Med Teach. 2009; 31(10):e477–e83. [PubMed: 19877856]

[39]. Fernandez-Duque D, Baird JA, Posner MI. Executive Attention and Metacognitive Regulation. Consciousness and Cognition. 2000; 9(2):288–307. [PubMed: 10924249]

[40]. Fleming SM, Dolan RJ. The neural basis of metacognitive ability. Philos Trans R Soc Lond B Biol Sci. 2012; 367(1594):1338–49. [PubMed: 22492751]

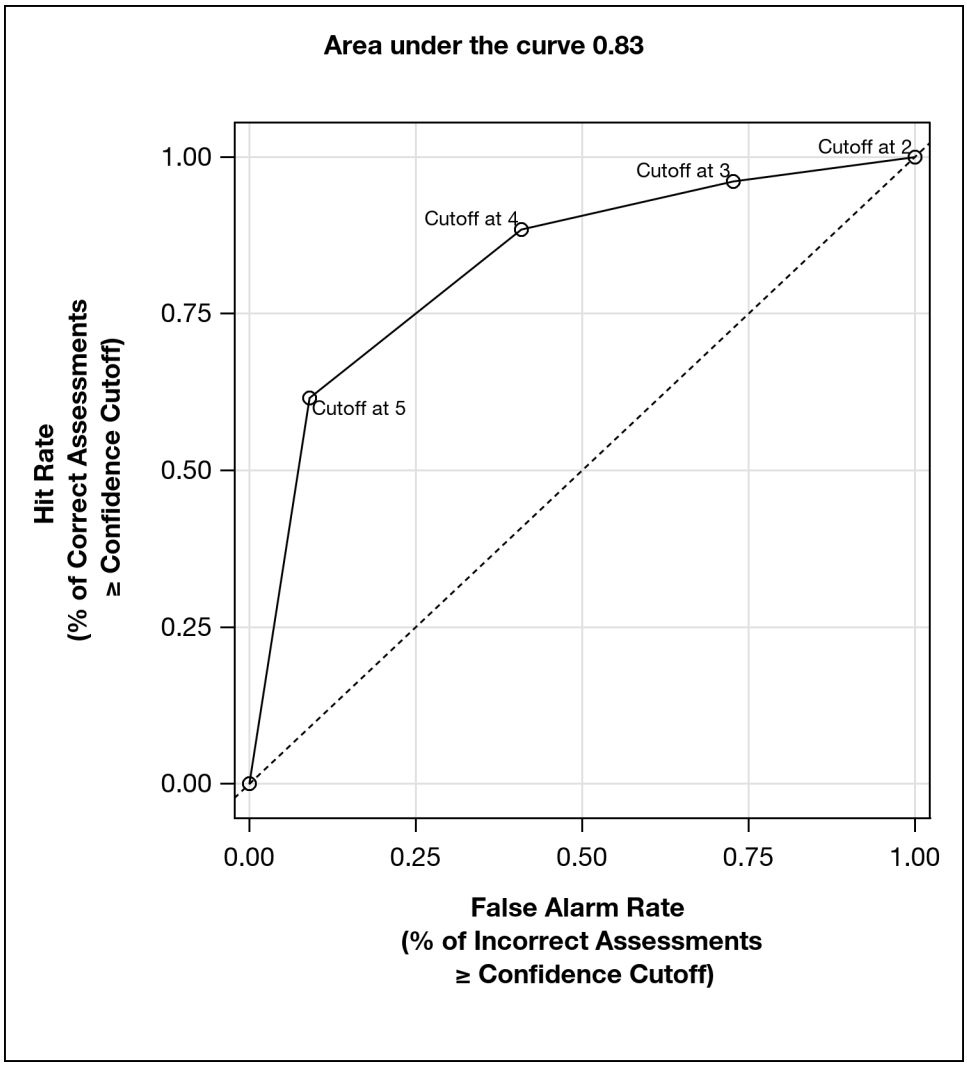
[41]. Shimamura AP. Toward a cognitive neuroscience of metacognition. Academic Press 2000.

[42]. Elder DE, Piepkorn MW, Barnhill RL, Longton GM, Nelson HD, Knezevich SR, et al. Pathologist characteristics associated with accuracy and reproducibility of melanocytic skin lesion interpretation. Journal of the American Academy of Dermatology. 2018; 79(1):52–9.e5. [PubMed: 29524584]

[43]. Kincannon JM, Gleber C, Kim J. The Effects of Metacognitive Training on Performance and Use of Metacognitive Skills in Self-Directed Learning Situations. 1999.

HIGHLIGHTS

- Metacognitive sensitivity varied across pathologists, with most showing higher sensitivity than expected by chance.
- None of the demographic or clinical characteristics we examined was significantly associated with metacognitive sensitivity.
- Pathologists with higher metacognitive sensitivity were more likely to request additional tests or second opinions for their inaccurate diagnoses.



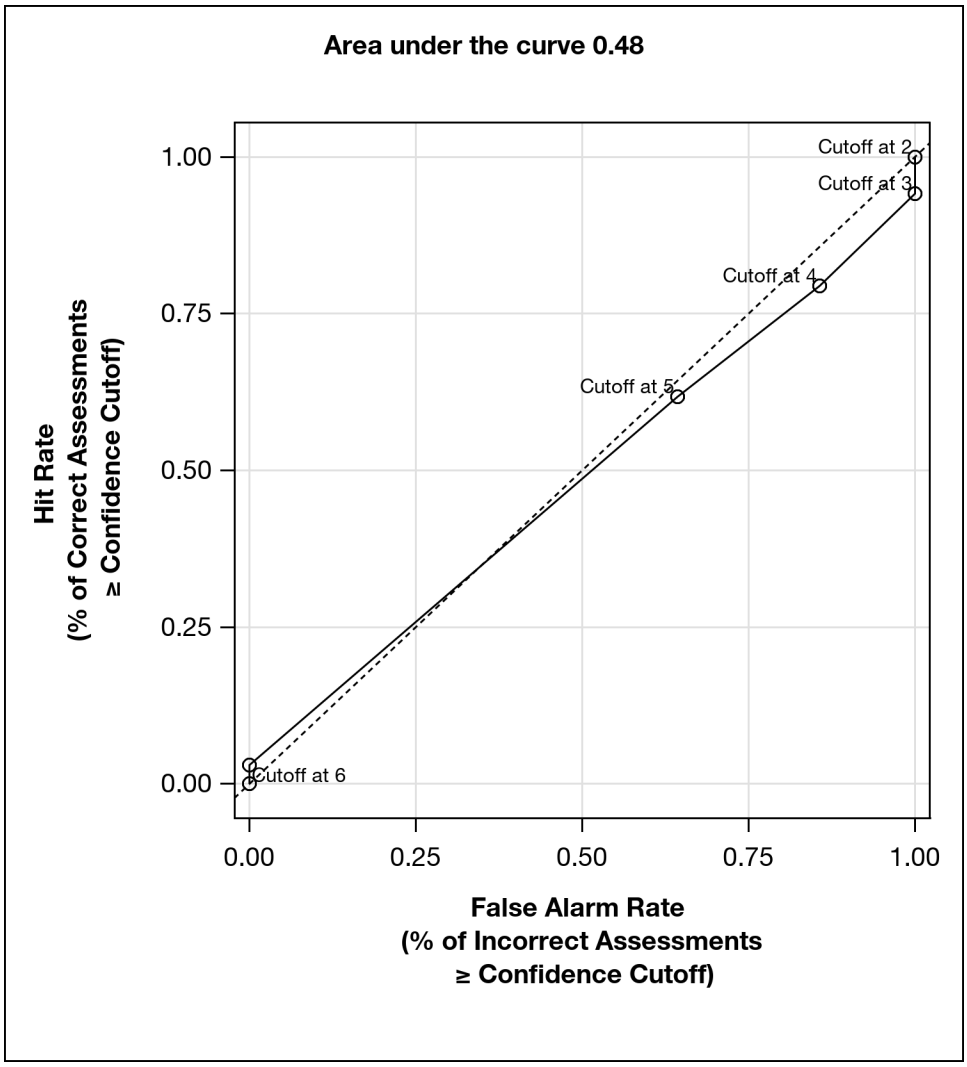


Figure 1/.

Two examples of ROC curves from M-Path participants interpreting skin biopsy slides. Each point plots the false alarm rate on the horizontal axis against the hit rate on the vertical axis for a given confidence cut-off.

Figure 1A/This example demonstrates high metacognitive sensitivity for a single subject with an AUC above the null value of 0.5.

Figure 1B/This example demonstrates a lower metacognitive sensitivity for a single subject near chance performance.

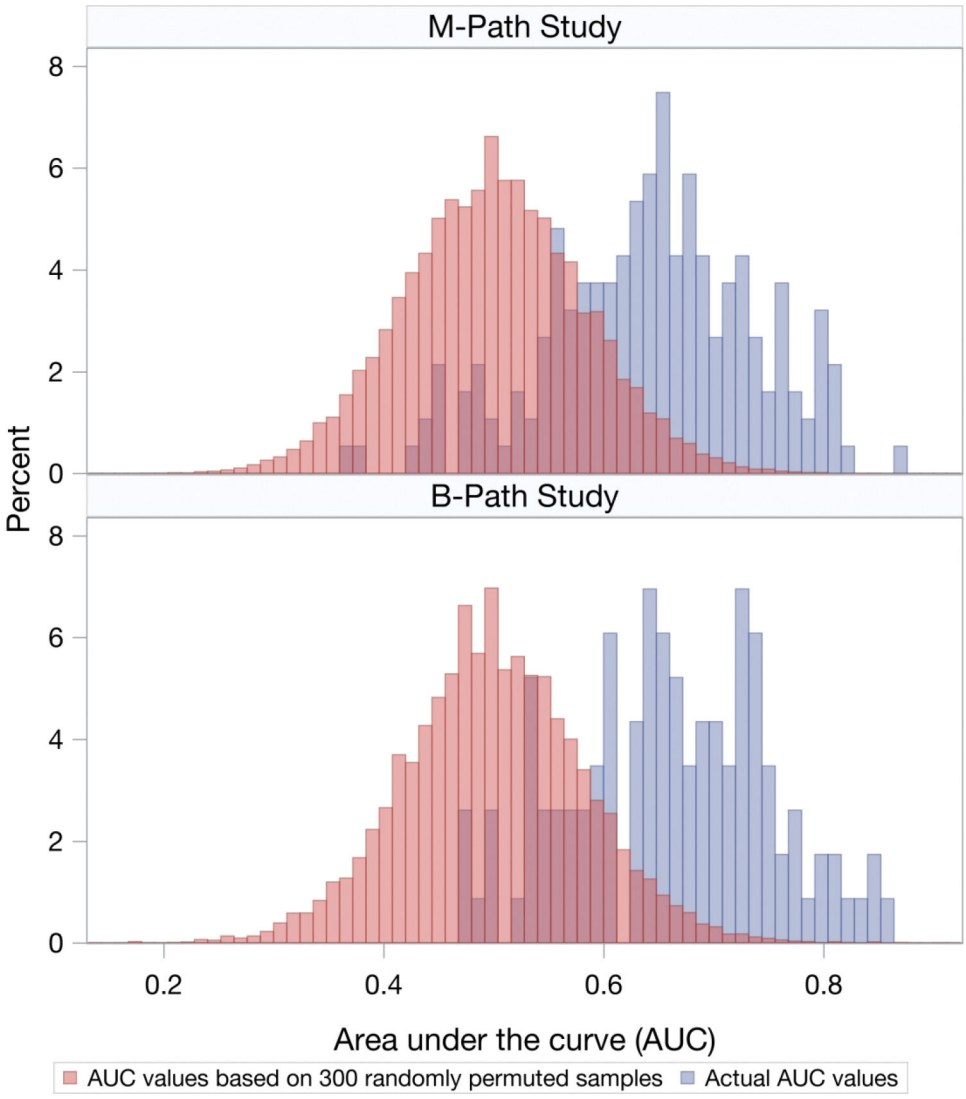
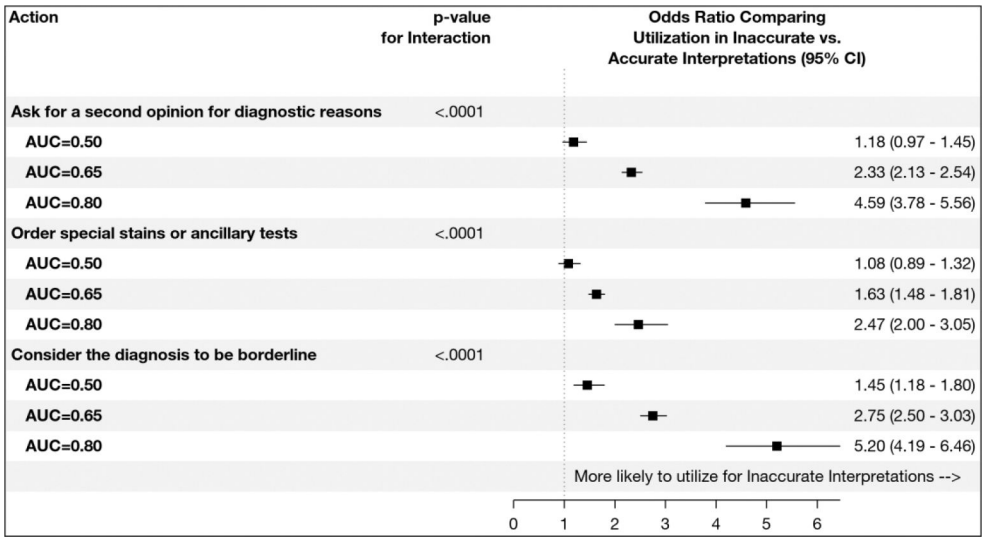


Figure 2/.

Distribution of AUC values based on 300 samples of randomly permuted confidence ratings within participants (red) and the distribution of actual pathologist metacognitive sensitivity estimates, measured by AUC, for study pathologists (blue). Top panel displays results for M-Path Study pathologists (N=187) and bottom panel displays results for B-Path Study pathologists (N=115).



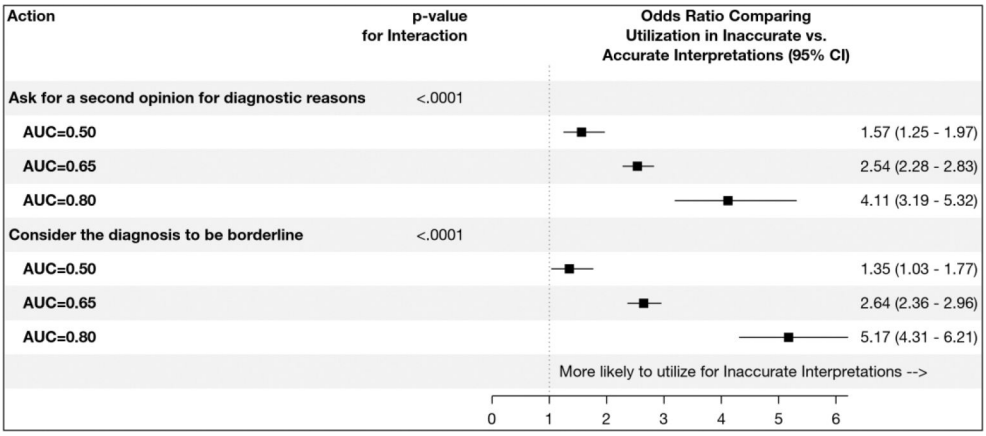


Figure 3/.

Analyses examining interaction between metacognitive sensitivity and diagnostic accuracy for secondary diagnostic actions and considering diagnoses to be borderline. Analytic models were logistic regression models with predictors AUC (continuous), accuracy (binary), and the AUC-accuracy interaction term. Although AUC was analyzed as a continuous measure, results are presented for three representative AUC values. Each outcome was more strongly associated with accuracy for highly metacognitive participants (AUC=0.8) than participants with lower (AUC=0.65) or null (AUC=0.5) metacognitive sensitivity.

Figure 3a/Interpreting skin biopsies in the M-Path Study (N = 8,976 interpretations)

Figure 3b/Interpreting breast biopsies in the B-Path Study (N = 6,900 interpretations)

Table 1

Pathologist Characteristics

Pathologist Characteristics	Skin Biopsies M-Path Study N (%)	Breast Biopsies B-Path Study N (%)
All participants	187	115
Demographics		
Age		
<40 years	32 (17%)	16 (14%)
40–49 years	55 (29%)	41 (36%)
50–59 years	63 (34%)	42 (37%)
≥60 years	37 (20%)	16 (14%)
Gender		
Male	114 (61%)	69 (60%)
Female	73 (39%)	46 (40%)
Training and Experience		
Affiliation with academic medical center		
No	134 (72%)	87 (76%)
Yes, adjunct/affiliated clinical faculty	34 (18%)	17 (15%)
Yes, primary appointment	19 (10%)	11 (10%)
Fellowship Training (M-Path)		
Dermatopathology	72 (39%)	-
Surgical Pathology	69 (37%)	-
Other	54 (29%)	-
No Fellowship	29 (16%)	-
Fellowship Training (B-Path)		
Breast pathology	-	6 (5%)
Surgical pathology	-	57 (50%)
No fellowship training in surgical or breast pathology	-	56 (49%)
Expertise in dermatopathology/breast pathology *		
Non-expert	113 (60%)	88 (77%)
Expert	74 (40%)	27 (23%)
Years interpreting melanocytic skin lesions/breast pathology cases		
<5 years	29 (16%)	22 (19%)
5–9 years	45 (24%)	23 (20%)
10–19 years	57 (30%)	34 (30%)
≥20 years	56 (30%)	36 (31%)
Percentage of caseload interpreting melanocytic skin lesions/breast pathology cases		
<10%	79 (42%)	59 (51%)
10–24%	72 (39%)	45 (39%)
≥25%	36 (19%)	11 (10%)
Number of melanocytic lesion interpretations in an average month		

Pathologist Characteristics	Skin Biopsies M-Path Study N (%)	Breast Biopsies B-Path Study N (%)
<25 cases per month	48 (26%)	-
25–99 cases per month	65 (35%)	-
100–249 cases per month	44 (24%)	-
≥250 cases per month	30 (16%)	-
Number of breast pathology cases interpreted in an average week		
<5 cases per week	-	31 (27%)
5–9 cases per week	-	44 (38%)
10–19 cases per week	-	31 (27%)
≥20 cases per week	-	9 (8%)
Attitudes towards interpreting melanocytic skin lesions/breast cases on baseline survey		
How confident are you in your assessments of melanocytic skin lesions (M-Path study) or breast lesions (B-Path study)?		
1 (Not at all Confident)	0 (0%)	0 (0%)
2	11 (6%)	0 (0%)
3	15 (8%)	8 (7%)
4	38 (20%)	27 (23%)
5	90 (48%)	66 (57%)
6 (Very Confident)	33 (18%)	14 (12%)

* For M-Path participants, expertise was defined as having board certification or fellowship training in dermatopathology. For B-Path participants, expertise was defined as having fellowship training in breast pathology or self-reported perception that their peers considered them to be experts in breast pathology.

Table 2

Summary statistics and associations with pathologist characteristics for metacognitive sensitivity (AUC). **P-values were derived from linear regression models using AUC as a continuous outcome and controlling for slide set.**

	Skin Biopsy M-Path Study (N = 187)		Breast Biopsy B-Path Study (N = 115)	
Pathologist Characteristics	**Mean of Participant AUC (SD)**	**p-value**	**Mean of Participant AUC (SD)**	**p-value**
Overall				
Total N	187		115	
All Participants	0.64 (0.10)		0.66 (0.09)	
Demographics				
Gender		0.69		0.135
Male	0.64 (0.09)		0.65 (0.09)	
Female	0.64 (0.11)		0.68 (0.09)	
Training and Experience				
Training		0.138		0.99
Expert	0.65 (0.09)		0.66 (0.10)	
Non-expert	0.63 (0.10)		0.66 (0.09)	
Years interpreting melanocytic skin lesions / breast pathology cases		0.091†		0.42†
<5 years	0.67 (0.09)		0.66 (0.10)	
5–9 years	0.64 (0.09)		0.68 (0.08)	
10–19 years	0.64 (0.10)		0.67 (0.09)	
≥20 years	0.63 (0.10)		0.64 (0.09)	
Percentage of caseload interpreting melanocytic skin lesions / breast pathology cases		0.93†		0.92†
<10%	0.64 (0.10)		0.66 (0.08)	
10–24%	0.65 (0.09)		0.65 (0.09)	
≥25%	0.63 (0.09)		0.67 (0.12)	
Number of melanocytic lesion interpretations in an average month		0.25†		
<25 cases per month	0.63 (0.11)			
25–99 cases per month	0.65 (0.08)			
100–249 cases per month	0.63 (0.10)			
≥250 cases per month	0.66 (0.09)			
Number of breast pathology cases interpreted in an average week				0.77†
<5 cases per week			0.67 (0.10)	
5–9 cases per week			0.65 (0.08)	
10–19 cases per week			0.67 (0.09)	
≥20 cases per week			0.64 (0.13)	

	Skin Biopsy M-Path Study (N = 187)		Breast Biopsy B-Path Study (N = 115)	
Pathologist Characteristics	**Mean of Participant AUC (SD)**	**p-value**	**Mean of Participant AUC (SD)**	**p-value**
Attitudes towards interpreting lesions				
General confidence reported on baseline survey before the study: How confident are you in your assessments of melanocytic skin lesions / breast pathology cases?		0.50†		0.50†
2	0.62 (0.07)		-	
3	0.63 (0.13)		0.68 (0.10)	
4	0.65 (0.11)		0.66 (0.10)	
5	0.64 (0.09)		0.66 (0.08)	
6 (Very Confident)	0.64 (0.09)		0.64 (0.12)	

†Ordinal variables were tested for significance of trend.