Original Research Article

MDM Medical Decision Making

Medical Decision Making
1–11

DOI: 10.1177/0272989X221126528
journals.sagepub.com/home/mdm
SAGE

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

Dayna A. Clayton, Megan M. Eguchi, Kathleen F. Kerr, Kiyofumi Miyoshi, Tad T. Brunyé, Trafton Drew, Donald L. Weaver, and Joann G. Elmore

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 area under the curve (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 with 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. In addition, while there is no gold standard for the correct diagnosis to determine the accuracy of pathologists' interpretations, our studies achieved a high-quality reference diagnosis by using the consensus diagnosis of 3 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.

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.

Corresponding Author
Joann G. Elmore, Department of General Internal Medicine, David Geffen School of Medicine, University of California, Los Angeles, 1100 Glendon Avenue, Suite 900, Los Angeles, CA, 90024, USA; (jelmore@-mednet.ucla.edu).

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

Date received: February 10, 2022; accepted: August 19, 2022

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.[4] outlined 4 ways in which metacognition contributes to problem solving: identifying the problem, mentally representing the problem, planning how to proceed, and evaluating one's performance. 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 2 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 use 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 1 slide set of 48 cases (M-Path) or 60 cases (B-Path). Slides sets were mailed to

Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, USA (DAC, MME, JGE); Department of Biostatistics, University of Washington, Seattle, WA, USA (KFK); Department of Psychology, University of California, Los Angeles, CA, USA (KM); Center for Applied Brain and Cognitive Sciences, Tufts University, Medford, MA, USA (TTB); Department of Psychology, University of Utah, Salt Lake City, UT, USA (TD); Department of Pathology & Laboratory Medicine, University of Vermont Larner College of Medicine, Burlington, VT, USA (DLW). The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support for this study was provided in part by grants from the National Institute of Health 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.

participants sequentially, and participants returned slide sets after they interpreted all cases using their own microscopes. Participants based their diagnosis off 1 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 1 of 4 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, 3 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 5 slide sets of 48 cases, balanced by MPATH-Dx class.

The 187 M-Path study participants were from 10 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 2 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, 3 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, atypical hyperplasia, and benign cases without atypia were included. The 240 cases were randomly assigned to 4 sets of 60 cases stratified by the consensus reference diagnosis, difficulty rating, breast density, and patient age.

The 115 B-Path participants were from 8 geographically diverse US States. Eligibility criteria included completion of residency, experience interpreting breast lesions for at least 1 y prior, and planning to interpret breast lesions for the following 2 y. 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 6-point scale ranging 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 cutoff 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. An 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 2 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 into whether AUCs differed from the distribution of AUCs with no metacognitive sensitivity, we developed a reference distribution of AUCs by creating 2000 permuted data sets. In each permuted data set, we randomly permuted each pathologist's confidence ratings across all his or 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, percentage 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 slide 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 2-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, USA) was used for analysis, and statistical significance was evaluated at a threshold of $P < 0.05$.

Results

The characteristics of the 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 y or longer (60% in M-Path and 61% in B-Path).

Example ROC curves are plotted in Figure 1 for 2 M-Path study participants, showing one pathologist with high metacognitive sensitivity (AUC 0.83) 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, 0.37–0.87). In the B-Path study, the mean AUC across all 115 pathologists was 0.66 (range, 0.47–0.85). 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. The mean AUC was similar across subgroups of pathologists (Table 2). For example, the 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 (ORs) comparing the utilization of each action for inaccurate versus accurate interpretations are shown at 3 selected values of AUC, 0.5, 0.65, and 0.8, representing null, average, and high AUC values in our study, respectively. In the M-Path study, the 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 2 different diagnoses.

Table 1 Pathologist Characteristics

Pathologist Characteristics	Skin Biopsies: M-Path Study, *n* (%)	Breast Biopsies: B-Path Study, *n* (%)
All participants	187	115
Demographics		
Age, y		
<40	32 (17)	16 (14)
40–49	55 (29)	41 (36)
50–59	63 (34)	42 (37)
≥60	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[a]		
Nonexpert	113 (60)	88 (77)
Expert	74 (40)	27 (23)
Years interpreting melanocytic skin lesions/breast pathology cases		
<5	29 (16)	22 (19)
5–9	45 (24)	23 (20)
10–19	57 (30)	34 (30)
≥20	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		
<25	48 (26)	—
25–99	65 (35)	—
100–249	44 (24)	—
≥250	30 (16)	—
Number of breast pathology cases interpreted in an average week		
<5	—	31 (27)
5–9	—	44 (38)
10–19	—	31 (27)
≥20	—	9 (8)
Attitudes toward 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)

[a]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.

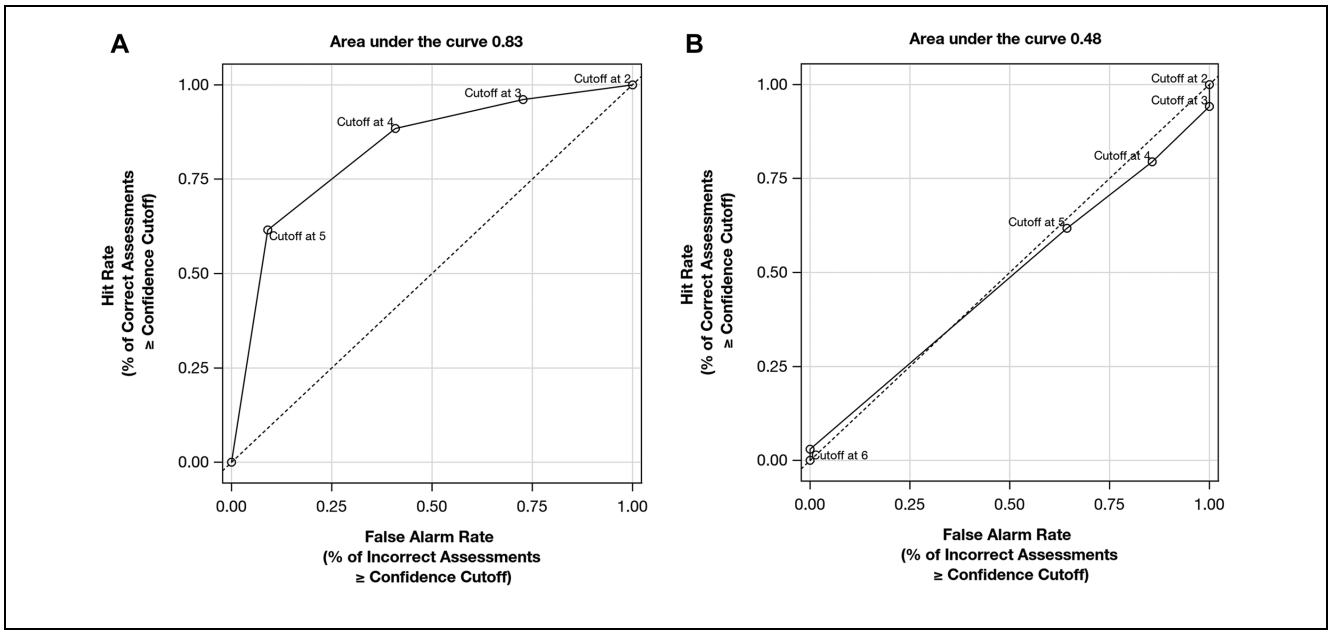

Figure 1 Two examples of receiver operating characteristic (ROC) curves from M-Path participants interpreting skin biopsy slides. Each point plots the false alarm rate on the horizontal axis against the hit rate on the vertical axis for a given confidence cutoff. (A) This example demonstrates high metacognitive sensitivity for a single subject with an area under the curve above the null value of 0.5. (B) This example demonstrates a lower metacognitive sensitivity for a single subject near chance performance.

Participants with relatively high metacognitive sensitivity were much more likely to request a secondary diagnostic action for inaccurate diagnoses compared with accurate diagnoses. For example, according to the fitted model, M-Path participants with AUC = 0.8 had a 4.59-fold higher odds of asking for a second opinion for an inaccurate diagnosis than an accurate diagnosis (95% confidence interval [CI]: 3.78–5.56) and 2.47-fold higher odds of ordering 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 versus accurate diagnoses, with ORs of 1.18 (95% CI: 0.97–1.45) and 1.08 (95% CI: 0.89–1.32). The ORs 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 an AUC of 0.8 were estimated to have 4.11-fold higher odds of asking for a second opinion for an inaccurate diagnosis than for an accurate diagnosis (95% CI: 3.19–5.32). Participants with AUC = 0.50 were also estimated to be more likely to ask for a second opinion for an inaccurate versus accurate diagnosis, but the association between accuracy and this action was much weaker (OR = 1.57, 95% CI: 1.25–1.97). Again, the ORs 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 2 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 2 independent studies in pathology, 1 in skin pathology and 1 in breast pathology. These tissues have very different histologic features, yet findings are strikingly similar in the 2 studies. We found that metacognitive sensitivity differs across pathologists and was not demonstrably better than chance for some

Table 2 Summary Statistics and Associations with Pathologist Characteristics for Metacognitive Sensitivity (AUC)[a]

Pathologist Characteristics	Skin Biopsy M-Path Study ($N = 187$)		Breast Biopsy B-Path Study ($N = 115$)	
	Mean of Participant AUC (s)	P Value	Mean of Participant AUC (s)	P Value
Overall				
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)	—
Nonexpert	0.63 (0.10)	—	0.66 (0.09)	—
Years interpreting melanocytic skin lesions/breast pathology cases	—	0.091[b]	—	0.42[b]
<5	0.67 (0.09)	—	0.66 (0.10)	—
5–9	0.64 (0.09)	—	0.68 (0.08)	—
10–19	0.64 (0.10)	—	0.67 (0.09)	—
≥20	0.63 (0.10)	—	0.64 (0.09)	—
Percentage of caseload interpreting melanocytic skin lesions/breast pathology cases	—	0.93[b]	—	0.92[b]
<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[b]	—	—
<25	0.63 (0.11)	—	—	—
25–99	0.65 (0.08)	—	—	—
100–249	0.63 (0.10)	—	—	—
≥250	0.66 (0.09)	—	—	—
Number of breast pathology cases interpreted in an average week	—	—	—	0.77[b]
<5	—	—	0.67 (0.10)	—
5–9	—	—	0.65 (0.08)	—
10–19	—	—	0.67 (0.09)	—
≥20	—	—	0.64 (0.13)	—
Attitudes toward 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[b]	—	0.50[b]
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)	—

[a] P values were derived from linear regression models using the area under the curve as a continuous outcome and controlling for slide set.
[b] Ordinal variables were tested for significance of trend.

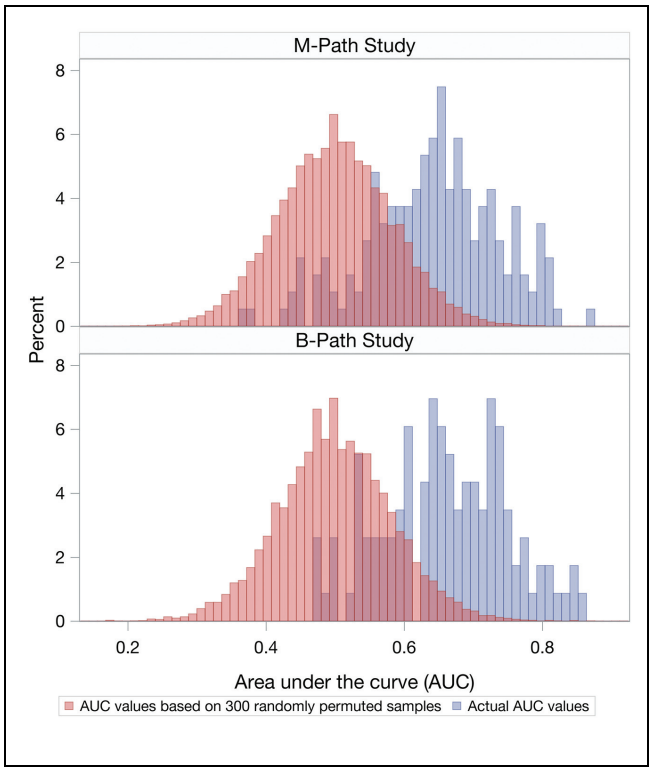

Figure 2 Distribution of the area under the curve (AUC) values based on 300 samples of randomly permuted confidence ratings within participants (red) and the distribution of actual pathologist metacognitive sensitivity estimates, measured by AUC, for study pathologists (blue). The top panel displays the results for M-Path study pathologists ($N = 187$), and bottom panel displays the results for the B-Path study pathologists ($N = 115$).

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 in which they were inaccurate and not for cases in which 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 data set (Supplementary Figure S1). 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.[6] found that when clinicians reported low confidence, they were likely to be incorrect. 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 judgments 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 with 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 2 alternatives on a continuous scale. In contrast, this study used 4 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.

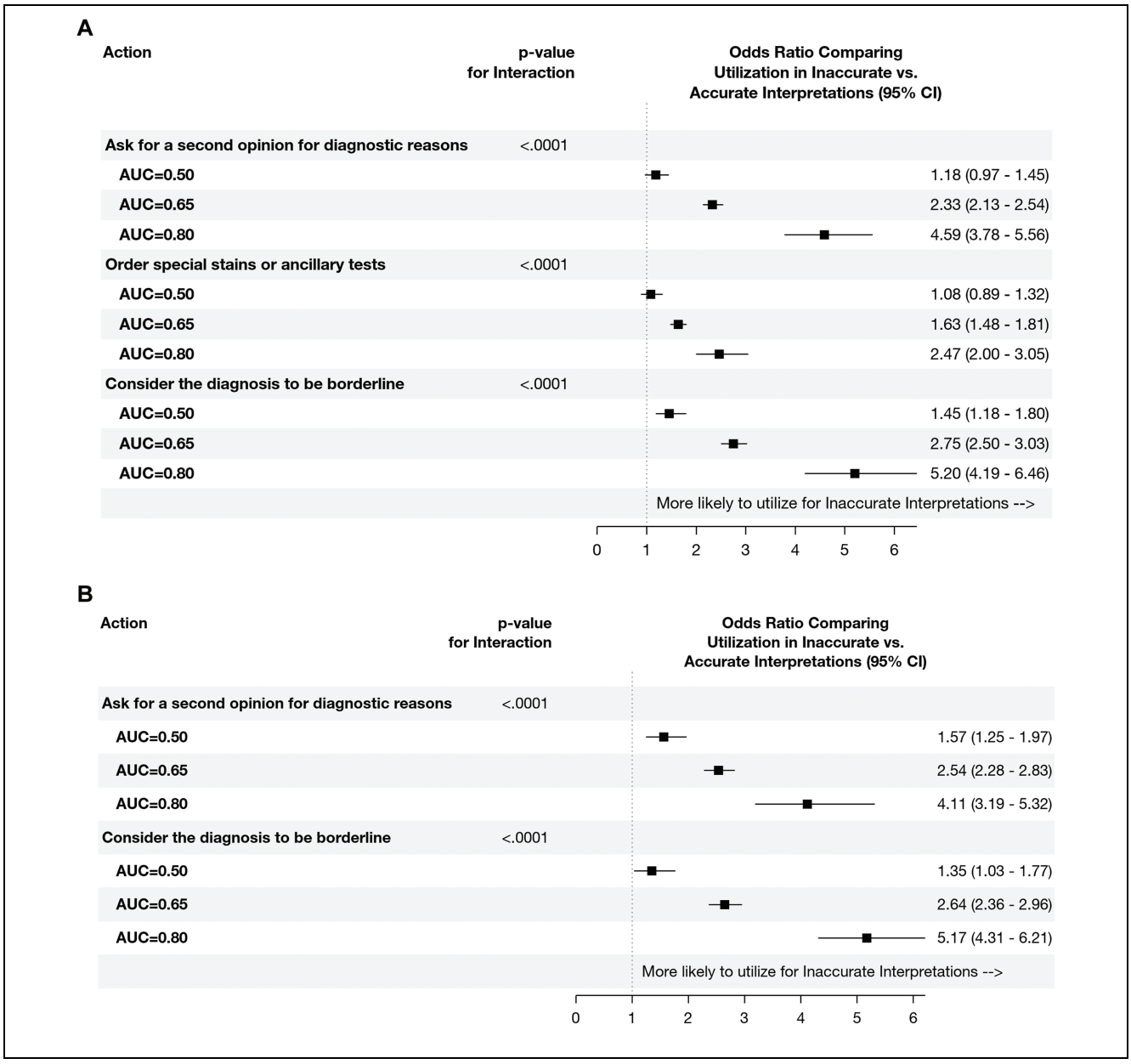

Figure 3 Analyses examining the interaction between metacognitive sensitivity and diagnostic accuracy for secondary diagnostic actions and considering diagnoses to be borderline. Analytic models were logistic regression models with predictors AUC (continuous), accuracy (binary), and the AUC-accuracy interaction term. Although AUC was analyzed as a continuous measure, results are presented for 3 representative AUC values. Each outcome was more strongly associated with accuracy for highly metacognitive participants (AUC = 0.8) than for participants with lower (AUC = 0.65) or null (AUC = 0.5) metacognitive sensitivity. (a) Interpreting skin biopsies in the M-Path study (N = 8976 interpretations). (b) Interpreting breast biopsies in the B-Path study (N = 6900 interpretations).

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 2 studies each used a panel of 3 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 continuing medical education), since feedback on diagnostic accuracy in clinical practice in pathology is often completely absent or substantially delayed.

ORCID iDs

Kathleen F. Kerr https://orcid.org/0000-0002-6438-9583
Tad T. Brunyé https://orcid.org/0000-0002-8788-8764
Joann G. Elmore https://orcid.org/0000-0002-7311-6835

Supplemental Material

Supplementary material for this article is available on the *Medical Decision Making* website at http://journals.sagepub.com/home/mdm.

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.
2. Nelson TO, Stuart RB, Howard C, Crowley M. Metacognition and clinical psychology: a preliminary framework for research and practice. *Clin Psychol Psychother*. 1999;6(2):73–9.
3. Foucar E. Diagnostic decision-making in anatomic pathology. *Pathol Patterns Rev*. 2001;116(suppl 1):S21–33.
4. Davidson JE, Deuser R, Sternberg RJ.The role of metacognition in problem solving. In: Metcalfe J, Shimamura AP, eds. *Metacognition: Knowing about Knowing*. Cambridge (MA): MIT Press; 1994. p 207–26.
5. Fleming SM, Lau HC. How to measure metacognition. *Front Hum Neurosci*. 2014;8:443.
6. Friedman CP, Gatti GG, Franz TM, et al. Do physicians know when their diagnoses are correct? *J Gen Intern Med*. 2005;20(4):334–9.
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.
8. Nelson TO. Metamemory: a theoretical framework and new findings. In: Bower GH, ed. *Psychology of Learning and Motivation*. Cambridge (MA): Academic Press; 1990: 125–73.
9. Metcalfe J, Shimamura AP. *Metacognition: Knowing about Knowing*. Cambridge (MA): MIT Press; 1994.
10. Maniscalco B, Lau H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Conscious Cogn*. 2012;21(1):422–30.
11. Sherbino J, Dore KL, Wood TJ, et al. The relationship between response time and diagnostic accuracy. *Acad Med*. 2012;87(6):785–91.
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.
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.
14. Desender K, Boldt A, Yeung N. Subjective confidence predicts information seeking in decision-making. *Psychol Sci*. 2018;29(5):761–78.
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.
16. Elmore JG, Barnhill RL, Elder DE, et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. *BMJ*. 2017;357:j2813.
17. Elmore JG, Longton GM, Carney PA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA*. 2015;313(11):1122–32.
18. 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): 131–41.
19. Allison KH, Reisch LM, Carney PA, et al. Understanding diagnostic variability in breast pathology: lessons learned from an expert consensus review panel. *Histopathology*. 2014;65(2):240–51.
20. Dalkey N, Brown B, Cochran S. *The Delphi Method, III: Use of Self-Ratings to Improve Group Estimates*. Santa Monica (CA): RAND Corporation; 1969.

21. 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):830–7.
22. Blanch DC, Hall JA, Roter DL, Frankel RM. Medical student gender and issues of confidence. *Patient Educ Couns*. 2008;72(3):374–81.
23. Blanch-Hartigan D. Medical students' self-assessment of performance: results from three meta-analyses. *Patient Educ Couns*. 2011;84(1):3–9.
24. Metcalfe J, Eich TS, Castel AD. Metacognition of agency across the lifespan. *Cognition*. 2010;116(2):267–82.
25. Cutting MF, Saks NS. Twelve tips for utilizing principles of learning to support medical education. *Med Teach*. 2012;34(1):20–4.
26. Elmore JG, Tosteson ANA, Pepe MS, et al. Evaluation of 12 strategies for obtaining second opinions to improve interpretation of breast histopathology: simulation study. *BMJ*. 2016;353:i3069.
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. *J Oncol Pract*. 2014;10(4):275–80.
28. Piepkorn MW, Longton GM, Reisch LM, et al. Assessment of second-opinion strategies for diagnoses of cutaneous melanocytic lesions. *JAMA Netw Open*. 2019;2(10): e1912597-e.
29. Tosteson ANA, Yang Q, Nelson HD, et al. Second opinion strategies in breast pathology: a decision analysis addressing over-treatment, under-treatment, and care costs. *Breast Cancer Res Treat*. 2018;167(1):195–203.
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.
31. Miller TM, Geraci L. Training metacognition in the classroom: the influence of incentives and feedback on exam predictions. *Metacogn Learn*. 2011;6(3):303–14.
32. Callender AA, Franco-Watkins AM, Roberts AS. Improving metacognition in the classroom through instruction, training, and feedback. *Metacogn Learn*. 2016;11(2): 215–35.
33. El Saadawi GM, Azevedo R, Castine M, 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.
34. Kulik JA, Kulik C-lC. Timing of feedback and verbal learning. *Rev Educ Res*. 1988;58(1):79–97.
35. Gonullu I, Artar M. Metacognition in medical education. *Educ Health*. 2014;27(2):225.
36. Medina MS, Castleberry AN, Persky AM. Strategies for improving learner metacognition in health professional education. *Am J Pharm Educ*. 2017;81(4):78.
37. Quirk M. *Intuition and Metacognition in Medical Education: Keys to Developing Expertise*. New York: Springer; 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):e477–83.
39. Fernandez-Duque D, Baird JA, Posner MI. Executive attention and metacognitive regulation. *Conscious Cogn*. 2000;9(2):288–307.
40. Fleming SM, Dolan RJ. The neural basis of metacognitive ability. *Philos Trans R Soc Lond B Biol Sci*. 2012; 367(1594):1338–49.
41. Shimamura AP. *Toward a Cognitive Neuroscience of Metacognition*. Cambridge (MA): Academic Press; 2000.
42. Elder DE, Piepkorn MW, Barnhill RL, et al. Pathologist characteristics associated with accuracy and reproducibility of melanocytic skin lesion interpretation. *J Am Acad Dermatol*. 2018;79(1):52-9.e5.
43. Kincannon JM, Gleber C, Kim J. *The Effects of Metacognitive Training on Performance and Use of Metacognitive Skills in Self-Directed Learning Situations*. Eric No. ED436146. February 1999.