Title

Pathologist characteristics associated with accuracy and reproducibility of melanocytic skin lesion interpretation.

Journal

Journal of The American Academy of Dermatology, 79(1)

Authors

Elder, David
Piepkorn, Michael
Barnhill, Raymond
et al.

Publication Date

2018-07-01

DOI

10.1016/j.jaad.2018.02.070

Peer reviewed

Published in final edited form as:

J Am Acad Dermatol. 2018 July ; 79(1): 52–59.e5. doi:10.1016/j.jaad.2018.02.070.

Pathologist Characteristics Associated with Accuracy and Reproducibility of Melanocytic Skin Lesion Interpretation

David E. Elder, **MB ChB, FRCPA**[1], **Mike Piepkorn, MD, PhD**[2,3], **Raymond L. Barnhill, MD, MSc**[4,5], **Gary M. Longton, MS**[6], **Heidi D. Nelson, MD, MPH**[7], **Stevan R. Knezevich, MD, PhD**[8], **Margaret S. Pepe, PhD**[6], **Patricia A. Carney, PhD**[9], **Linda J. Titus, PhD, MA**[10], **Tracy Onega, PhD**[11,12], **Anna N. A. Tosteson, ScD**[13], **Martin A. Weinstock, MD, PhD**[14,15], and **Joann G. Elmore, MD, MPH**[16]

[1]Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania, Philadelphia, PA, USA

[2]Division of Dermatology, Department of Medicine, University of Washington School of Medicine, Seattle, WA, USA

[3]Dermatopathology Northwest, Bellevue, WA, USA

[4]Department of Pathology, Institut Curie, Paris, France

[5]University of Paris Descartes, Paris, France

[6]Program in Biostatistics and Biomathematics, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[7]Departments of Medical Informatics and Clinical Epidemiology and Medicine, Oregon Health & Science University

[8]Pathology Associates, Clovis, CA, USA

[9]Department of Family Medicine, Oregon Health and Science University, Portland, OR, USA

[10]Departments of Epidemiology and Pediatrics, Geisel School of Medicine at Dartmouth, Norris Cotton Cancer Center, Lebanon, NH, USA

[11]Department of Biomedical Data Science, Department of Epidemiology, Norris Cotton Cancer Center

[12]Geisel School of Medicine at Dartmouth, The Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, NH, USA

Correspondence to: Joann G. Elmore, MD, MPH, University of Washington, Mailbox 359780, 325 Ninth Avenue, Seattle, WA 98104, Telephone: (206) 744-3632, Fax: (206) 744-6097, jeadmin@uw.edu.

IRB statement: This study was approved by the Institutional Review Boards of the University of Washington, Fred Hutchinson Cancer Research Center, Oregon Health & Science University, Rhode Island Hospital, and Dartmouth College.

[13]Departments of Medicine and Community and Family Medicine, The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Norris Cotton Cancer Center, Lebanon, NH, USA

[14]Center for Dermatoepidemiology, VA Medical Center, Providence Department of Dermatology, Rhode Island Hospital

[15]Departments of Dermatology and Epidemiology, Brown University, Providence, RI, USA

[16]Department of Medicine, University of Washington School of Medicine, Seattle, WA, USA

Abstract

Background—Diagnostic interpretations of melanocytic skin lesions vary widely among pathologists, yet the underlying reasons remain unclear.

Objective—Identify pathologist characteristics associated with rates of accuracy and reproducibility.

Methods—Pathologists independently interpreted the same set of biopsies of melanocytic lesions on two occasions. Diagnoses were categorized into one of five classes according to the MPATH-Dx© system. Reproducibility was determined by pathologists' concordance of diagnoses across two occasions. Accuracy was defined by concordance with a consensus reference standard. Associations of pathologist characteristics with reproducibility and accuracy were assessed individually and in multivariable logistic regression models.

Results—Rates of diagnostic reproducibility and accuracy were highest among pathologists with board certification and/or fellowship training in dermatopathology, and those with ≥5 years of experience. In addition, accuracy was high among pathologists with higher caseload composition and volume of melanocytic lesions.

Limitations—Data gathered in a test set situation using a classification tool not currently in clinical use.

Conclusion—Diagnoses are more accurate among pathologists with specialty training and those with more experience interpreting melanocytic lesions. These findings support the practice of referring difficult cases to more experienced pathologists to improve diagnostic accuracy, although the impact on patient outcomes of these referrals requires additional research.

Keywords

dermatopathology; pathologist characteristics; diagnosis; discordance; observer variability; melanoma; melanocytic lesions

INTRODUCTION

Background

The accuracy and reproducibility of diagnostic interpretations of melanocytic skin lesions vary widely among pathologists in some portions of the histological disease spectrum, yet the underlying reasons for this remain unclear.[1] Objective: This study aims to identify pathologist characteristics associated with this variability. Previous studies have identified

considerable diagnostic variation; however, these have in general involved small series of selected cases and for the most part have been conducted among specialists in academic rather than in community settings.[2–7] These studies have not addressed differences among pathologists with varying degrees of training and experience, and few have categorized lesions according to characteristics that might be associated with diagnostic difficulty. We hypothesized that greater experience and higher levels of formal training would be associated with increased diagnostic accuracy and reproducibility.

METHODS

Human Subjects

Detailed information about the study design and data collection is provided elsewhere.[8–12] This study was approved by the Institutional Review Boards of the University of Washington, Fred Hutchinson Cancer Research Center, Oregon Health & Science University, Rhode Island Hospital, and Dartmouth College. Participating pathologists provided informed consent.

Participants

Pathologists who had completed their residency and/or fellowship training, who had interpreted cutaneous melanocytic lesions within the previous year in their own clinical practices, and were expected to interpret such lesions for the following two years were eligible to participate.

Skin Biopsy Cases

The 240 test cases were selected from routine cases that included shave, punch, and excisional skin biopsies of melanocytic lesions. A new hematoxylin and eosin stained glass slide was prepared for each case. Three experienced dermatopathologists independently reviewed each case, followed by an in-person consensus review using a modified Delphi approach,[13] to achieve a consensus reference diagnosis. Each case was assigned to one of the five classes using the Melanocytic Pathology Assessment Tool and Hierarchy for Diagnosis (MPATH-Dx©) classification tool, which incorporates treatment recommendations.[11] It was assumed that the specimen margin was positive for the purpose of management recommendations. Examples of potential diagnostic terms for each MPATH-Dx© Class, along with associated treatment recommendations, are: I) nevus/mild atypia - no further treatment required; II) moderate atypia/dysplasia -consider narrow but complete excision margin <5mm; III) severe dysplasia/melanoma *in situ* - excision with 5 mm margins; IV) pT1a invasive melanoma - wide excision ≥1cm margin; and V); ≥pT1b invasive melanoma - wide excision ≥1cm with possible additional treatment. These examples are not inclusive of the vast number of terms that can be used in diagnosis of melanocytic lesions, [1,14] and are subject to further development and revision by consensus groups. The exact wording of the MPATH-Dx© questionnaire is included in online supplementary material.

Because we hypothesized that diagnostic difficulty would be concentrated in lesions with atypia including “thin” low-risk melanomas, and because of the relative rarity of these lesions in routine practice, the distribution of cases in the sample sets intentionally included

a higher proportion of cases in MPATH-Dx© Classes II-V (the so-called “intermediate” lesions and the low-stage melanomas) than typically encountered in practice: 10.4% (n=25) Class I, 15.0% (n=36) Class II, 25.0% (n=60) Class III, 24.2% (n=58) Class IV, and 25.4% (n=61) Class V. Participants were not informed of the distribution of biopsy cases.

The 240 cases were divided into five different sets that each included the full range of MPATH-Dx© classes. All pathologists interpreted 48 cases using glass slides in Phase I of the study; in Phase II they interpreted a test set including 36 or 48 cases in either a glass or digital format. Data from the 40 pathologists assigned to the digital format reading during Phase II were used in a separate study; therefore, only data from their Phase I interpretations are included in this analysis.

Data Collection

All participants completed a baseline survey assessing their demographic and clinical practice characteristics before being randomized to a test set.[8,10] The slides were arranged in a random order for each participant. The patient’s age, sex, biopsy type, and anatomic location of the biopsy site were provided for each case. In order to limit the number of slides for the review, and preclude any need for additional sections to be provided (e.g. levels through the block), pathologists were asked to assume that the glass slide was representative of the entire melanocytic lesion. In order to allow us to request a treatment recommendation for each lesion, they were also asked to assume that the margin was involved in all cases. Pathologists used their own microscopes and provided their diagnoses using an online histology form that included the MPATH-Dx© system to classify their diagnoses into one of the five MPATH-Dx© classes.

After a wash-out period of at least 8 months, participants were invited to Phase II, in which they viewed the same cases presented in a different randomly assigned order. Participants were not told that they were sent the same cases in Phase II that they had already seen in Phase I.

Analysis

We compared each pathologist’s diagnosis to the consensus reference diagnosis for each case to estimate accuracy (rate of inter-observer concordance). Over-interpretation was defined as the participant diagnosing a case at a higher MPATH-Dx© class relative to the consensus reference diagnosis; under-interpretation was defined as the participant diagnosing a case at a lower MPATH-Dx© class. Accurate diagnoses were those in agreement with the reference diagnosis. Confidence intervals accounted for both within- and across-participant variability by employing variance estimates of the form $\{\text{var}(\text{rate}_p)+[\text{ave}(\text{rate}_p)\times(1-\text{ave}(\text{rate}_p))]/n_c\}/n_p$ where $\text{ave}(\text{rate}_p)$ is the average rate among pathologists, $\text{var}(\text{rate}_p)$ is the sample variance of rates among pathologists, n_c is the number of cases interpreted by each pathologist, and n_p is the number of pathologists.[15]

We compared each pathologist’s diagnosis for a single case in Phase I to their diagnosis on the same case in Phase II to estimate reproducibility (rate of intra-observer concordance). Reproducibility was defined as the proportion of interpretations on the same case that received the same MPATH-Dx© class diagnosis by the same pathologist in both Phases I and

II. Confidence intervals were estimated using a logit transformation and robust standard error that accounted for pathologist-level clustering.

The associations of pathologist characteristics with estimates of reproducibility and accuracy were assessed by comparing concordance rates between subgroups of pathologists (e.g. pathologist level of experience or training). Logistic regression models were fit to determine the best combination of characteristics predicting accuracy and reproducibility and to assess their effects with adjustment for the effects of other related characteristics. Models used robust estimators of variance to account for correlation of case interpretations from the same pathologist. Stata statistical software (StataCorp), version 14, was used.

RESULTS

Study Population

Of 301 pathologists eligible for the study, 187 participants completed Phase I interpretations and 118 completed glass slide interpretations during Phase II (50 pathologists were randomized to interpret the cases in whole slide digital imaging format in Phase II). Participating pathologists were predominantly male (61%), ≥50 years of age (54%), not affiliated with an academic medical center (72%), and reported ≥10 years of experience interpreting melanocytic skin lesions (60%). All pathologists were required to interpret melanocytic skin lesions in their clinical practice in order to participate in the study; 19% reported that more than a quarter of their caseloads included melanocytic lesions, and 40% were board certified and/or fellowship trained in dermatopathology. The majority of pathologists (86%) reported being moderately to extremely confident in their assessments of melanocytic lesions, however more than half (69%) of participants also reported that interpreting melanocytic skin lesions makes them more nervous than other types of pathology.

Pathologist Characteristics Associated with Accuracy

Accuracy rates of Phase I interpretations by pathologist characteristics are shown in Table 1. No pathologist characteristics were significantly associated with rates of over-interpretation. Under-interpretation rates were lower among pathologists with academic affiliations, those with a higher percentage of melanocytic skin lesions in their caseloads, a higher volume of melanocytic skin lesions within a month's caseload, reported expertise in diagnosis of melanocytic lesions, and board certification and/or fellowship training in dermatopathology.

Pathologists' characteristics that remained significantly associated with accuracy in multivariable logistic regression models (Figure 1) include: dermatopathology board certification and/or fellowship training (OR 1.41; 95% CI 1.22, 1.63); melanocytic skin lesion case load of 60 or more per month (OR 1.18; 95% CI 1.05, 1.34); 5 or more years of experience interpreting melanocytic skin lesions (OR 1.22; 95% CI 1.04, 1.44); and a composition of >10% melanocytic cases in their practice (OR 1.14: 95% CI .99, 1.30).

Pathologist Characteristics Associated with Reproducibility

Reproducibility rates, indicating agreement of Phase I and Phase II interpretations of the same case, of 118 pathologists who completed both phases of the study are shown in Table 2. Rates were significantly higher among pathologists who had completed either a dermatopathology board certification and/or fellowship program, and among those with 10% or more of their caseload consisting of melanocytic lesions and who reported interpreting 60 or more melanocytic lesions per month. Figure 2 shows the multivariable model for reproducibility, with significantly higher reproducibility among board certified and/or fellowship trained pathologists and those with more than 5 years interpreting melanocytic skin lesions. Caseload volume and composition no longer contribute to prediction of increased reproducibility when the effects of board certification/fellowship training and years of experience are accounted for.

DISCUSSION

In multivariable analysis, pathologists with board certification or fellowship training in dermatopathology and 5 or more years of experience had higher rates of reproducibility and accuracy. In addition, pathologists with higher caseload volume and interpreting more melanocytic lesions in practice demonstrated higher accuracy. While the differences noted between groups reached statistical significance, the absolute quantitative differences were perhaps modest for some comparisons. However, when considering the increasing number of skin biopsies obtained each year, the impact of our findings at a population level needs to be considered. It is estimated that 23% of all skin biopsies obtained among adults in the U.S. are of melanocytic lesions.[16]

These findings are generally similar to those from other, less comprehensive, studies.[16–20] The consensus points to dermatopathologists being the best suited to interpret challenging melanocytic skin lesions. However, the influence of experience and higher case volume, associated with better outcomes in other fields of medicine[22], was not addressed in these studies, and these attributes were found to be significant in this present study. Although the attributes are highly correlated,,our data nevertheless suggest that clinical experience is an important attribute of accuracy, supporting the continuing involvement of “legacy” practitioners with years of experience, or who have acquired diagnostic skills outside of a formal dermatopathology training program.

Our present and prior studies suggest directions for needed improvement in the field of diagnostic melanocytic pathology to better serve patients. The field could benefit substantially from efforts to simplify the diverse and confusing nomenclatures in current existence. Our MPATH-Dx© mapping tool based on observer perceptions of risks of given lesions and the appropriate surgical interventions for them is a reasonable first iteration to prompt the community of melanocytic lesion pathologists to move forward collectively to refine such a stratification schema incorporating general consensus viewpoints of the medical community. In addition, when new concepts and tools such as supplemental molecular tests to complement traditional histopathological criteria are introduced, attempts should be promptly instituted to determine the evidence basis for those putative

advancements, including determinations of their specificity, sensitivity, predictive value and reproducibility.

LIMITATIONS AND STRENGTHS OF THE STUDY

Despite this being a large study, every study has limitations. The current data was gathered in a test situation, with only 1 slide per case and no opportunity for participating pathologists to consult with a colleague for second opinion, nor the opportunity to request additional stains, deeper level sections, or tests. For this reason, generalizability to real practice where pathologists do have these opportunities is not exact. Nevertheless, it is not uncommon in practice for the results of these investigations to add little to the diagnostic specificity, and the participants were asked to assume that the single slide was representative. Additionally, although our findings are reported by classifying interpretations into the MPATH-Dx© categories, this classification tool was not in clinical use at the time of our study. Finally, while the definition of accuracy is based on the consensus diagnosis among three experienced pathologists, which would be considered ideal in a clinical setting, the natural outcome of the cases is unknown. Given that the diagnostic truth is in the biology, our consensus reference diagnosis is the best (and yet imperfect, as it is subjective) proxy available to study accuracy.

The strengths of the current study included a large number of cases randomly selected to fill the full range of diagnostic classes from benign to invasive melanoma, while weighted to include a higher proportion of the diagnostically more difficult “intermediate” lesions. We recruited a large and diverse population of pathologists to participate in this study, and gathered a multitude of relevant variables on pathologist characteristics. Participating pathologists also had a broad range of levels of clinical experience.

CONCLUSION

Diagnoses of melanocytic tumors are more accurate among pathologists with specialty training and/or more experience interpreting melanocytic lesions. These findings support the practice of referring difficult cases to more experienced pathologists to improve diagnostic accuracy, although the impact on patient outcomes of these referrals requires additional research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding sources: This work was supported by the National Cancer Institute (R01 CA151306, R01 CA201376 and K05 CA104699). The funding agency had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Conflicts of interest: All authors had financial support from the National Cancer Institute for this study. RLB had a financial relationship with Myriad Genetics outside of the submitted work, and GML had grants from Fred Hutchinson Cancer Research Center while this study was conducted.

References

1. Elmore JG, Barnhill RL, Elder DE, et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. Bmj-Brit Med J. 2017:357.
2. Barnhill RL, Argenyi ZB, From L, et al. Atypical Spitz nevi/tumors: lack of consensus for diagnosis, discrimination from melanoma, and prediction of outcome. Hum Pathol. 1999; 30(5):513–520. [PubMed: 10333219]
3. Barnhill RL, Argenyi Z, Berwick M, et al. Atypical cellular blue nevi (cellular blue nevi with atypical features): lack of consensus for diagnosis and distinction from cellular blue nevi and malignant melanoma ("malignant blue nevus"). Am J Surg Pathol. 2008; 32(1):36–44. [PubMed: 18162768]
4. Gerami P, Busam K, Cochran A, et al. Histomorphologic assessment and interobserver diagnostic reproducibility of atypical spitzoid melanocytic neoplasms with long-term follow-up. Am J Surg Pathol. 2014; 38(7):934–940. [PubMed: 24618612]
5. Duncan LM, Berwick M, Bruijn JA, Byers HR, Mihm MC, Barnhill RL. Histopathologic recognition and grading of dysplastic melanocytic nevi: an interobserver agreement study. J Invest Dermatol. 1993; 100(3):318s–321s. [PubMed: 8440913]
6. Duray PH, DerSimonian R, Barnhill R, et al. An analysis of interobserver recognition of the histopathologic features of dysplastic nevi from a mixed group of nevomelanocytic lesions. J Am Acad Dermatol. 1992; 27(5 Pt 1):741–749. [PubMed: 1430397]
7. Corona R, Mele A, Amini M, et al. Interobserver variability on the histopathologic diagnosis of cutaneous melanoma and other pigmented skin lesions. J Clin Oncol. 1996; 14(4):1218–1223. [PubMed: 8648377]
8. Onega T, Reisch L, Frederick P, et al. Use of Digital Whole Slide Imaging in Dermatopathology. J Digit Imaging. 2016; 29(2):243–253. [PubMed: 26546178]
9. Knezevich SR, Barnhill RL, Elder DE, et al. Variability in mitotic figures in serial sections of thin melanomas. J Am Acad Dermatol. 2014; 71(6):1204–1211. [PubMed: 25239732]
10. Carney PA, Frederick PD, Reisch LM, et al. How concerns and experiences with medical malpractice affect dermatopathologists' perceptions of their diagnostic practices when interpreting cutaneous melanocytic lesions. J Am Acad Dermatol. 2016; 74(2):317–324. quiz 324 e311-318. [PubMed: 26559597]
11. 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. [PubMed: 24176521]
12. 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. [PubMed: 27247109]
13. Dalkey, NC., Brown, B., Cochran, N. The Delphi Method, III Use of Self Ratings to Improve Group Estimates. Santa Monica, CA: Rand Corp; 1969.
14. Lott JP, Elmore JG, Zhao GA, et al. Evaluation of the Melanocytic Pathology Assessment Tool and Hierarchy for Diagnosis (MPATH-Dx) classification scheme for diagnosis of cutaneous melanocytic neoplasms: Results from the International Melanoma Pathology Study Group. J Am Acad Dermatol. 2016; 75(2):356–363. [PubMed: 27189823]
15. Elmore JG, Longton GM, Carney PA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA. 2015; 313(11):1122–1132. [PubMed: 25781441]
16. Lott J, Boudreau D, Barnhill R, et al. Population-Based Analysis of Histologically Confirmed Melanocytic Proliferations Using Natural Language Processing. JAMA Dermatology. 2017
17. Gaudi S, Zarandona JM, Raab SS, English JC, Jukic DM. Discrepancies in dermatopathology diagnoses: the role of second review policies and dermatopathology fellowship training. J Am Acad Dermatol. 2013; 68(1):119–128. [PubMed: 22892284]
18. Trotter SC, Sroa N, Winkelmann RR, Olencki T, Bechtel M. A Global Review of Melanoma Follow-up Guidelines. J Clin Aesthet Dermatol. 2013; 6(9):18–26.
19. Grant-Kels JM. The whys and wherefores of who reads dermatopathology slides. J Am Acad Dermatol. 2005; 53(4):703–704. [PubMed: 16198799]

20. Ackerman AB. Dermatologist not equal dermatopathologist: No place in a profession for pretenders. J Am Acad Dermatol. 2005; 53(4):698–699. [PubMed: 16198796]

21. Moy R. The reason that dermatologists should not send all their slides to dermatopathologists is a scope of practice argument, not an ethical arguement. JAAD. 2005; 53:700. [PubMed: 16198797]

22. Halm E, Lee CH, Chassin M. Is Volume Related to Outcome in Health Care? A Systematic Review and Methodologic Critique of the Literature. Ann Intern Med. 2002; 137(6):511–520. [PubMed: 12230353]

23. Gimotty PA, Elder DE, Fraker DL, et al. Identification of high-risk patients among those diagnosed with thin cutaneous melanomas. J Clin Oncol. 2007; 25(9):1129–1134. [PubMed: 17369575]

24. Balch CM, Gershenwald JE, Soong SJ, et al. Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol. 2009; 27(36):6199–6206. [PubMed: 19917835]

Capsule Summary

- Diagnostic interpretations of melanocytic skin lesions vary widely among pathologists
- Experienced, dermatopathology trained physicians' interpretations of melanocytic skin lesions are more accurate and reproducible
- Accuracy in clinical practice may be improved with referrals of difficult cases to experienced dermatopathologists.

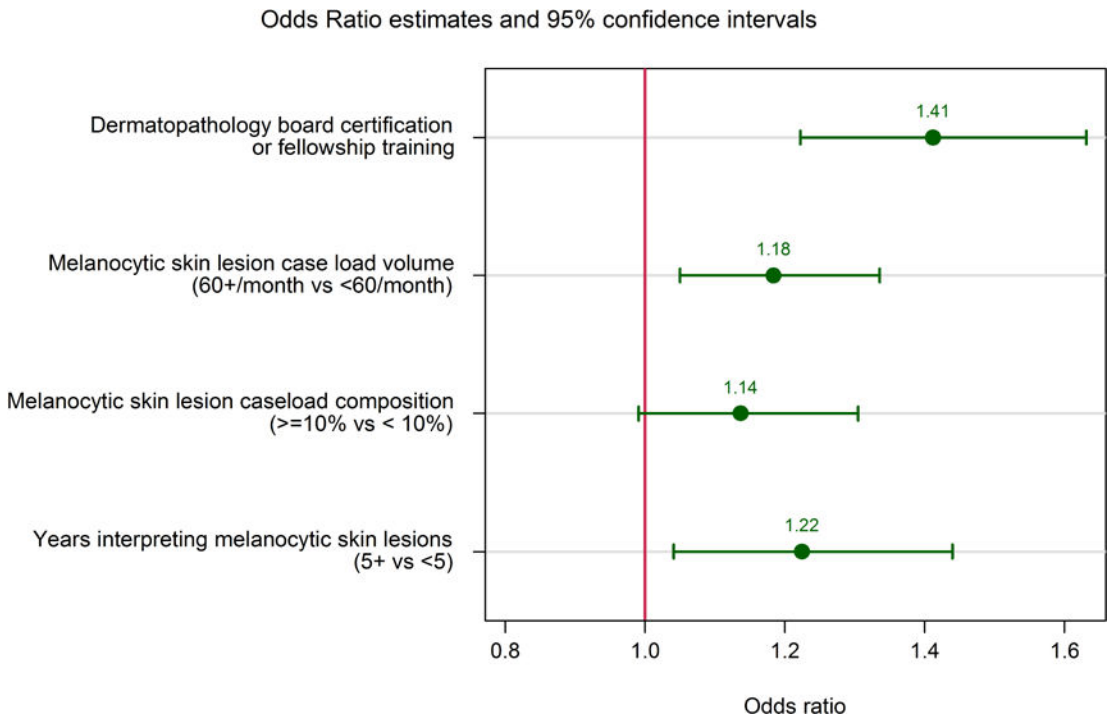
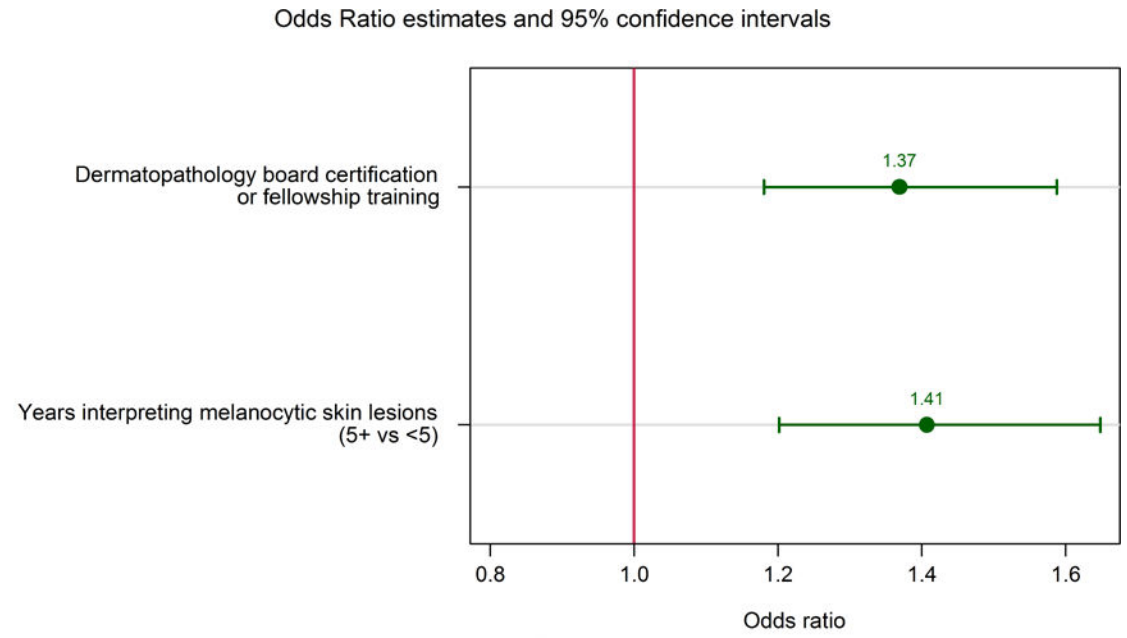


Figure 1.

Multivariable logistic regression model of **accuracy** as a function of pathologist characteristics.[1]

[1]Outcome of accuracy is defined as participant concordance with the reference diagnosis; OR>1 indicates increased concordance/accuracy. Pathologist level factors considered for inclusion in the multivariable model of accuracy were dermatopathology board certification, fellowship training, years of experience interpreting melanocytic skin lesions (MSL), affiliation with an academic medical center, practice size, melanocytic caseload composition (% MSL), melanocytic caseload volume (# MSL cases/month), and self-perceived MSL expertise among peers.



* The outcome is the agreement of the participant phase II interpretation with his/her phase I interpretation of the same case. OR>1 = improved intra-observer agreement.

Figure 2.

Multivariable logistic regression model of **reproducibility** as a function of pathologist characteristics.[2]

[2]Pathologist level factors considered for inclusion in the multivariable model of reproducibility were dermatopathology board certification, fellowship training, years of experience interpreting melanocytic skin lesions (MSL), affiliation with an academic medical center, melanocytic caseload composition (% MSL), melanocytic caseload volume (# MSL cases/month), and self-perceived MSL expertise among peers.

Table 1

Pathologist Characteristics and Rates of Accuracy (Over-interpretation, Under-interpretation, and Concordance)[a]

Pathologist Characteristics (N=187 Participants)	Total (N)	Discordance				Concordance	
		Over-Interpretation Rate (95%CI)	P-value	Under-Interpretation Rate (95%CI)	P-value	Concordance Rate (95%CI)	P-value
Academic affiliation[b]							
None	134	6% (5%, 7%)		44% (41%, 46%)		50% (48%, 52%)	
Adjunct Affiliation	34	6% (5%, 9%)		37% (32%, 42%)		57% (52%, 62%)	
Primary Academic	19	6% (4%, 9%)	.97	36% (31%, 43%)	.004	57% (52%, 63%)	.001
Estimated years interpreting melanocytic skin lesions[c]							
<5	29	5% (3%, 7%)		46% (41%, 51%)		49% (44%, 54%)	
5-9	45	6% (5%, 8%)		38% (33%, 42%)		56% (52%, 60%)	
10-19	57	6% (5%, 8%)		41% (37%, 45%)		53% (50%, 57%)	
20+	56	7% (6%, 9%)	.09	44% (40%, 47%)	.81	49% (46%, 52%)	.29
% of caseload that is melanocytic skin lesions[d]							
< 10%	79	7% (5%, 8%)		47% (44%, 50%)		46% (44%, 49%)	
10-24%	72	5% (4%, 7%)		40% (37%, 44%)		54% (52%, 57%)	
>= 25%	36	8% (6%, 10%)	.80	33% (28%, 37%)	< .001	60% (56%, 64%)	< .001
Melanocytic Skin Lesion case load (# cases/month)[c]							
< 25	48	7% (6%, 9%)		47% (43%, 51%)		46% (42%, 49%)	< .001
25-59	50	6% (4%, 8%)		46% (42%, 50%)		48% (45%, 52%)	
60-179	42	5% (4%, 7%)		38% (34%, 42%)		57% (53%, 60%)	
180+	47	6% (5%, 8%)	.32	35% (32%, 39%)	< .001	58% (55%, 61%)	
Board certification or fellowship trained in dermatopathology							
No	113	6% (5%, 7%)		47% (44%, 49%)		47% (45%, 49%)	
Yes	74	6% (5%, 8%)	.72	34% (31%, 37%)	<.001	60% (57%, 62%)	< .001

[a]Reference consensus diagnosis was obtained from the consensus of three experienced dermatopathologists.

[b] P-values for academic affiliation are for a test of no academic affiliation vs. any affiliation (adjunct/affiliated clinical faculty or primary appointment)

[c] P-values for years interpreting melanocytic skin lesions and for monthly melanocytic caseload are based on a trend test using a single 4 category ordinal variable in the context of a logistic regression model of misclassification, with clustering to account for within and between participant variability.

[d] P-values for melanocytic skin lesion composition of caseload are based on a trend test using a single 3 category ordinal variable in the context of a logistic regression model of misclassification, with clustering to account for within and between participant variability.

Table 2

Pathologist Characteristics and Rates of Reproducibility

	Reproducibility (Agreement of Phase II interpretation with Phase I interpretation by the same pathologist)		
Pathologist Characteristics	**Pathologists (n)**	**% (95% CI)**	**P-value**
Dermatopathology board certification and/or fellowship training			
No	69	65% (62%, 67%)	
Yes	49	70% (68%, 73%)	.001
Years interpreting melanocytic skin lesions			
< 5	18	62% (58%, 66%)	
5-9	27	68% (65%, 72%)	.47[b]
10-19	39	69% (66%, 72%)	
20+	34	65% (62%, 69%)	
Academic affiliation			
No	87	66% (64%, 68%)	
Yes	31	69% (66%, 73%)	.14
% of caseload that is melanocytic skin lesions			
< 10%	52	65% (62%, 67%)	
>/=10%	66	69% (67%, 71%)	.014
# of melanocytic cases per month			
<60	59	65% (62%, 67%)	
>/=60	59	69% (67%, 72%)	.009

[a] Reproducibility outcome is the agreement of the participant phase II interpretation with his/her phase I interpretation of the same case

[b] P-value for years interpreting MSL of .47 is based on a trend test using a single 4 category ordinal variable in the context of a logistic regression model of misclassification, with clustering to account for within and between participant variability. P-value of .01 based on test of 5+ vs <5 years interpreting MSL. This dichotomization is used in the multivariable model.