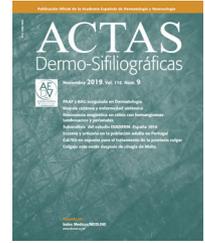




ACADEMIA ESPAÑOLA
DE DERMATOLOGÍA
Y VENEREOLÓGIA

ACTAS Dermo-Sifiliográficas

Full English text available at
www.actasdermo.org



ORIGINAL

[Artículo traducido] Variabilidad interobservador en el análisis histológico del melanoma: análisis de 60 casos



V. Sanz-Motilva^a, A. Martorell^a, E. Manrique-Silva^b, L. Terradez-Mas^c, C. Requena^b,
V. Traves^d, O. Sanmartín^b, J.L. Rodríguez-Peralto^e y E. Nagore^{b,*}

^a Department of Dermatology, Hospital de Manises, Valencia, España

^b Department of Dermatology, Fundación Instituto Valenciano de Oncología, Valencia, España

^c Department of Pathology, Hospital Clínico Universitario de Valencia, Valencia, España

^d Department of Pathology, Fundación Instituto Valenciano de Oncología, Valencia, España

^e Department of Pathology, Hospital Universitario 12 de Octubre, Madrid, España

Recibido el 13 de septiembre de 2023; aceptado el 15 de mayo de 2024

Disponible en Internet el 3 de abril de 2025

PALABRAS CLAVE

Melanoma;
Características
histológicas;
Reproducibilidad;
Acuerdo
interobservador

Resumen

Introducción: La dermatopatología no es una ciencia exacta estando sujeta a la subjetividad personal, lo que en ocasiones provoca variabilidad interobservador en cuanto al diagnóstico y la valoración de determinadas características histológicas. Respecto al melanoma, algunas variables como la regresión presentan baja concordancia interobservador. Por el contrario, otras variables como la medición del espesor de Breslow muestran una alta reproducibilidad. **Objetivo:** El principal objetivo de nuestro estudio fue investigar la reproducibilidad interobservador de múltiples características sobre un total de 60 casos consecutivos de melanoma. **Métodos y resultados principales:** Se realizó un estudio observacional y descriptivo en el Hospital de Manises (Manises), la Fundación IVO (Valencia) y el Hospital 12 de Octubre (Madrid). La concordancia media de las variables del estudio fue moderada. Las mayores tasas de concordancia se obtuvieron para la morfología polipoide, la pigmentación, la ulceración y la elastosis solar. Por el contrario, la concordancia más baja se dio para la presencia de pleomorfismo celular y la necrosis tumoral. Es necesario al realizar un estudio de investigación, analizar las variables de estudio y demostrar su validez para medir o clasificar una determinada característica. Adicionalmente, es recomendable garantizar que las variables sean reproducibles para poder utilizarlas en otros estudios o en la práctica clínica habitual.

Véase contenido relacionado en DOI: <https://doi.org/10.1016/j.ad.2024.05.023>

* Autor para correspondencia.

Correo electrónico: eduardo.nagore@ucv.es (E. Nagore).

<https://doi.org/10.1016/j.ad.2025.04.004>

0001-7310/© 2024 AEDV. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licencias/by-nc-nd/4.0/>).

KEYWORDS

Malignant melanoma;
Histological features;
Reproducibility;
Interobserver
agreement

Conclusiones: Nuestra concordancia media fue moderada, lo que refleja que algunas de las características medidas como el pleomorfismo celular o la presencia de necrosis no pueden ser utilizadas para futuros estudios o, por el contrario, deben ser redefinidas y restablecida su reproducibilidad.

© 2024 AEDV. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licencias/by-nc-nd/4.0/>).

Interobserver Variability in the Histopathological Evaluation of Melanoma: Analysis of 60 Cases

Abstract

Introduction: Since the field of dermatopathology is not an exact science, it is subject to personal subjectivity, which sometimes causes disagreements on the diagnosis and assessment of some histological features. In the case of melanoma, some variables such as regression are associated with low interobserver agreement. On the contrary, other variables such as the measurement of Breslow thickness show high reproducibility.

Objective: The main objective of our study was to investigate multiple features of 60 consecutive cases of melanoma to establish interobserver reproducibility.

Methods and main results: We conducted an observational and descriptive study at Hospital de Manises, Valencia, Spain, IVO Foundation, Valencia, Spain, and Hospital 12 de Octubre, Madrid, Spain. The mean level of agreement of all study variables was moderate (Cohen's kappa coefficient statistic = 0.5). The highest agreement corresponded to polypoid morphology, pigmentation, ulceration, and solar elastosis. On the other hand, the lowest level agreement was reached for the presence of cellular pleomorphism and tumor necrosis.

Conclusions: Our mean level of agreement was moderate, which reflects that some of the measured characteristics such as cellular pleomorphism or the presence of necrosis cannot be used for future studies or must be redefined and their reproducibility, reestablished. When conducting a research study, it is necessary to analyze the study variables to demonstrate their validity to measure or classify a certain feature. It is also advisable to warrant that the variables are reproducible to be able to use them for other studies or in the routine clinical practice.

© 2024 AEDV. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

Para realizar el diagnóstico histopatológico del melanoma cutáneo se suelen utilizar múltiples características morfológicas. Estas pueden clasificarse en dos tipos principales: características arquitectónicas (infiltración pagetoide y elastosis solar) y citológicas (atipia melanocítica y tasa de mitosis, entre otras).

Varias de estas características tienen implicaciones pronósticas, siendo las más importantes la presencia de ulceración y el grosor de Breslow.

El estudio del perfil molecular está cambiando la forma de ver y clasificar las lesiones melanocíticas, y en los últimos años se han producido avances interesantes en este campo¹.

Además, llama la atención lo heterogéneo que puede ser el melanoma en cuanto a su aspecto histopatológico. Aunque esto se debe probablemente a múltiples factores, el estado mutacional está ganando interés, ya que podría determinar algunas características clínicas e histopatológicas, como el grado de diseminación pagetoide o la forma y el tamaño de las células del melanoma. En un estudio anterior, Viros et al.² definieron algunas características

histopatológicas asociadas a la presencia de mutaciones en *BRAF*. Los melanomas mutados en *BRAF* mostraban una mayor migración ascendente y formación de nidos de melanocitos intraepidérmicos, una demarcación más nítida con la piel circundante, engrosamiento de la epidermis implicada, melanocitos más redondeados, grandes y pigmentados, y engrosamiento de la epidermis implicada.

Los informes histopatológicos de las características tumorales están sujetos a la variabilidad interobservador, por lo que el diagnóstico y la descripción del melanoma a menudo pueden ser problemáticos. En las lesiones melanocíticas, varios estudios han informado de un bajo nivel de concordancia para algunas características semicuantitativas, como la infiltración linfocitaria o la regresión³⁻⁶. Además, el diagnóstico diferencial entre nevos displásicos y melanoma invasivo en estadios tempranos algunas veces también puede resultar difícil^{7,8}.

Por otro lado, algunas características pronósticas suelen tener una alta reproducibilidad interobservador, como el grosor de Breslow y la presencia de ulceración^{4,6,9}.

El objetivo principal del presente estudio era evaluar la concordancia entre los dos grupos para cada característica patológica evaluada.

Material y métodos

Panel de clínicos y procedimiento de revisión

Nuestro estudio fue realizado por dos grupos de investigadores; el primero incluía dos dermatopatólogos de la misma institución (IVO) y el segundo, un patólogo (HUDO) y un dermatólogo (HM).

El dermatólogo realizó la evaluación histológica siempre bajo la supervisión y estando acompañado por el patólogo: así mismo, este tenía experiencia en dermatopatología (ya que realizó una estancia de 4 meses en la unidad de dermatopatología del Hospital Universitario 12 de Octubre, en Madrid, España y otra estancia corta de 15 días en la *Ackerman Academy*, Nueva York, NY, Estados Unidos).

El estudio se llevó a cabo en dos etapas. La primera etapa fue la preselección de las características histológicas a estudiar, así como sus definiciones. Esto se hizo teniendo en cuenta la literatura previa, concretamente el trabajo de investigación de Viros et al.², y las reuniones preliminares del grupo de investigación en las que se examinó un subconjunto de 40 muestras con diferentes estados mutacionales. Se decidió incluir algunas características adicionales basándose en su prevalencia. Esta etapa también incluyó la evaluación del nivel de concordancia en una serie de 10 muestras de práctica para poder identificar y resolver discrepancias en la evaluación de las variables definitivas. Esta evaluación fue realizada por separado por cada uno de los dos grupos.

Por último, la última etapa se realizó de la misma manera, pero con el conjunto de muestras finales sobre el que se llevó a cabo el presente estudio.

Casos estudiados

Los cortes histológicos se obtuvieron de manera consecutiva de los archivos de anatomía patológica del IVO, desde enero de 2004 hasta diciembre de 2004.

Los datos demográficos (edad, sexo, localización, tipo de melanoma, grosor de Breslow y nivel de Clark) se obtuvieron de la base de datos de melanomas del IVO.

Los fragmentos de tejido se fijaron en formol, se procesaron de forma rutinaria y se tiñeron con hematoxilina y eosina. Se excluyeron todos los casos que no eran óptimos para la revisión.

Características histopatológicas

Las características específicas analizadas por Viros et al.² que se incluyeron en nuestro análisis fueron:

La extensión ascendente de los melanocitos intraepidérmicos o extensión pagetoide (100% de melanocitos en la unión dermoepidérmica, 75% hasta 100% de melanocitos en la unión dermoepidérmica, 50% de extensión pagetoide, > 50% de extensión pagetoide), la formación de nidos de melanocitos intraepidérmicos (sin nidos, < 25% de melanocitos en nidos, 25% hasta 50% de melanocitos en nidos, > 50% de melanocitos en nidos), la pigmentación de los melanocitos (ausencia, ligera, moderada, alta, muy alta), el contorno epidérmico (atrófico, adelgazado, normal, engro-

sado, hipertrófico) y la circunscripción lateral (discontinua, gradual pero continua, abrupta).

Dado su potencial interés también se incluyeron algunas características adicionales:

El patrón de crecimiento (expansivo vs. infiltrativo), el grado de elastosis solar (categorizada de dos formas: 1 - ausencia, mínima, moderada y elevada; y 2 - bajo frente a elevado), la presencia de ulceración (ausencia frente a presencia), el tipo de ulceración (expansiva frente a infiltrativa), la forma polipoide (ausencia frente a presencia), la regresión (ausencia, $\leq 50\%$, $> 50\%$), la presencia de necrosis (ausencia frente a presencia) y la presencia de pleomorfismo (ausencia frente a presencia).

Las definiciones se describen detalladamente en el [anexo, en los datos adicionales](#).

Análisis estadístico

Los resultados se exportaron a una tabla de Excel. Cuando una característica o un caso completo por cualquier motivo no podía ser evaluado, se consideró como no aplicable.

El acuerdo entre los dos grupos se evaluó mediante el coeficiente kappa de Cohen (κ), que es un índice conocido para medir la concordancia corregida por el azar en una escala nominal u ordinal. Según Landis y Koch¹⁰, los valores superiores a 0,75 representan una concordancia excelente más allá de la casualidad, los valores entre 0,75 y 0,40 representan una concordancia entre regular y buena más allá de la casualidad, y los valores $< 0,40$ representan una concordancia deficiente más allá de la casualidad. Un valor de κ cercano a uno significa un acuerdo casi perfecto¹⁰.

En el caso de las variables ordinales, utilizamos el kappa ponderado de Cohen —una modificación del estadístico kappa original— propuesto para variables nominales en presencia de dos observadores¹¹.

Para cada característica considerada, se construyó una tabla diagnóstica de 2×2 utilizando categorías dicotómicas, y se calcularon los valores específicos de κ .

Resultados

Se seleccionaron los 60 primeros casos consecutivos diagnosticados en el servicio de dermatología del IVO desde enero de 2004 hasta diciembre de 2004.

Se excluyeron 5 casos, 4 de ellos por no poder ser interpretados debido a su pequeño tamaño y uno por duplicidad (2 secciones de un mismo caso). Así, finalmente se consideraron 55 casos válidos para establecer los valores kappa.

La población del estudio incluyó 55 pacientes, 28 (50,9%) hombres y 27 (49,1%) mujeres, con una edad media en el momento del diagnóstico de 58,9 años (rango, 23-82 años). Los datos sobre la localización del tumor primario, el tipo histológico y el estadio tumoral se muestran en la [tabla 1 de los datos adicionales](#).

La concordancia general fue moderada (valor kappa medio, 0,5). Los valores máximos correspondieron a la presencia de la forma polipoide (0,8), la pigmentación (0,7), la presencia de ulceración (0,7), la elastosis (CSD alto/CSD bajo) (0,7) y el grado de elastosis (0,7).

Los valores más discordantes fueron la presencia de pleomorfismo (0,2) y de necrosis (0,3).

Tabla 1 Valores kappa de los estudios actuales y anteriores

	Viros et al. ²	Broekaert et al. ¹²	Nuestro estudio
Extensión	0,7	0,7	0,5
Formación de nidos	0,6	0,4	0,5
Pigmentación	0,7	0,7	0,7
Circunscripción lateral	0,4	0,4	0,3
Contorno epidérmico	0,5	0,4	0,4
Elastosis solar (4 grados)	0,8	0,6	0,7
Patrón de crecimiento	N/A	N/A	0,4
Elastosis (alto CSD/bajo CSD)	N/A	N/A	0,7
Ulceración	N/A	N/A	0,7
Tipo de ulceración	N/A	N/A	0,4
Forma polipoide	N/A	N/A	0,8
Regresión	N/A	N/A	0,5
Necrosis	N/A	N/A	0,3
Pleomorfismo	N/A	N/A	0,2

La **tabla 1** muestra los valores kappa de diferentes estudios, concretamente nuestros resultados y los de los estudios realizados por Viros et al.² y Broekaert et al.¹².

Por último, realizamos un ajuste de las variables de la extensión pagetoide y de la formación de nidos intraepidérmicos y evaluamos la concordancia interobservador. En concreto, redujimos el número de categorías de 4 a 2.

El valor de κ para la extensión pagetoide fue de 0,7 y el valor de κ para la formación de nidos, de 0,4.

Discusión

En el presente estudio evaluamos la reproducibilidad o concordancia interobservador de algunas características de 60 casos de melanoma maligno, algunas de ellas con implicaciones pronósticas. La concordancia global fue moderada. Los valores kappa más elevados correspondieron a la forma polipoide y la elastosis solar (CSD alto/CSD bajo). Por otro lado, los valores más bajos fueron para la necrosis y el pleomorfismo.

El melanoma es un tumor heterogéneo, ya que presenta diferentes características clínicas e histopatológicas, lo que a veces dificulta el diagnóstico.

Estudios anteriores han evaluado la reproducibilidad interobservador de los criterios diagnósticos en el melanoma y otras lesiones melanocíticas con resultados heterogéneos^{2,4,8,9,13-16}.

En estos estudios las secciones patológicas de melanomas cutáneos o lesiones melanocíticas se distribuyen a diferentes combinaciones de patólogos y dermatopatólogos para así categorizar diversas variables histopatológicas.

Las mayores tasas de concordancia se alcanzaron para la variable más importante en cuanto al valor pronóstico, el grosor de Breslow. Esto es algo predecible al tratarse de una variable cuantitativa^{4,6,9,14,17-20}.

Por el contrario, algunas otras variables con importancia pronóstica muestran una reproducibilidad baja o moderada, como la valoración del nivel de Clark^{9,14,19,20}.

Nuestro estudio se basa en varias variables que fueron definidas en el estudio de Viros et al.². Este estudio mostró mejores valores kappa para la mayoría de las variables

evaluadas en comparación con nuestros resultados² (ver **tabla 1**).

Otro estudio realizado por el mismo grupo (Broekaert et al.)¹² presentó valores kappa similares, excepto para la formación de nidos, el contorno epidérmico y la circunscripción que fueron inferiores a los valores obtenidos en el estudio realizado por Viros et al.² (**tabla 1**).

Una explicación de algunas de las discordancias señaladas sería que el melanoma puede ser una lesión muy extensa y mostrar una heterogenicidad manifiesta *per se*. Por lo tanto, para determinar la formación de nidos o la extensión pagetoide, podría existir una zona con elevado número de nidos o extensión intraepidérmica seguida de otra con ausencia total, lo que complica la obtención de un resultado que cuantifique este hallazgo. Además, algunos rasgos sólo pueden detectarse en algunos cortes histológicos.

En cuanto a la circunscripción lateral y el contorno epidérmico, hemos detectado que, a menudo, el melanoma puede mostrar diferentes transiciones de la porción de crecimiento intraepidérmico del tumor a la piel normal de uno a otro lado del corte del cristal.

Cabe destacar que la elastosis solar fue más reproducible que otras variables cuando se evaluó en dos categorías (CSD alto/CSD bajo), una forma más sencilla pero válida de clasificar la elastosis, definida por primera vez por Landi et al.²¹ y considerada por la Organización Mundial de la Salud (OMS) como una característica importante que categoriza los distintos tipos de melanoma²².

La concordancia kappa para la ulceración fue sustancial, como han afirmado estudios anteriores^{6,9,17,19,20}. La evaluación de la ulceración sólo es posible en algunas secciones histológicas de melanoma; por ejemplo, la ulceración sólo puede afectar una pequeña porción de melanoma y pasar desapercibida en algunos cortes.

Cuando existe una pérdida focal de epidermis, puede ser problemático establecer la ulceración, ya que se podría tratar de una ulceración real o de un artefacto del corte. A menos que haya evidencia de una cicatriz dérmica o una biopsia previa, puede ser problemático distinguir entre una ulceración traumática y no traumática⁹.

En nuestra experiencia, la forma polipoide fue muy reproducible, ya que es fácil de evaluar.

En cuanto a la regresión, en nuestro estudio, su evaluación mostró una concordancia moderada, habiéndose comunicado tasas heterogéneas en estudios anteriores. La mayoría de los estudios que evaluaban diferentes características mostraron una baja reproducibilidad^{6,19,20}, mientras que otros mostraron tasas más altas²³. Kang et al. mostraron una mejor reproducibilidad para la regresión en su estudio, sin embargo, debemos mencionar que este estudio se centró únicamente en la regresión. Cabe esperar que un estudio asociado a una sola característica alcance mayor concordancia que los estudios que evalúan varias variables, especialmente si existe un entrenamiento previo. La categorización de la regresión y los criterios han cambiado con el tiempo, lo que puede ser una causa de la concordancia variable. De hecho, algunos estudios consideran regresión precoz, intermedia o tardía²³, mientras que otros sólo consideran dos grados (presencia o leve/ausencia)^{19,20}.

La literatura es escasa sobre la reproducibilidad interobservador en lo que respecta a la necrosis y el pleomorfismo.

La variable pleomorfismo es muy subjetiva, y las células de melanoma son pleomórficas en general, sobre todo en algunas zonas del tumor. Esta característica puede ser problemática ya que puede ser difícil de establecer debido a su subjetividad.

En cuanto a la evaluación de la necrosis, en el estudio de Urso et al.¹³ se estableció la concordancia interobservador para 55 casos de melanoma y sólo se encontró necrosis en un caso. Por lo tanto, al ser una característica menos prevalente, puede ser un hallazgo que sólo se presente en un área pequeña y pase desapercibido.

Por último, hay un hecho realmente importante que debemos mencionar que es la subjetividad, ya que se ha confirmado que incluso en algunas observaciones existe discordancia intraobservador.

En consonancia con esto, Elmore et al. mostraron discordancia intraobservador (y lógicamente menor discordancia interobservador) para categorizar la atipia en lesiones melanocíticas, especialmente al categorizar aquellas con valores diferentes a los valores extremos⁸.

Proponemos que, si es posible, se redefinan las variables y se reagrupen con menos categorías. Por ejemplo, la extensión pagetoide podría categorizarse en 2 grupos en lugar de 4, considerando 0: ausencia o extensión pagetoide menor (< 25% de las células) y 1: extensión pagetoide manifiesta (> 25% de las células). Del mismo modo, la formación de nidos podría redefinirse en 2 categorías: 0: ausencia o formación mínima de nidos (< 25% de células en nidos intraepidérmicos) frente a 1: formación marcada de nidos (> 25% de células en nidos intraepidérmicos). La forma de las células, el contorno epidérmico y la pigmentación también podrían simplificarse en menos categorías.

Por el contrario, hay algunas variables con baja concordancia que no pueden ajustarse, como la circunscripción lateral. Por lo tanto, es probable que no utilicemos esta variable en el futuro.

Cualquier ajuste de las categorías de la variable, como su reagrupación, si es posible, deberá someterse de nuevo a pruebas de reproducibilidad.

Según nuestra experiencia, tras reagrupar y estudiar la concordancia interobservadores de la extensión pagetoide y la formación de nidos, sólo en la extensión pagetoide esta

mejoró. Es un ejemplo de que la reagrupación de categorías de variables puede a veces mejorar su reproducibilidad o, al menos, mantenerla, asegurando al mismo tiempo que su relevancia permanezca intacta.

Quizá aún sea pronto, pero las nuevas tecnologías y la inteligencia artificial podrían suponer un antes y un después en la categorización histológica de las lesiones melanocíticas. Para ello, es necesario generar algoritmos útiles que deben ser entrenados hasta que su sensibilidad diagnóstica se aproxime a la de un patólogo experimentado²⁴.

De hecho, se ha probado un modelo informático histológico tridimensional de melanoma maligno con resultados prometedores. En él se evaluaban diferentes niveles de tejido y se evitaba el problema de que algunas características estuvieran presentes en un subconjunto limitado de cristales. Aunque se trata de tecnologías prometedoras en el campo del diagnóstico histopatológico, todavía existen algunas limitaciones que deben abordarse²⁵.

En conclusión, nuestro estudio demostró una reproducibilidad media moderada de varias características histopatológicas del melanoma, lo que significa que, si es posible, algunas variables deben ser redefinidas y evaluadas para lograr un acuerdo interobservador en futuras investigaciones.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Anexo. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en [doi:10.1016/j.ad.2024.05.023](https://doi.org/10.1016/j.ad.2024.05.023).

Bibliografía

1. Fernandez-Flores A. Modern concepts in melanocytic tumors. *Actas Dermosifiliogr.* 2023;114:402–12, <http://dx.doi.org/10.1016/j.ad.2023.01.001>.
2. Viros A, Fridlyand J, Bauer J, Lasithiotakis K, Garbe C, Pinkel D, et al. Improving melanoma classification by integrating genetic and morphologic features. *PLoS Med.* 2008;5:e120, <http://dx.doi.org/10.1371/journal.pmed.0050120>.
3. 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;100:318S–21S, <http://dx.doi.org/10.1111/1523-1747.ep12470215>.
4. Murali R, Hughes MT, Fitzgerald P. Interobserver variation in the histopathologic reporting of key prognostic parameters, particularly Clark level, affects pathologic staging of primary cutaneous melanoma. *Ann Surg.* 2009;249:641–7, <http://dx.doi.org/10.1097/SLA.0b013e31819ed973>.
5. Duray PH, DerSimonian R, Barnhill R, Stenn K, Ernstoff MS, Fine J, 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;27:741–9, [http://dx.doi.org/10.1016/0190-9622\(92\)70248-e](http://dx.doi.org/10.1016/0190-9622(92)70248-e).
6. Corona R, Mele A, Amini M, De Rosa G, Coppola G, Piccardi P, et al. Interobserver variability on the histopathologic diagnosis of cutaneous melanoma and other

- pigmented skin lesions. *J Clin Oncol*. 1996;14:1218–23, <http://dx.doi.org/10.1200/JCO.1996.14.4.1218>.
7. Gerami P, Busam K, Cochran A, Cook MG, Duncan LM, Elder DE, et al. Histomorphologic assessment and interobserver diagnostic reproducibility of atypical spitzoid melanocytic neoplasms with long-term follow-up. *Am J Surg Pathol*. 2014;38:934–40, <http://dx.doi.org/10.1097/PAS.000000000000198>.
 8. 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, <http://dx.doi.org/10.1136/bmj.j2813>.
 9. Scolyer RA, Bs MB, Shaw HM, Thompson JF, Li LL, Colman MH, et al. Interobserver reproducibility of histopathologic prognostic variables in primary cutaneous melanomas. *Am J Surg Pathol*. 2003;27:1571–6, <http://dx.doi.org/10.1097/00000478-200312000-00011>.
 10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159, <http://dx.doi.org/10.2307/2529310>.
 11. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70:213–20, <http://dx.doi.org/10.1037/h0026256>.
 12. Broekaert SMC, Roy R, Okamoto I, Van Den Oord J, Bauer J, Garbe C, et al. Genetic and morphologic features for melanoma classification. *Pigment Cell Mel Res*. 2010;23:763–70, <http://dx.doi.org/10.1111/j.1755-148X.2010.00778.x>.
 13. Urso C, Rongioletti F, Innocenzi D, Saieva C, Batolo D, Chimenti S, et al. Interobserver reproducibility of histological features in cutaneous malignant melanoma. *J Clin Pathol*. 2005;58:1194–8, <http://dx.doi.org/10.1136/jcp.2005.026765>.
 14. Eriksson H, Hedblad MFM, Hellborg HJLK, Krawiec K, Lundh B, Eva R, et al. Interobserver variability of histopathological prognostic parameters in cutaneous malignant melanoma: impact on patient management. *Acta Derm Venerol*. 2013;93:411–6, <http://dx.doi.org/10.2340/00015555-1517>.
 15. Monshizadeh L, Hanikeri M, Beer TW, Heenan PJ. A critical review of melanoma pathology reports for patients referred to the Western Australian Melanoma Advisory Service. *Pathology*. 2012;44:441–7, <http://dx.doi.org/10.1097/PAT.0b013e328355767e>.
 16. Niebling MG, Haydu LE, Karim RZ, Thompson JF, Scolyer RA. Reproducibility of AJCC staging parameters in primary cutaneous melanoma: an analysis of 4,924 cases. *Ann Surg Oncol*. 2013;20:3969–75, <http://dx.doi.org/10.1245/s10434-013-3092-5>.
 17. Colloby PS, West KP, Fletcher A. Observer variation in the measurement of Breslow depth and Clark's level in thin cutaneous malignant melanoma. *J Pathol*. 1991;163:245–50, <http://dx.doi.org/10.1002/path.1711630310>.
 18. Prade M, Sancho-Garnier H, Cesarini JP, Cochran A. Difficulties encountered in the application of Clark classification and the Breslow thickness measurement in cutaneous malignant melanoma. *Int J Cancer*. 1980;26:159–63, <http://dx.doi.org/10.1002/ijc.2910260206>.
 19. Heenan PJ, Matz LR, Blackwell JB, Kelsall GR, Singh A, ten Seldam RE, et al. Inter-observer variation between pathologists in the classification of cutaneous malignant melanoma in western Australia. *Histopathology*. 1984;8:717–29, <http://dx.doi.org/10.1111/j.1365-2559.1984.tb02388.x>.
 20. Lock-Andersen J, Hou-Jensen K, Hansen JP, Jensen NK, Sogaard H, Andersen PK. Observer variation in histological classification of cutaneous malignant melanoma. *Scand J Plast Reconstr Surg Hand Surg*. 1995;29:141–8, <http://dx.doi.org/10.3109/02844319509034330>.
 21. Landi MT, Bauer J, Pfeiffer RM, Elder DE, Hulley B, Minghetti P, et al. MC1R germline variants confer risk for BRAF-mutant melanoma. *Science*. 2006;313:521–2, <http://dx.doi.org/10.1126/science.1127515>.
 22. Elder DE, Bastian BC, Cree IA, Massi D, Scolyer RA. The 2018 World Health Organization classification of cutaneous, mucosal, and uveal melanoma: detailed analysis of 9 distinct subtypes defined by their evolutionary pathway. *Arch Pathol Lab Med*. 2020;144:500–22, <http://dx.doi.org/10.5858/arpa.2019-0561-RA>.
 23. Kang S, Barnhill RL, Mihm MCJ, Sober AJ. Histologic regression in malignant melanoma: an interobserver concordance study. *J Cutan Pathol*. 1993;20:126–9, <http://dx.doi.org/10.1111/j.1600-0560.1993.tb00228.x>.
 24. Brinker TJ, Schmitt M, Kriehoff-Henning EI, Barnhill R, Beltraminelli H, Braun SA, et al. Diagnostic performance of artificial intelligence for histologic melanoma recognition compared to 18 international expert pathologists. *J Am Acad Dermatol*. 2022;86:640–2, <http://dx.doi.org/10.1016/j.jaad.2021.02.009>.
 25. Kurz A, Krahl D, Kutzner H, Barnhill R, Perasole A, Fernández-Figueras MT, et al. A 3-dimensional histology computer model of malignant melanoma and its implications for digital pathology. *Eur J Cancer*. 2023;193:113294, <http://dx.doi.org/10.1016/j.ejca.2023.113294>.