Estudios y Perspectivas Revista Científica y Académica
ISSN en línea 3005-2599, enero-marzo 2026,
Volumen 6, Número 1.

DOI: https://doi.org/10.61384/r.c.a.v6i1

INTELIGENCIA ARTIFICIAL PARA EL DIAGNÓSTICO DE MELANOMA: RENDIMIENTO, SESGOS Y GENERALIZACIÓN. REVISIÓN SISTEMÁTICA BASADA EN PRISMA

ARTIFICIAL INTELLIGENCE FOR MELANOMA DIAGNOSIS: PERFORMANCE, BIAS, AND GENERALIZABILITY. A PRISMA-BASED SYSTEMATIC REVIEW

Victoria Marisol San Martin Ovalle
Universidad Autónoma de Tamaulipas, México

Mary Ana Adonia Machore
Universidad de Montemorelos, México

Victoria Lopez Estudillo
Universidad Autónoma Metropolitana, México

Niccole Czestohowa Trejo Hernandez
Universidad Autónoma de Guadalajara, México

E. COUTURE
Educación y Ciencias Sociales

Estudios y Perspectivas
Revista Científica y Académica

Estudios y Perspectivas
Revista Científica Multidisciplinaria

DOI: https://doi.org/10.61384/r.c.a..v6i1.1881

Inteligencia Artificial para el Diagnóstico de Melanoma: Rendimiento, Sesgos Y Generalización. Revisión Sistemática Basada En PRISMA

Victoria Marisol San Martin Ovalle	**Mary Ana Adonia Machore**
marysolovalle11@gmail.com	maryadonia@gmail.com
https://orcid.org/0009-0002-7356-0597	https://orcid.org/0009-0003-7652-4134
Universidad Autónoma de Tamaulipas, México	Universidad de Montemorelos, México
Victoria Lopez Estudillo	**Niccole Czestohowa Trejo Hernandez**
vikyle051100@gmail.com	niccolecz97@gmail.com
https://orcid.org/0009-0003-2519-7743	https://orcid.org/0009-0003-5785-3143
Universidad Autónoma Metropolitana, México	Universidad Autónoma de Guadalajara, México

RESUMEN

Introducción: La inteligencia artificial (IA), especialmente el aprendizaje profundo aplicado a imágenes dermatoscopias y clínicas, ha demostrado alto rendimiento en el diagnóstico de melanoma. No obstante, su traslación clínica depende no solo de métricas de exactitud, sino también de la calidad metodológica, los sesgos de datos, la validez externa y la interacción humano–IA. El objetivo fue sintetizar la evidencia sobre el rendimiento diagnóstico de la IA, describir sesgos relevantes y evaluar su generalización en escenarios clínicos reales

Métodos: Se realizó una revisión sistemática conforme a PRISMA 2020. Se incluyeron estudios de exactitud diagnóstica, validaciones prospectivas y revisiones sistemáticas/meta-análisis. Dos revisores independientes efectuaron la selección y extracción de datos. El riesgo de sesgo se evaluó mediante QUADAS-2 u herramientas equivalentes

Resultados: Meta-análisis recientes reportan un AUC cercano a 0,96, con sensibilidad y especificidad aproximadas de 0,89 y 0,92. La IA muestra rendimiento comparable o superior al de dermatólogos expertos en entornos controlados. Estudios prospectivos confirman su utilidad clínica, aunque con variabilidad según prevalencia y condiciones de uso. Persisten sesgos por subrepresentación de fototipos altos, desequilibrio de clases y atajos algorítmicos. La generalización es heterogénea y depende de la validación externa y diversidad de datos.

Conclusiones: La IA es una herramienta prometedora en el diagnóstico de melanoma, pero su implementación segura requiere diversidad de datos, evaluación prospectiva, métricas calibradas y monitoreo continuo de sesgos

***Palabras clave*:** melanoma, inteligencia artificial, dermatoscopia, exactitud diagnóstica, sesgo algorítmico

Artificial Intelligence for Melanoma Diagnosis: Performance, Bias, And Generalizability. A PRISMA-Based Systematic Review

ABSTRACT

Introduction: Artificial intelligence (AI), particularly deep learning applied to dermoscopic and clinical images, has demonstrated high performance in melanoma diagnosis. However, its clinical translation depends not only on accuracy metrics but also on methodological quality, data biases, external validity, and human–AI interaction. The aim of this study was to synthesize the evidence on AI diagnostic performance, describe relevant biases, and evaluate its generalizability in real-world clinical settings

Methods: A systematic review was conducted in accordance with PRISMA 2020. Studies on diagnostic accuracy, prospective validations, and systematic reviews/meta-analyses were included. Two independent reviewers performed study selection and data extraction. Risk of bias was assessed using QUADAS-2 or equivalent tools

Results: Recent meta-analyses report an area under the curve (AUC) close to 0.96, with sensitivity and specificity of approximately 0.89 and 0.92, respectively. AI shows performance comparable to or exceeding that of expert dermatologists in controlled settings. Prospective studies confirm its clinical utility, although variability exists depending on prevalence and real-world conditions. Persistent biases include underrepresentation of darker skin types, class imbalance, and algorithmic shortcuts. Generalizability is heterogeneous and depends on external validation and data diversity

Conclusions: AI is a promising tool for melanoma diagnosis; however, its safe implementation requires diverse datasets, prospective evaluation, calibrated metrics, and continuous monitoring of bias

***Keywords*:** melanoma, artificial intelligence, dermoscopy, diagnostic accuracy, algorithmic bias

Recibido: 20- octubre-2025 / Aprobado: 30-enero-2026

INTRODUCCIÓN

El melanoma cutáneo representa una de las neoplasias de piel con mayor potencial letal; su pronóstico depende en gran medida de la detección temprana y la excisión oportuna. La dermatoscopia y la experiencia clínica han demostrado mejorar la exactitud diagnóstica; sin embargo, persisten variaciones interobservador y limitaciones en contextos de atención primaria y en regiones con acceso restringido a servicios dermatológicos. En este escenario, la inteligencia artificial (IA), particularmente las redes neuronales convolucionales y las arquitecturas multimodales, se ha propuesto como una herramienta de apoyo diagnóstico para el triaje, la toma de decisiones sobre biopsia y la priorización de derivaciones (Esteva et al., 2017; Haenssle et al., 2018).

Desde los primeros modelos con rendimiento comparable al de dermatólogos utilizando imágenes clínicas, hasta los sistemas actuales entrenados con dermatoscopia y metadatos clínicos, la evidencia ha mostrado métricas elevadas en entornos controlados (Esteva et al., 2017; Haenssle et al., 2018). Asimismo, estudios comparativos con especialistas y evaluaciones de algoritmos integrados al flujo clínico refuerzan el potencial de la IA para reducir omisiones diagnósticas, especialmente en lesiones sutiles (Salinas et al., 2024; Ertürk Zararsız et al., 2025; Heinlein et al., 2024; Marchetti et al., 2023).

No obstante, el entusiasmo inicial ha dado paso a preocupaciones metodológicas relevantes, entre ellas los sesgos de selección, la falta de validación externa, el entrenamiento sobre conjuntos de datos poco diversos, el uso de métricas no ajustadas a la prevalencia real y el riesgo de “atajos” algorítmicos que pueden comprometer la equidad y la seguridad clínica (Alipour et al., 2024; Ruga, 2025; Bissoto et al., 2019; Daneshjou et al., 2021).

A diferencia de otras tecnologías diagnósticas, la IA es altamente dependiente del dominio de los datos; por ello, variaciones en la cámara, la iluminación, la población o el espectro de lesiones pueden afectar significativamente su rendimiento. En este contexto, la presente revisión tiene como objetivo sintetizar la evidencia reciente sobre el desempeño diagnóstico de la IA en melanoma, así como analizar los sesgos identificados y los principales desafíos para su generalización en la práctica clínica.

METODOLOGÍA

Se realizó una revisión sistemática siguiendo las directrices de **PRISMA 2020** (Page et al., 2021). Dos revisores independientes llevaron a cabo el proceso en tres etapas: (1) búsqueda y selección de estudios mediante títulos y resúmenes, (2) evaluación de los textos completos y (3) extracción de datos. Las discrepancias se resolvieron por consenso entre los revisores.

Tabla 1. Estrategia de búsqueda y criterios de selección.

Elemento	Descripción
Bases de datos	MEDLINE/PubMed, Embase, Cochrane Library, Scopus, Web of Science, IEEE Xplore, arXiv, ClinicalTrials.gov, WHO-ICTRP
Periodo de búsqueda	2015–2025
Términos MeSH/keywords	"melanoma", "skin neoplasms", "dermoscopy", "artificial intelligence", "deep learning", "convolutional neural network", "diagnostic accuracy", "external validation", "bias", "fairness", "skin tone", "multimodal"
Criterios de inclusión	Estudios primarios de exactitud diagnóstica de IA en melanoma; estudios prospectivos o ensayos con aplicación clínica; revisiones sistemáticas y meta-análisis relevantes; estudios que aborden sesgos y generalización
Criterios de exclusión	Modelos no aplicados a melanoma, estudios en animales o no humanos, series con <10 lesiones sin verificación histológica, duplicados
Variables extraídas	Tipo de imagen, arquitectura del modelo, conjunto de entrenamiento, tipo de validación (interna, externa o prospectiva), sensibilidad, especificidad o AUC, comparador humano, análisis de sesgo y desempeño por subgrupos

Nota: Cuando alguna variable no fue reportada en los estudios incluidos, se consignó como *no reportado (NR)*.

RESULTADOS Y DISCUSIÓN

Características generales de la evidencia incluida

La literatura analizada se compone de:

a) estudios retrospectivos de clasificación en bases dermatoscópicas,

b) comparaciones humano–máquina en series controladas,

c) meta-análisis de exactitud diagnóstica, y

d) validaciones prospectivas en entornos clínicos y de atención primaria (Salinas et al., 2024; Ertürk Zararsız et al., 2025; Heinlein et al., 2024; Marchetti et al., 2023; Papachristou et al., 2024; Krakowski et al., 2024; Naseri et al., 2024; Yan et al., 2025).

La mayoría de los estudios emplea imágenes dermatoscópicas provenientes de repositorios, con heterogeneidad en la anotación, la prevalencia y la calidad de los datos (Patel et al., 2023; Alipour et al., 2024; Ruga, 2025; Bissoto et al., 2019; Daneshjou et al., 2021; Helenason et al., 2024; Kaur et al., 2025).

Tabla 2. Tipos de estudios y escenarios evaluados.

Tipo de evidencia	Escenario típico	Observaciones
Retrospectivos de clasificación	Bases dermatoscópicas públicas	Alta exactitud interna; riesgo de sobreajuste y espectro incompleto (Patel et al., 2023; Bissoto et al., 2019)
Comparaciones humano-IA	Lectura ciega con expertos/no expertos	IA comparable a dermatólogos expertos; mayor utilidad en no expertos (Salinas et al., 2024; Haenssle et al., 2018; Tschandl et al., 2019)
Meta-análisis de exactitud	Síntesis de múltiples estudios	Alta AUC y sensibilidad/especificidad; heterogeneidad metodológica (Salinas et al., 2024; Ertürk Zararsız et al., 2025)
Prospectivos clínicos	Dermatología/atención primaria multicéntrica	Confirman utilidad clínica; variabilidad según dispositivos y prevalencia (Heinlein et al., 2024; Papachristou et al., 2024; Marchetti et al., 2023)

Fuente: Autoría propia.

Rendimiento diagnóstico global

Un meta-análisis centrado en IA para melanoma reportó un área bajo la curva (AUC) aproximada de 0,96, con sensibilidad de 0,89 y especificidad de 0,92 en conjuntos de prueba, destacando un mejor desempeño de modelos híbridos multimodales frente a redes puras (Ertürk Zararsız et al., 2025).

De forma complementaria, una síntesis comparativa entre IA y clínicos en cáncer de piel evidenció que, en promedio, la IA supera a médicos generales y presenta un rendimiento similar al de especialistas, aunque con variabilidad según el tipo de tarea diagnóstica y la clase de lesión (Salinas et al., 2024).

Tabla 3. Métricas de rendimiento reportadas en meta-análisis recientes.

Fuente	Modalidad principal	Sensibilidad	Especificidad	AUC
Meta-análisis melanoma IA (Ertürk Zararsız et al., 2025)	Dermatoscopia	~0.89	~0.92	~0.96
IA vs clínicos cáncer de piel (Salinas et al., 2024)	Dermatoscopia/mixta	~0.87	~0.77	NR

Nota: Los valores aproximados se reportan conforme a los estudios originales; NR indica que la métrica no fue uniformemente reportada.

Comparación con dermatólogos y efecto del apoyo clínico

En evaluaciones ciegas, la IA alcanza niveles comparables a dermatólogos experimentados y supera a médicos no especialistas, especialmente en tareas binarias de clasificación melanoma/no melanoma (Esteva et al., 2017; Haenssle et al., 2018; Tschandl et al., 2019).

Asimismo, la integración de sistemas de IA como herramienta de apoyo mejora la exactitud diagnóstica de los clínicos sin incrementar significativamente la tasa de falsos positivos en la mayoría de los estudios, lo que sugiere una interacción sinérgica entre humano e inteligencia artificial (Krakowski et al., 2024; Tschandl et al., 2019).

Validación prospectiva y utilidad en el mundo real

Los estudios prospectivos evidencian la transición de la IA hacia la práctica clínica. En un ensayo multicéntrico, un sistema basado en dermatoscopia mejoró la detección de melanomas complejos y

superó el rendimiento promedio de dermatólogos en lectura asistida, manteniendo su eficacia pese a la variabilidad en los dispositivos de captura (Heinlein et al., 2024).

En atención primaria, herramientas basadas en IA demostraron alta capacidad de descarte y utilidad en la priorización de referencias, aunque su desempeño dependió de factores como la prevalencia real de melanoma y la calidad en la adquisición de imágenes (Papachristou et al., 2024).

Adicionalmente, la validación de algoritmos de código abierto ha mostrado resultados prometedores, aunque se resalta la necesidad de ensayos clínicos aleatorizados antes de su implementación generalizada (Marchetti et al., 2023).

Sesgos identificados

Los principales sesgos identificados se agrupan en tres dominios:

- Sesgo de datos: subrepresentación de fototipos cutáneos altos (V–VI), limitada variabilidad anatómica y espectro incompleto de lesiones benignas;
- Sesgo de diseño y validación: ausencia de cohortes consecutivas, uso de prevalencias artificiales y falta de validación externa;
- Sesgo algorítmico: aprendizaje de correlatos espurios (como marcas o artefactos de imagen) que inflan la exactitud interna sin trasladarse a la práctica clínica (Alipour et al., 2024; Ruga, 2025; Bissoto et al., 2019; Daneshjou et al., 2021).

Tabla 4. Sesgos principales y estrategias de mitigación.

Sesgo	Manifestación	Impacto	Mitigación propuesta
Falta de diversidad de piel	Predominio de fototipos claros en el entrenamiento	Disminución del desempeño en piel oscura (Alipour et al., 2024; Ruga, 2025; Daneshjou et al., 2021)	Repositorios balanceados por fototipo y auditorías de diversidad
Desequilibrio de clases	Predominio de lesiones benignas frente a melanoma	Inflación de la exactitud por clase mayoritaria (Patel et al., 2023; Alipour et al., 2024)	Re-muestreo, funciones de pérdida ponderadas y métricas por clase

Sesgo	Manifestación	Impacto	Mitigación propuesta
Atajos algorítmicos	Uso de artefactos o contexto no clínico	Sobreestimación del AUC interno (Bissoto et al., 2019)	Evaluaciones con imágenes descontextualizadas y uso de IA explicable
Prevalencia artificial	Conjuntos con prevalencia mayor a la real	Sensibilidad/especificidad poco trasladables (Ertürk Zararsız et al., 2025; Papachristou et al., 2024)	Validaciones prospectivas consecutivas
Falta de validación externa	Solo validación interna	Generalización incierta (Heinlein et al., 2024; Marchetti et al., 2023; Yan et al., 2025)	Estudios multicéntricos y evaluación en múltiples dominios

Fuente: Autoría propia.

Generalización: qué funciona y qué falla

Los modelos entrenados con datos multiinstitucionales y evaluados mediante validación externa o estudios prospectivos muestran mayor estabilidad frente a cambios en dispositivos, poblaciones y contextos clínicos (Heinlein et al., 2024; Papachristou et al., 2024; Marchetti et al., 2023; Yan et al., 2025).

En particular, los modelos multimodales que integran imágenes y datos clínicos han demostrado una menor degradación del rendimiento ante cambios de dominio, lo que sugiere una mayor robustez (Yan et al., 2025).

En contraste, los modelos entrenados en bases de datos homogéneas tienden a perder exactitud al ser aplicados en entornos clínicos distintos, especialmente ante la presencia de lesiones atípicas, localizaciones poco frecuentes o fototipos subrepresentados (Naseri et al., 2024; Patel et al., 2023; Alipour et al., 2024; Kaur et al., 2025).

La evidencia actual sitúa a la inteligencia artificial (IA) en un punto de alta madurez técnica y en plena transición hacia su aplicación clínica.

Los meta-análisis recientes no solo confirman una elevada exactitud diagnóstica, sino que también establecen un umbral de desempeño que resulta clínicamente relevante para el triaje y el apoyo a la toma de decisiones en dermatoscopia (Salinas et al., 2024; Ertürk Zararsız et al., 2025; Ye et al., 2024). Este hallazgo es consistente con la evolución histórica de los modelos fundacionales, que han pasado de alcanzar un rendimiento comparable al de especialistas a incorporar enfoques híbridos y multimodales más avanzados (Esteva et al., 2017; Haenssle et al., 2018; Tschandl et al., 2019).

No obstante, estas cifras globales pueden ocultar una importante fragilidad frente al cambio de dominio. El melanoma presenta una alta heterogeneidad en términos de morfología, localización y manifestación según el fototipo cutáneo; por lo tanto, si los conjuntos de entrenamiento no reflejan esta diversidad, los algoritmos tienden a aprender representaciones parciales de la enfermedad. Diversas revisiones sobre diversidad en datos dermatológicos y análisis de equidad coinciden en que los sistemas actuales muestran un menor desempeño en pieles más oscuras, no debido a limitaciones biológicas, sino a la falta de representatividad en los datos de entrenamiento (Alipour et al., 2024; Ruga, 2025; Daneshjou et al., 2021). Asimismo, investigaciones sobre sesgo en bases de datos dermatoscópicas han demostrado que los modelos pueden explotar correlatos no clínicos presentes en las imágenes, lo que subraya la necesidad de implementar auditorías rigurosas, pruebas de estrés y mecanismos de transparencia antes de su adopción en la práctica clínica (Bissoto et al., 2019).

Por otra parte, los estudios prospectivos destacan que la IA no opera en un entorno aislado. Su rendimiento en condiciones reales depende de múltiples factores, entre ellos la calidad en la captura de imágenes, el contexto clínico, la prevalencia de la enfermedad y la forma en que los profesionales interpretan las recomendaciones del sistema. En este sentido, la interacción humano–IA ha demostrado mejorar la exactitud diagnóstica de profesionales no especialistas y reducir la variabilidad interobservador; sin embargo, su implementación efectiva requiere capacitación adecuada y el desarrollo de marcos regulatorios que definan responsabilidades y aseguren la trazabilidad de las decisiones clínicas (Heinlein et al., 2024; Papachristou et al., 2024; Krakowski et al., 2024).

En síntesis, el debate ya no se centra en si la IA es capaz de detectar melanoma, sino en las condiciones bajo las cuales su uso resulta seguro, equitativo y generalizable. La consolidación de su aplicación clínica exige avanzar en la diversificación de los conjuntos de datos, la validación externa y prospectiva, la incorporación de enfoques de inteligencia artificial explicable, la calibración de los modelos según la prevalencia real y la implementación de estrategias de vigilancia continua tras su despliegue.

CONCLUSIONES

1. La inteligencia artificial aplicada al diagnóstico de melanoma presenta una alta exactitud, con valores de AUC cercanos a 0,96 y niveles elevados de sensibilidad y especificidad reportados en metaanálisis contemporáneos.
2. En entornos controlados, su rendimiento resulta comparable al de dermatólogos expertos y contribuye a mejorar el desempeño de clínicos no especialistas cuando se emplea como herramienta de apoyo diagnóstico.
3. Persisten sesgos relevantes asociados a la falta de diversidad en los fototipos cutáneos, el desequilibrio de clases y la presencia de atajos algorítmicos; estos factores afectan negativamente el desempeño en piel oscura y en lesiones atípicas.
4. Una adecuada generalización del modelo se vincula con el uso de datos multicéntricos, la validación externa y el desarrollo de estudios prospectivos, siendo la multimodalidad un enfoque prometedor para mejorar la robustez de los sistemas.
5. La implementación clínica de la inteligencia artificial debe realizarse de forma progresiva, incorporando auditorías de equidad, transparencia en los modelos y mecanismos de vigilancia continua que garanticen su uso seguro y efectivo.

REFERENCIAS BIBLIOGRÁFICAS

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Salinas, M. P., Sepúlveda, J., Hidalgo, L., Peirano, D., Morel, M., Uribe, P., et al. (2024). Artificial

intelligence versus clinicians for skin cancer diagnosis: Systematic review and meta-analysis. *npj Digital Medicine, 7*, 103. https://doi.org/10.1038/s41746-024-01103-x

Ertürk Zararsız, G. E., Yerlitaş Taştan, S. I., Çelik Gürbulak, E., Erakcaoğlu, A., Yılmaz Işıkhan, S., Demirbaş, A., et al. (2025). Diagnosis of melanoma with artificial intelligence systems: A meta-analysis and systematic review. *Journal of the European Academy of Dermatology and Venereology, 39*(11), 1912–1922. https://doi.org/10.1111/jdv.20781

Ye, Z., et al. (2024). Deep learning algorithms for melanoma detection using dermoscopic images: Systematic review and meta-analysis. *Artificial Intelligence in Medicine, 148*, 102934. https://doi.org/10.1016/j.artmed.2024.102934

Heinlein, L., Maron, R. C., Hekler, A., Haggenmüller, S., Wies, C., Utikal, J. S., et al. (2024). Prospective multicenter study using artificial intelligence to improve dermoscopic melanoma diagnosis in patient care. *Communications Medicine, 4*, 177. https://doi.org/10.1038/s43856-024-00598-5

Marchetti, M. A., et al. (2023). Prospective validation of dermoscopy-based open-source artificial intelligence for melanoma diagnosis (PROVE-AI study). *npj Digital Medicine, 6*, 152. https://doi.org/10.1038/s41746-023-00872-1

Papachristou, P., Söderholm, M., et al. (2024). Evaluation of an artificial intelligence-based decision support system for detection of cutaneous melanoma in primary care: A prospective real-life clinical trial. *British Journal of Dermatology, 191*(1), 125–134. https://doi.org/10.1093/bjd/ljae021

Krakowski, I., et al. (2024). Human–AI interaction in skin cancer diagnosis: Systematic review and meta-analysis. *npj Digital Medicine, 7*, 92. https://doi.org/10.1038/s41746-024-01031-w

Naseri, H., et al. (2024). Diagnosis and prognosis of melanoma from dermoscopy images using machine learning and deep learning: A systematic review. *BMC Cancer, 24*, 13423. https://doi.org/10.1186/s12885-024-13423-y

Patel, R. H., et al. (2023). Analysis of artificial intelligence-based approaches in melanoma

classification. *Cancers, 15*(19), 4694. https://doi.org/10.3390/cancers15194694

Yan, S., et al. (2025). A multimodal vision foundation model for clinical dermatology. *Nature Medicine*. https://doi.org/10.1038/s41591-025-03747-y

Alipour, N., et al. (2024). Skin type diversity in skin lesion datasets: A review. *Journal of Medical Systems, 48*. https://doi.org/10.1007/s13671-024-00440-0

Ruga, T. (2025). Bias in dermatological datasets: Underrepresentation of dark skin tones in melanoma classification images. In *Proceedings of ICCS 2025*. https://doi.org/10.1007/978-3-031-97554-7_32

Bissoto, A., Fornaciali, M., Valle, E., & Avila, S. (2019). (De)constructing bias on skin lesion datasets. *arXiv preprint*.

Daneshjou, R., Vodrahalli, K., Novoa, R. A., et al. (2021). Disparities in dermatology AI performance across skin tones. *Science Advances, 7*(50), eabj7821. https://doi.org/10.1126/sciadv.abj7821

Helenason, J., et al. (2024). Exploring the feasibility of an artificial intelligence-based clinical decision support system to detect cutaneous melanoma in primary care. *Acta Dermato-Venereologica*. https://doi.org/10.1080/02813432.2023.2283190

Kaur, R., et al. (2025). Advanced deep learning models for melanoma diagnosis and decision support: A comparative study. *Sensors, 25*(3), 594. https://doi.org/10.3390/s25030594

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115–118. https://doi.org/10.1038/nature21056

Haenssle, H. A., Fink, C., Toberer, F., Winkler, J., Stolz, W., Deinlein, T., et al. (2018). Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition. *Annals of Oncology, 29*(8), 1836–1842. https://doi.org/10.1093/annonc/mdy166

Tschandl, P., et al. (2019). Human–computer collaboration for skin cancer recognition. *The Lancet Oncology, 20*(7), 938–947. https://doi.org/10.1016/S1470-2045(19)30433-X