Comparative and complementary diagnostic value of dermatoscopy and clinical close-up photography in skin cancer diagnosis: A study from the MILK10k dataset

Christoph Müller, MD,[a] Anna Wolber, MD,[a] Aimilios Lallas, MD,[b] Giuseppe Argenziano, MD,[c] Iris Zalaudek, MD,[d] H. Peter Soyer, MD,[e,f] Susana Puig, MD,[g] Bengu Nisa Akay, MD,[h] Ana Maria Forsea, MD,[i,j] Josep Malvehy, MD,[k] Ashfaq Marghoob, MD,[l] Zoe Apalla, MD,[m] Cliff Rosendahl, MD,[n] Konstantinos Liopyris, MD,[o] Ofer Reiter, MD,[p] Caterina Longo, MD,[q] Philipp Tschandl, MD,[a,r] and Harald Kittler, MD[a,r]

Background: While dermatoscopy's added value to clinical imaging is known, the standalone value of dermatoscopy and clinical images remains unclear.

Objective: To quantify the standalone and complementary value of dermatoscopy and clinical close-up imaging across lesion types and reader expertise.

Methods: In an online reader study, 283 participants diagnosed 1567 paired images (clinical close-up/dermatoscopic) of skin cancers and mimics. The presentation order was randomized. Diagnostic accuracy, sensitivity, and specificity were compared. The effects of modality, reader expertise, and presentation order were analyzed using a generalized linear mixed model.

Results: Dermatoscopy alone showed higher sensitivity (85.0% vs 74.2%) but lower specificity (66.8% vs 71.9%) compared with clinical close-up images alone. It improved accuracy for melanoma and basal cell carcinoma, but not for nevi. Adding either modality increased the odds of a correct diagnosis.

From the Department of Dermatology, Medical University of Vienna, Vienna, Austria[a]; First Department of Dermatology, School of Medicine, Faculty of Health Sciences, Aristotle University, Thessaloniki, Greece[b]; Dermatology Unit, University of Campania "Luigi Vanvitelli", Naples, Italy[c]; Department of Dermatology, Medical University of Trieste, Trieste, Italy[d]; Frazer Institute, The University of Queensland, Dermatology Research Centre, Brisbane, Queensland, Australia[e]; Department of Dermatology, Princess Alexandra Hospital, Brisbane, Australia[f]; Melanoma Unit, Dermatology Department, Hospital Clínic de Barcelona, IDIBAPS, Universitat de Barcelona, Barcelona, Spain[g]; Department of Dermatology, Faculty of Medicine, Ankara University, Ankara, Turkey[h]; Department of Oncologic Dermatology, Carol Davila University of Medicine and Pharmacy Bucharest, Elias University Hospital Bucharest, Bucharest, Romania[i]; Health and Medical University Potsdam, Potsdam, Germany[j]; Dermatology Department, Hospital Clínic de Barcelona, University of Barcelona, IDIBAPS, CIBER de enfermedades raras, Instituto Carlos III, Barcelona, Spain[k]; Memorial Sloan Kettering Cancer Center, Hauppauge, New York[l]; Second Department of Dermatology, Aristotle University, Thessaloniki, Greece[m]; Mayne Academy of General Practice, Medical School, The University of Queensland, Brisbane, Queensland, Australia[n]; Department of Dermatology, Andreas Syngros Hospital of Venereal and Dermatological Diseases, University of Athens, Athens, Greece[o]; Dermatology Division, Rabin Medical Center, Petach Tikva, and Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel[p]; Department of Dermatology, University of Modena and Reggio Emilia, Modena, Italy[q]; and Comprehensive Center for Artificial Intelligence in Medicine, Medical University of Vienna, Vienna, Austria.[r]

Funding sources: None.

Patient consent: As the study involved no direct interventions with patients and only anonymized image data were used, patient consent was not required.

IRB approval status: Reviewed and approved by the Ethics Committee of the Medical University of Vienna, Vienna, Austria, approval number 1070/2023.

Accepted for publication March 23, 2026.

Correspondence to: Christoph Müller, MD, Department of Dermatology, Medical University of Vienna, Währinger Gürtel 18-20, Vienna, 1090, Austria. E-mail: christoph.mueller@meduniwien.ac.at.

Published online March 26, 2026.

0190-9622

https://doi.org/10.1016/j.jaad.2026.03.069

Dermatoscopy raised the odds by 52% (OR = 1.52; 95% CI: 1.36-1.70; $P < .001$), while clinical close-ups increased the odds by 40% (OR = 1.40; 95% CI: 1.20-1.61; $P < .001$).

Limitations: Image-based evaluation simulated teledermatology rather than face-to-face assessment; potential selection and verification bias cannot be excluded.

Conclusion: Dermatoscopy alone yields higher sensitivity for malignant lesions but lower specificity for nevi than clinical close-up images. The latter provides complementary diagnostic cues, underscoring the value of integrating both modalities for optimal assessment. (J Am Acad Dermatol 2026;95:70-7.)

Key words: dermatoscopy; diagnostic accuracy; diagnostic modality; sensitivity; skin cancer; specificity.

INTRODUCTION

Dermatoscopy is an integral component of modern dermatology. As a noninvasive imaging technique that reveals subsurface skin structures, it has markedly improved the detection and management of skin cancer.[1-6] Its importance is emphasized in clinical guidelines such as the German S3 guideline,[7-10] which recommend its routine use and structured training for clinicians. The diagnostic benefit of dermatoscopy is well established, with numerous studies showing that it significantly increases both sensitivity and specificity compared to unaided visual inspection.[3,11]

CAPSULE SUMMARY

- This study confirms that dermatoscopy enhances sensitivity for skin cancer detection when added to clinical close-up, while clarifying its trade-off of reduced specificity for nevi.
- Dermatoscopic and clinical close-up provide complementary diagnostic insights. Each enhances information available from the other, highlighting the importance of integrating both views in diagnostic evaluation.

However, much of the existing evidence is limited to selected lesion types, mainly melanocytic nevi and melanoma, and to highly experienced experts in academic settings, thus not fully representing the broader community of practitioners. Most prior studies have examined only the complementary diagnostic value of dermatoscopy by assessing performance gains when dermatoscopic images are added to clinical close-ups.[12] This approach captures the incremental benefit of dermatoscopy but overlooks 2 important aspects: the comparative (standalone) diagnostic performance of each modality and the potential added value of clinical close-up images themselves. These distinctions are particularly relevant in teledermatology, where the order and combination of image modalities may vary.

While the diagnostic benefits of dermatoscopy are well-documented, concerns have been raised regarding its potential influence on biopsy rates. Specifically, the increased sensitivity of dermatoscopy may come at the cost of reduced specificity in certain contexts, potentially leading to a higher number of false positives.[13]

This trade-off underscores the need to evaluate the performance of dermatoscopy in more diverse settings and across varying levels of expertise. The aim of this study is to quantify both the comparative (standalone) and complementary (combined) diagnostic value of dermatoscopy and clinical close-up imaging, using a large dataset encompassing multiple lesion types and a diverse international reader base with different expertise.

MATERIAL AND METHODS

Cases

For this study, 1567 clinical (close-up) and dermatoscopic images of solitary skin lesions, including actinic keratosis/intraepidermal carcinoma, basal cell carcinoma, benign keratinocytic lesions (seborrheic keratoses, solar lentigines), dermatofibroma, melanoma, melanocytic nevus, squamous cell carcinoma/keratoacanthoma and vascular lesions, were used. The images used in this study are a selected subset from the MILK10k dataset (all 10,480 images available at the ISIC Archive: https://challenge.isic-archive.com/data/#milk10k).[14] Diagnostic ground truth was based on histopathology ($n = 1428$) or, when unavailable, expert consensus or long-term follow-up ($n = 139$). Numbers of all lesions used in this study and examples of nevi with differences in performance

Abbreviations used:

IDS: International Dermoscopy Society
OR: odds ratio

are provided in the Supplementary Material available via Mendeley at https://data.mendeley.com/datasets/bxy9272429/1.

Online reader study

This study, including the recruitment of raters, was conducted under the framework of an International Dermoscopy Society (IDS) project. Data collection was based on voluntary participation in a quiz hosted on the online platform Dermachallenge from July 21, 2024, to June 13, 2025 (https://dermonaut.meduniwien.ac.at/dermachallenge).

Each quiz round consisted of a random selection of 10 lesions, with block randomization ensuring representation of all diagnostic categories. For each lesion, participants were shown both clinical close-up and dermatoscopic images in a randomized order, that is, either the clinical or the dermatoscopic image could appear first. After viewing the first image, participants were required to make an initial diagnosis. Upon presentation of the second image, they were prompted to revise or confirm their diagnosis before final submission.

Readers could complete multiple quiz rounds at their discretion, with each round containing a different random set of lesions. Prior to participation, each registered user completed a series of standardized, independent assessments consisting of 30 test cases, with a simple scoring system awarding 1 point for each correct diagnosis. Based on this standardized test, participants were classified into low-, medium-, and high-expertise groups according to the tertiles of the score distribution.

Ethical considerations

This study was performed in accordance with the Declaration of Helsinki and received approval from the Ethics Committee of the Medical University of Vienna (approval number 1070/2023). As the study involved no direct interventions with patients and only de-identified image data were used, patient consent was not required. Participants in the online questionnaire were asked to provide their consent online for the anonymous use of the collected data for statistical analysis.

Statistics

All analyses were conducted in R (version 4.5.0). Incomplete quiz rounds were excluded. To determine the raw diagnostic accuracy of each modality (clinical or dermatoscopic) independently, only the responses to the first image of each pair were considered; that is, diagnoses made without access to the corresponding image of the other modality. To evaluate the added diagnostic value of each modality, we calculated the change in accuracy after participants viewed the second image and could revise their initial diagnosis.

Categorical variables are given as absolute frequencies and frequencies in percentages. Professions were grouped as dermatology residents, dermatology specialists, general practitioners and 'others' (see Table I). As described above each round contained 20 questions, presented as 10 pairs (clinical close-up and dermatoscopy). For each pair, we recorded response transitions (eg, wrong to correct, correct to wrong). Differences between orders were visualized with bar plots. Accuracy was defined as the proportion of correct answers. Comparisons between image modalities (clinical close-up and dermatoscopy) were tested using two-sample proportion tests, overall and within subgroups (diagnosis, modality, expertise).

To test whether adding a second diagnostic modality increased accuracy, a McNemar test was performed. To reduce bias, only the first assessment for each image was included in the analysis. Mixed-effects logistic regression (modelling correctness as binary) was used to account for clustering by participant and image. Extended models included

Table I. Descriptive data of participants

Characteristic	Category	Count (N = 283)	Percentage (%)
Gender	Female	180	63.6
	Male	100	35.3
	Other/Not specified	3	1.1
Country of origin (top 5)	Austria	37	13.1
	Italy	21	7.4
	Australia	20	7.1
	Germany	20	7.1
	United Kingdom	18	6.4
	Other (40 countries)	167	59.0
Profession	Dermatology Resident	121	42.8
	Board-Certified Dermatologist	73	25.8
	General Practitioner	53	18.7
	Other*	36	12.7
Expertise level	Low	117	41.3
	Medium	72	25.4
	High	94	33.2

*Non-dermatological residents, physician assistants, medical specialists, nurse practitioners, medical students and non-medical persons.

expertise-level, sequence, and interaction terms.[15] To test for order effects, final accuracy on the second image was compared between "close-up first" and "dermatoscopy first" rounds using chi-squared and proportion tests, with stratification by expertise levels. All tests were two-sided ($P < .05$).

RESULTS

Participants

A total of 283 participants from 45 countries completed at least 1 quiz round. The majority of participants were dermatology residents (n = 121), followed by dermatology specialists (n = 73), general practitioners (n = 53), and "others" (n = 36); as shown in Table I.

Comparative diagnostic accuracy of image modalities

Dermatoscopy alone had a significantly higher sensitivity compared to clinical close-up image alone (85.0% vs 74.2%; OR = 1.97; 95% CI: 1.69-2.30; $P < .001$), albeit at the expense of specificity (66.8% vs 71.9%; OR = 0.79; 95% CI: 0.70-0.89; $P < .001$). Subgroup analyses demonstrated a consistent diagnostic advantage of dermatoscopy across most lesion types, including basal cell carcinoma (OR = 1.28; 95% CI: 1.04-1.5; $P = .015$), melanoma (OR = 1.84; 95% CI: 1.53-2.21; $P < .001$), actinic keratosis/intraepidermal carcinoma (OR = 1.96; 95% CI: 1.20-3.21; $P = .005$), dermatofibroma (OR = 2.16; 95% CI: 1.09-4.32; $P = .025$), and benign keratinocytic lesion (OR = 1.60; 95% CI: 1.22-2.11; $P < .001$). In contrast, nevi were more frequently classified correctly using clinical close-up images alone (OR = 0.62; 95% CI: 0.54-0.71; $P < .001$). No significant difference was observed for squamous cell carcinoma/keratoacanthoma (OR = 1.14; 95% CI: 0.59-2.21; $P = .754$). Detailed results are provided in Supplementary Material available via Mendeley at https://data.mendeley.com/datasets/bxy9272429/1.

In the subgroup analysis by expertise level, no significant difference in diagnostic performance between dermatoscopy and clinical close-up images was observed among participants with low or medium expertise levels (OR = 0.98; 95% CI: 0.84-1.14; $P = .771$ and OR = 0.91; 95% CI: 0.78-1.06; $P = .225$, respectively). Among participants with high dermatoscopic expertise, however, diagnostic accuracy using dermatoscopic images alone was significantly higher than with clinical close-up images alone (OR = 0.87; 95% CI: 0.76-0.99; $P = .032$).

To assess the influence of experience on diagnostic performance, results were stratified by participants' expertise level and lesion type (Fig 1). For melanoma and basal cell carcinoma, diagnostic accuracy with dermatoscopy alone was consistently higher than with clinical close-up images alone across all expertise levels. In contrast, nevi were more accurately classified based on clinical close-up images alone in all groups. A clear trend of improved performance with increasing expertise was observed for both modalities.

Complementary diagnostic value of each modality

To evaluate whether adding a second imaging modality enhanced diagnostic performance, we compared accuracy before and after the other modality was provided. The inclusion of dermatoscopy led to a significant improvement in diagnostic accuracy by 5.7% (OR = 1.52; 95% CI: 1.36-1.70; $P < .001$), while adding a clinical close-up image increased accuracy by 2.6% (OR = 1.40; 95% CI: 1.20-1.61; $P < .001$).

A subgroup analysis by dermatoscopic expertise level showed that adding dermatoscopy significantly improved diagnostic accuracy across all groups: low expertise (OR = 1.25; 95% CI: 1.03-1.52; $P = .025$), medium expertise (OR = 1.67; 95% CI: 1.36-2.05; $P < .001$), and high expertise (OR = 1.69; 95% CI: 1.40-2.04; $P < .001$). In contrast, adding the clinical image yielded a significant benefit only among participants with low (OR = 1.42; 95% CI: 1.13-1.80; $P = .003$) and medium dermatoscopy expertise levels (OR = 1.63; 95% CI: 1.23-2.15; $P < .001$), whereas no significant improvement was observed for high expertise participants (OR = 1.21, 95% CI: 0.94-1.54, $P = .151$). Detailed results are presented in Table II.

Analysis by lesion type demonstrated that adding dermatoscopy provided a highly significant net benefit for diagnosing malignant lesions, including melanoma (OR = 3.54; 95% CI: 2.68-4.67; $P < .001$), basal cell carcinoma (OR = 2.35; 95% CI: 1.74-3.17; $P < .001$), and actinic keratosis/intraepidermal carcinoma (OR = 3.76; 95% CI: 1.90-7.43; $P < .001$). In contrast, adding the clinical image did not yield a significant improvement for these neoplasms (see Table III).

Order effects

To examine the dynamics of diagnostic revisions, we analyzed how participant responses changed after viewing the second image in each pair (Supplementary Material available via Mendeley at https://data.mendeley.com/datasets/bxy9272429/1). In both presentation orders, the addition of the second image prompted a substantial number of beneficial corrections, where an initially incorrect diagnosis was changed to a correct one. In contrast,

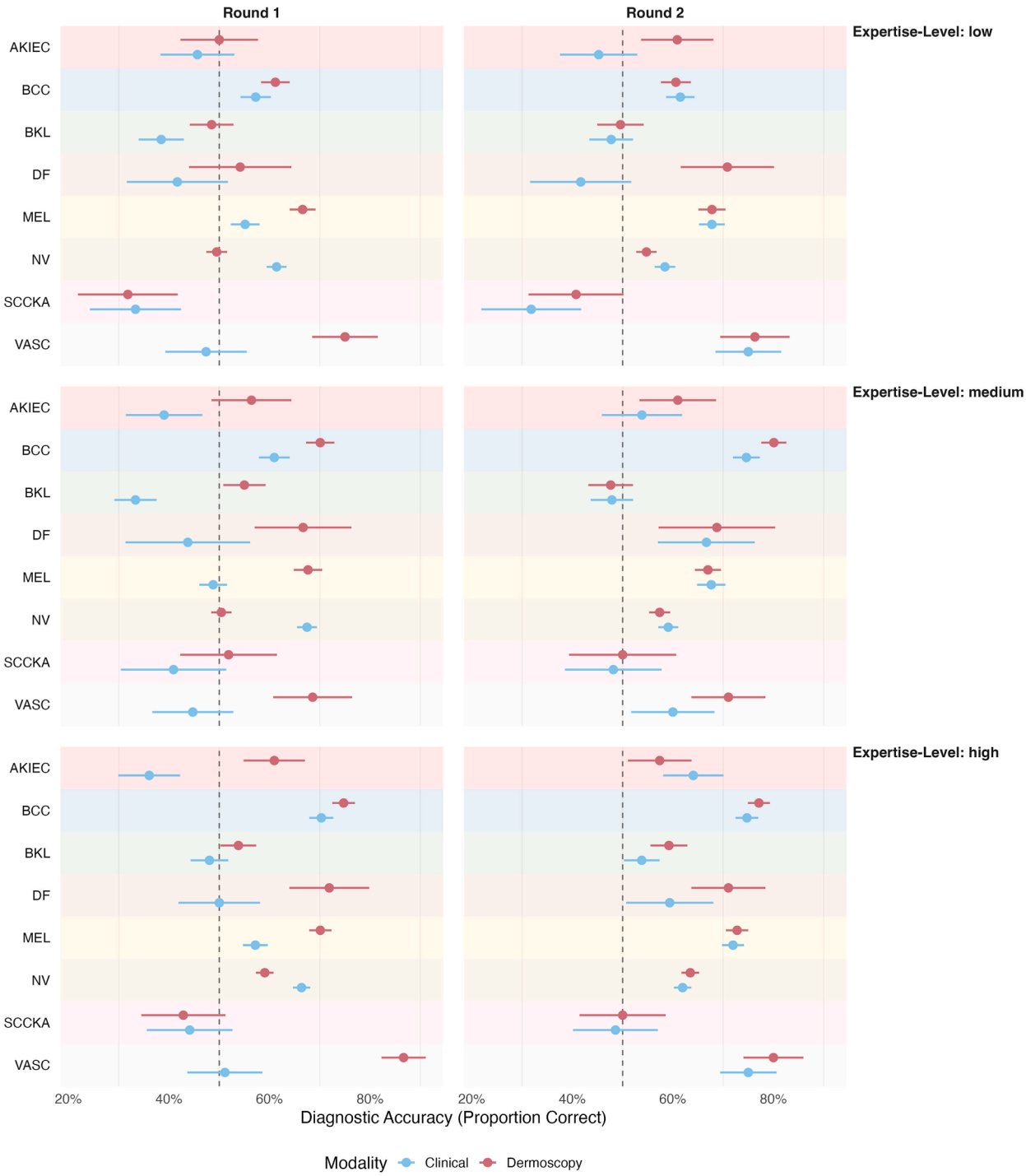

Fig 1. Diagnostic accuracy by modality, lesion type, and expertise. The plot displays the diagnostic accuracy based on the evaluation of the first image only (round 1) and after addition of the second image (round 2). The results are faceted by expertise level (low, medium, high). Within each facet, the accuracy for the clinical view alone (Clinical, light blue) is compared to the accuracy when using dermatoscopy alone (Dermatoscopy, red). Points represent the mean accuracy; error bars indicate the 95% confidence intervals. *AKIEC*, Actinic keratosis/intraepidermal carcinoma; *BCC*, basal cell carcinoma; *BKL*, benign keratinocytic lesions; *DF*, dermatofibroma; *MEL*, melanoma; *NV*, nevus; *SCCKA*, squamous cell carcinoma/keratoacanthoma; *VASC*, vascular lesion.

Table II. Subgroup analysis expertise level for accuracy improvement from combined modalities

Expertise level	Sequence	OR	95% CI	*P*-value
Low	Close-up first, dermatoscopy added	1.25	1.03-1.52	**.0250**
Low	Dermatoscopy first, close-up added	1.42	1.13-1.80	**.0034**
Medium	Close-up first, dermatoscopy added	1.67	1.36-2.05	**<.001**
Medium	Dermatoscopy first, close-up added	1.63	1.23-2.15	**<.001**
High	Close-up first, dermatoscopy added	1.69	1.40-2.04	**<.001**
High	Dermatoscopy first, close-up added	1.21	0.94-1.54	.1506

The odds ratio (OR) show total benefit after adding the second imaging modality. Bolded *P* values indicate statistical significance ($P < .05$).

Table III. Subgroup analysis diagnostic categories for accuracy improvement from combined modalities

Diagnosis	Sequence	OR	95% CI	*P*-value
Benign keratinocytic lesions	Close-up first, dermatoscopy added	2.36	1.61-3.46	**<.001**
Benign keratinocytic lesions	Dermatoscopy first, close-up added	0.71	0.44-1.17	.2148
Melanoma	Close-up first, dermatoscopy added	3.54	2.68-4.67	**<.001**
Melanoma	Dermatoscopy first, close-up added	1.17	0.85-1.60	.3754
Nevus	Close-up first, dermatoscopy added	0.63	0.53-0.76	**<.001**
Nevus	Dermatoscopy first, close-up added	2.29	1.81-2.89	**<.001**
Basal cell carcinoma	Close-up first, dermatoscopy added	2.35	1.74-3.17	**<.001**
Basal cell carcinoma	Dermatoscopy first, close-up added	1.21	0.86-1.69	.3053
Actinic Keratosis/ Intraepidermal Carcinoma	Close-up first, dermatoscopy added	3.76	1.90-7.43	**<.001**
Actinic Keratosis/ Intraepidermal Carcinoma	Dermatoscopy first, close-up added	0.93	0.44-1.95	1.0000
Dermatofibroma	Close-up first, dermatoscopy added	39.00	2.35-645.92	**<.001**
Dermatofibroma	Dermatoscopy first, close-up added	0.33	0.10-1.12	.0961
Vascular Lesion	Close-up first, dermatoscopy added	5.00	2.39-10.44	**<.001**
Vascular Lesion	Dermatoscopy first, close-up added	0.35	0.13-0.94	**.0442**
Squamous Cell Carcinoma/ Keratoacanthoma	Close-up first, dermatoscopy added	2.09	0.76-5.78	.2113
Squamous Cell Carcinoma/ Keratoacanthoma	Dermatoscopy first, close-up added	1.15	0.40-3.30	1.0000

The odds ratio (OR) show total benefit after adding the second imaging modality. Bolded *P* values indicate statistical significance ($P < .05$).

only a small proportion of responses shifted from correct to incorrect. The overall pattern of these response changes was highly consistent across both sequences, irrespective of whether the clinical close-up or the dermatoscopic image was presented first.

Generalized linear mixed model

A generalized linear mixed model was applied to account for potential learning effects and to ensure that frequent participants did not disproportionately influence the overall results. Compared with the low-expertise group, participants with high expertise ($\beta = 0.43$; OR = 1.54; 95% CI: 1.30-1.83; $P < .001$) and intermediate expertise ($\beta = 0.19$; OR = 1.21; 95% CI: 1.00-1.45; $P = .050$) were more likely to provide correct answers. A significant main effect of case sequence was observed, indicating a learning effect: the odds of a correct response increased with each additional image pair ($\beta = 0.27$; OR = 1.31; 95% CI: 1.18-1.45; $P < .001$). In contrast, there was no significant main effect of modality ($\beta = 0.15$; OR = 1.16; 95% CI: 0.90-1.50; $P = .241$) and no significant interaction between modality and its order of presentation ($\beta = -0.03$; OR = 0.97; 95% CI: 0.84-1.12; $P = .671$), suggesting that diagnostic performance was independent of whether the clinical or dermatoscopic image was shown first.

DISCUSSION

It is well established that dermatoscopy enhances diagnostic accuracy for pigmented skin lesions compared with clinical examination alone,[16] and that this improvement depends strongly on the examiner's level of expertise.[17] However, because dermatoscopy is typically performed after naked eye examination, the potential influence of the sequence

of modalities has not been systematically investigated. In telemedicine, where image order can vary, it is important to clarify whether the observed benefit reflects intrinsic superiority or the complementary value of combining 2 perspectives.[18]

In our study, when each modality was assessed in isolation, viewing a dermatoscopic image alone did not guarantee higher overall diagnostic accuracy compared with viewing a clinical image alone. Thus, there was no general superiority of dermatoscopy over clinical imaging when both were evaluated independently. A closer examination, however, reveals a more nuanced interpretation. When used as a standalone examination, dermatoscopy showed significantly higher accuracy for most malignancies, particularly melanoma and basal cell carcinoma. This advantage came with a trade-off: lower accuracy for nevi compared with clinical close-up images. In other words, dermatoscopy alone provided higher sensitivity, whereas clinical imaging offered higher specificity for nevi. Since melanocytic nevi constitute the majority of benign lesions in clinical practice, this trade-off is clinically relevant and should not be overlooked. Improving specificity, especially for common lesions like nevi, requires integrating morphological assessment with biological context, particularly through comparative assessment and monitoring for change over time, a process inherently more challenging in a static image-based teledermatology setting than in a face-to-face consultation.[19]

Less experienced readers were particularly affected by this loss of specificity when evaluating nevi with dermatoscopy.[20] Nevi may display subtle dermatoscopic features that mimic melanoma, which less experienced users may overinterpret as malignant.[21] Non-experts may also adopt lower diagnostic thresholds out of caution, whereas experts apply more calibrated judgments based on pattern recognition and experience.[17,22]

These findings support the concept that the diagnostic value of dermatoscopy is inseparable from user expertise. As an interpretive technique, it requires structured training to recognize patterns accurately. Misinterpretation by untrained users may increase false positives, a point rightly noted by critics, but this reflects insufficient expertise rather than a limitation of the method itself. Dermatoscopy achieves its full diagnostic potential only when applied by trained and experienced clinicians.

In clinical practice, the standalone accuracy of each modality is of limited relevance, as diagnoses typically rely on the combined assessment of clinical and dermatoscopic images. What matters is the complementary value each modality contributes. In this regard, we confirmed that dermatoscopy provides substantial additional diagnostic benefit, particularly by improving sensitivity for melanoma/skin cancer, and that this effect was consistent across all expertise levels. Adding a clinical image after dermatoscopic assessment yielded further, though smaller, diagnostic gains, especially among less experienced readers who benefited from the contextual information provided by the clinical image.

Although dermatoscopy and clinical close-up imaging achieved comparable accuracy when used alone, the gain in accuracy was greater when dermatoscopy followed clinical imaging than when the sequence was reversed, even though the final combined accuracy was independent of presentation order. This seemingly paradoxical finding can be explained by the informational asymmetry between the 2 modalities. Dermatoscopy reveals subsurface structures and subtle diagnostic clues that are highly informative for detecting malignancies such as melanoma but may mislead less experienced readers when evaluating nevi. In contrast, clinical close-up images provide fewer potentially confusing details and are therefore less prone to overinterpretation, reinforcing their complementary rather than corrective role in the diagnostic process.

CONCLUSION

In conclusion, our study shows that diagnostic improvement arises not from the superiority of a single modality but from the synergistic integration of both. The process of diagnostic revision, enabled by incorporating information from the second modality, leads to a significant net gain in accuracy. Optimal performance is achieved through the combined use of 2 complementary visual perspectives, clinical close-up and dermatoscopic imaging, which together enhance diagnostic reasoning regardless of viewing order. The complementary diagnostic gain provided by dermatoscopy is strongly dependent on user expertise and lesion type, with a corresponding increase in false positives observed among less experienced readers, particularly for nevi.

LIMITATIONS

Potential selection and verification bias cannot be excluded and might explain the observed lower specificity for nevi; however, this loss of specificity did not occur with other benign lesions (eg, vascular lesions or dermatofibromas). To mitigate bias, many nevi verified as unequivocally benign via long-term follow-up or expert consensus were also included. Furthermore, the study used a controlled quiz environment (simulating teledermatology) and thus does not capture the complexity of real-world

clinical practice, where contextual cues may further inform the evaluation. Additionally, the study design involved paired image assessment. Although participants were instructed to diagnose based on the currently visible image, the knowledge that a second complementary image would follow may have influenced their initial diagnostic thresholds compared to a strictly isolated single-modality assessment. Finally, there is the possibility that some diagnostic classes, especially those with a low number of cases may not be fully representative.

Declaration of generative AI and AI-Assisted technologies in the writing process

During the preparation of this work the authors used Google Gemini for proofreading and to improve readability of the final manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Conflicts of interest

None disclosed.

REFERENCES

1. Kittler H, Marghoob AA, Argenziano G, et al. Standardization of terminology in dermoscopy/dermatoscopy: results of the third consensus conference of the International Society of Dermoscopy. *J Am Acad Dermatol*. 2016;74(6):1093-1106.
2. Kittler H. Evolution of the clinical, dermoscopic and pathologic diagnosis of melanoma. *Dermatol Pract Concept*. 2021;11(Suppl 1):e2021163S.
3. Vestergaard ME, Macaskill P, Holt PE, Menzies SW. Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. *Br J Dermatol*. 2008;159(3):669-676.
4. Bafounta ML, Beauchet A, Aegerter P, Saiag P. Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests. *Arch Dermatol*. 2001;137(10):1343-1350.
5. Dinnes J, Deeks JJ, Chuchu N, et al. Dermoscopy, with and without visual inspection, for diagnosing melanoma in adults. *Cochrane Database Syst Rev*. 2018;12(12):CD011902.
6. Dinnes J, Deeks JJ, Chuchu N, et al. Visual inspection and dermoscopy, alone or in combination, for diagnosing keratinocyte skin cancers in adults. *Cochrane Database Syst Rev*. 2018;12(12):CD011901.
7. Melanoma: assessment and management. NICE. Accessed November 6, 2025. https://www.nice.org.uk/guidance/ng14
8. Leitlinie zur Diagnostik, Therapie und Nachsorge des Melanoms. *J Dtsch Dermatol Ges*. 2020;18(10):1. https://doi.org/10.1111/ddg.14307_g
9. Amaral T, Ottaviano M, Arance A, et al. Cutaneous melanoma: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*. 2025;36(1):10-30.
10. Swetter SM, Tsao H, Bichakjian CK, et al. Guidelines of care for the management of primary cutaneous melanoma. *J Am Acad Dermatol*. 2019;80(1):208-250.
11. Sinz C, Tschandl P, Rosendahl C, et al. Accuracy of dermatoscopy for the diagnosis of nonpigmented cancers of the skin. *J Am Acad Dermatol*. 2017;77(6):1100-1109.
12. Ferrándiz L, Ojeda-Vila T, Corrales A, et al. Internet-based skin cancer screening using clinical images alone or in conjunction with dermoscopic images: a randomized teledermoscopy trial. *J Am Acad Dermatol*. 2017;76(4):676-682.
13. Welch HG, Mazer BL, Adamson AS. The rapid rise in cutaneous melanoma diagnoses. *N Engl J Med*. 2021;384(1):72-79.
14. Tschandl P, Akay BN, Rosendahl C, et al. MILK10k: a hierarchical multimodal imaging-learning toolkit for diagnosing pigmented and nonpigmented skin cancer and its simulators. *J Invest Dermatol*. 2026;146:357-364.e7. https://doi.org/10.1016/j.jid.2025.06.1594
15. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *arXiv [statCO]*. Published online June 23, 2014. http://arxiv.org/abs/1406.5823
16. Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. *Lancet Oncol*. 2002;3(3):159-165.
17. Binder M, Puespoeck-Schwarz M, Steiner A, et al. Epiluminescence microscopy of small pigmented skin lesions: short-term formal training improves the diagnostic performance of dermatologists. *J Am Acad Dermatol*. 1997;36(2):197-202.
18. Finnane A, Dallest K, Janda M, Soyer HP. Teledermatology for the diagnosis and management of skin cancer: a systematic review. *JAMA Dermatol*. 2017;153(3):319-327.
19. Kittler H, Guitera P, Riedl E, et al. Identification of clinically featureless incipient melanoma using sequential dermoscopy imaging. *Arch Dermatol*. 2006;142(9):1113-1119.
20. Secker LJ, Buis PAJ, Bergman W, Kukutsch NA. Effect of a dermoscopy training course on the accuracy of primary care physicians in diagnosing pigmented lesions. *Acta Derm Venereol*. 2017;97(2):263-265.
21. Carli P, Mannone F, De Giorgi V, Nardini P, Chiarugi A, Giannotti B. The problem of false-positive diagnosis in melanoma screening: the impact of dermoscopy. *Melanoma Res*. 2003;13(2):179-182.
22. Chuchu N, Dinnes J, Takwoingi Y, et al. Teledermatology for diagnosing skin cancer in adults. *Cochrane Database Syst Rev*. 2018;12(12):CD013193.