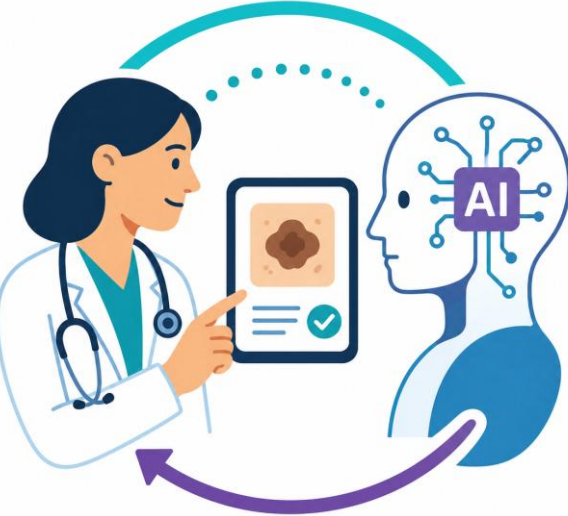


Diagnostic performance of human-artificial intelligence (AI) collaboration in teledermoscopy: a retrospective validation study of a novel 8-class AI tool

Carolina Nätterdahl¹, Isabelle Krakowski², Kari Nielsen^{1,3}, Teo Helkkula¹, Nikki Adler⁴, Johan Kappelin^{1,5}, Cristina Vico-Alonso⁴, Cecilia Sand¹, Karim Saleh¹, Albert Duvetorp^{6,7}, Katharine Hopkins^{1,7}, Kevin Yang¹, Hugh Roberts⁸, Sinja Kristiansen⁷, Alex Chamberlain⁴, Irina Baranovskaya⁷, Tine Vestergaard⁸, Magdalena Claesson⁹, Bertil Persson¹, Victoria Mar⁴, Jan Lapins², Åsa Ingvar¹

1. Department of Dermatology and Venereology, Skåne University Hospital, Lund, Sweden, Dermatology and Venereology, Lund University Skin Cancer Research Group (LUSCaR), Department of Clinical Sciences Lund, Lund University, Lund, Sweden, 2. Dermatology and Venereology Division, Department of Medicine Solna and Huddinge, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden, Stockholm, Sweden, 3. Department of Dermatology and Venereology, Helsingborg Hospital, Sweden, 4. School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia, Victorian Melanoma Service, Alfred Hospital, Melbourne, Australia, 5. Department of Dermatology, Landskrona Hospital, Landskrona, Sweden, 6. LEO Foundation Skin Immunology Research Center, Department of Immunology and Microbiology, University of Copenhagen, Denmark, 7. Department of Dermatology and Venereology, Skåne University Hospital, Malmö, Sweden, 8. Department of Dermatology and Allergy Centre, Odense University Hospital, Odense, Denmark, 9. Department of Dermatology and Venereology, Sahlgrenska University Hospital, Department of Dermatology and Venereology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

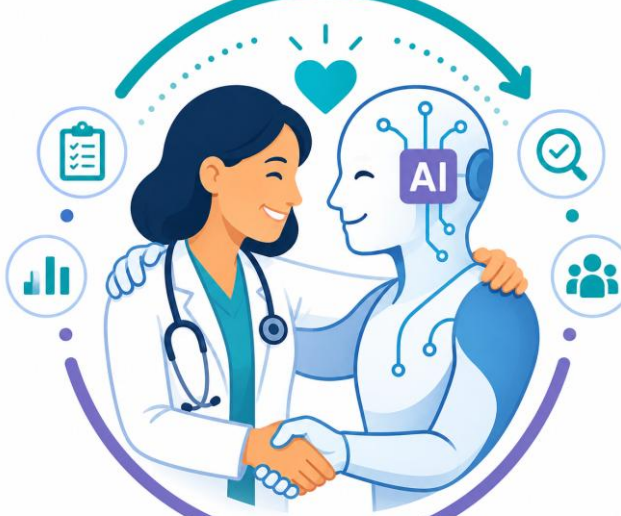
Objective



To evaluate whether **dermatologist–AI collaboration improves diagnostic accuracy** in skin lesion assessment using a novel 8-class AI model.

”This study aims to evaluate the most likely future use of AI in dermatology: AI as a support to clinicians”

Conclusion



Access to an 8-class AI tool added value to teledermoscopic assessments by dermatologists. Although overall gains in accuracy were modest, potentially reflecting the already high baseline performance of experienced assessors, **all significant changes represented improvements with no indications of performance deterioration.**

These findings highlight the potential of human-AI collaboration and support integrating AI assistance into teledermoscopy workflows. Prospective real-world studies are now needed to establish its clinical impact and guide safe implementation.

Patients and Methods

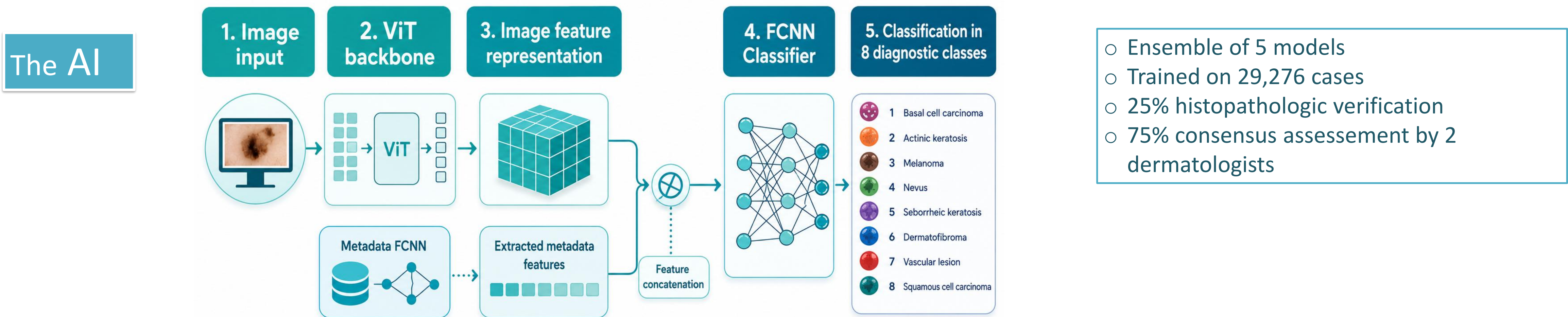
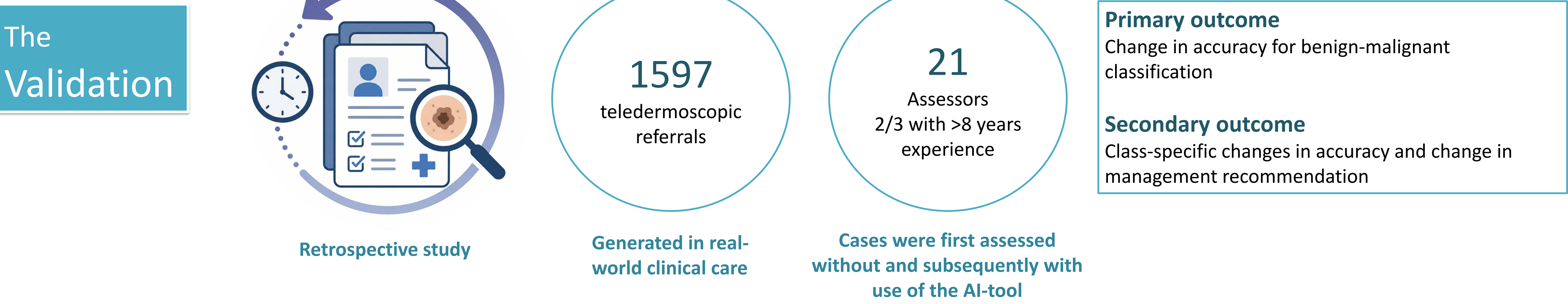


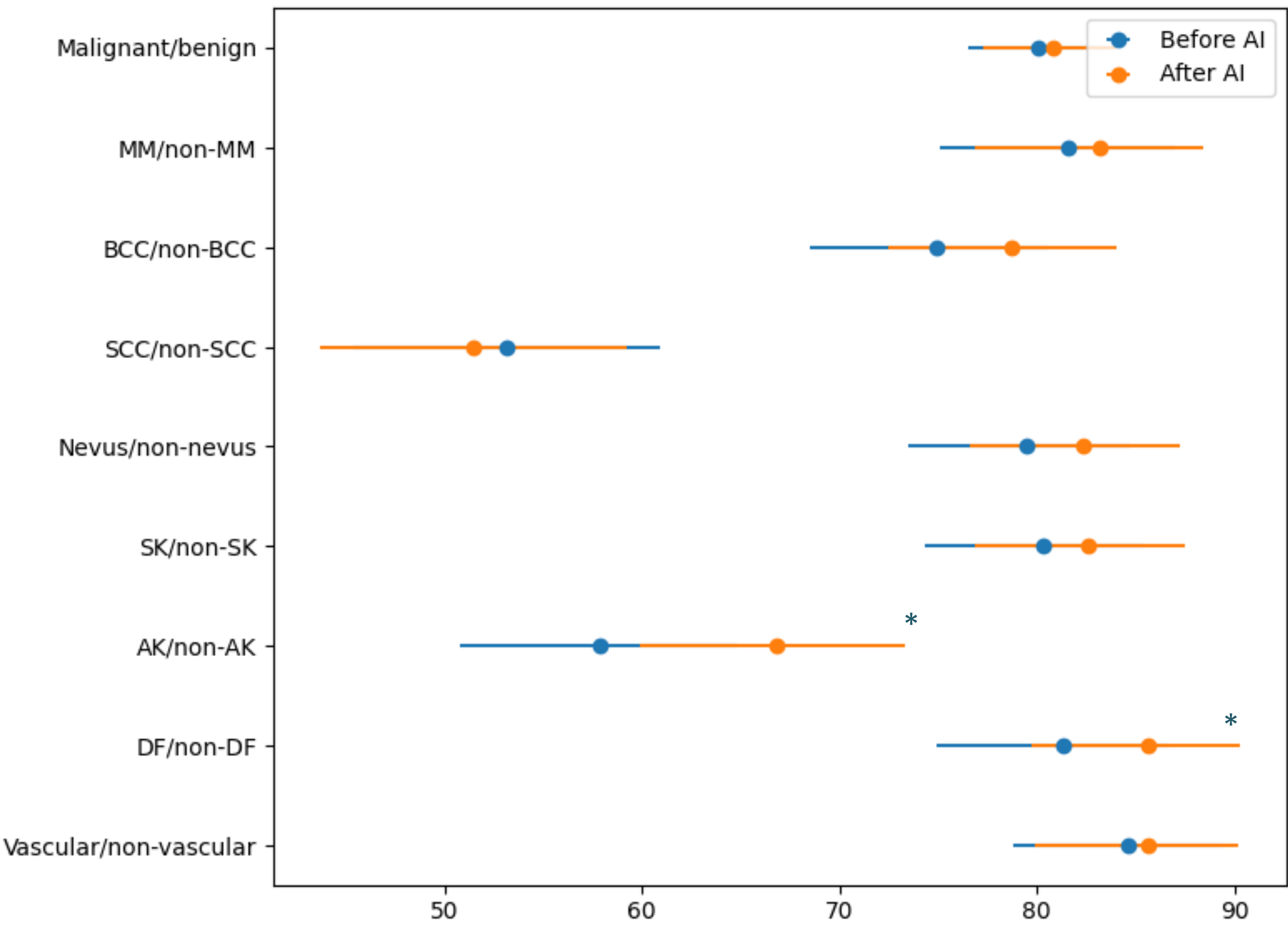
Figure 1. Illustration of the multimodal model architecture for skin lesion classification.



Results

Baseline sensitivity and specificity for benign-malignant classification were 80% and 90%, increasing to 81% (p = 0.43) and 91% (p = 0.05) after AI integration. Class-specific analyses showed significant improvements in specificity for nevi, seborrheic keratoses, and vascular lesions, and significant improvements in sensitivity for actinic keratoses and dermatofibromas.

Sensitivity



Specificity

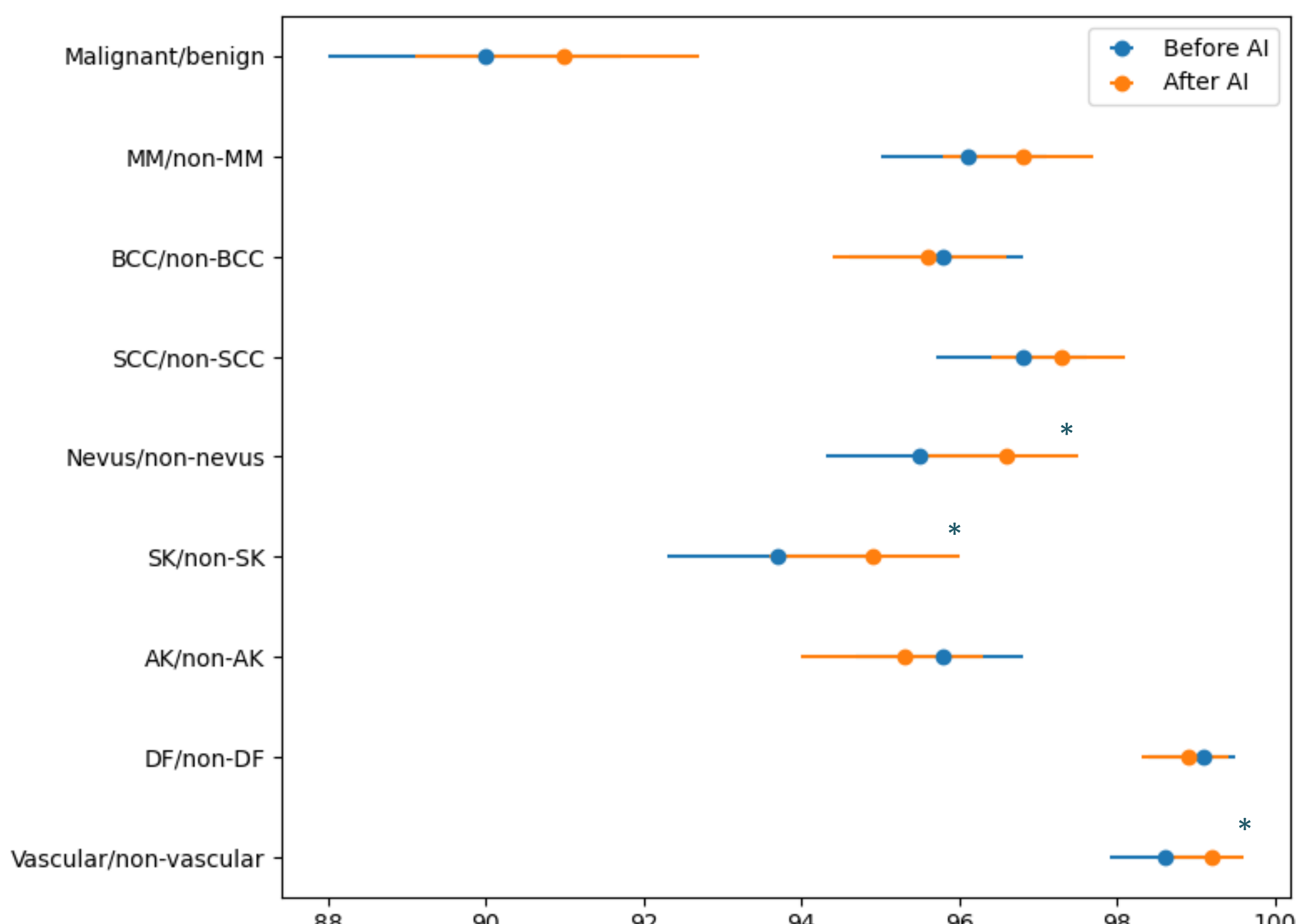


Figure 2. Sensitivity and specificity of skin lesion classification of dermatologists before (blue) and after (orange) use of 8-class diagnostic AI-tool. Dots are point estimates and lines represent 95% confidence intervals. MM = malignant melanoma; BCC = basal cell carcinoma; SCC = squamous cell carcinoma; SK = seborrheic keratosis; AK = actinic keratosis; DF = dermatofibroma. * p-value <0.05.

Additional results “at a glance”:

- Correct management improved: 85% → 86% (NS).
 - AI rescued management of 7 initially missed melanomas
- Diagnostic confidence increased (8.4 → 8.7/10, p<0.001)
- With clinician–AI disagreement:
 - 41% of incorrect diagnoses corrected
 - 15% of correct diagnoses overturned
- No accuracy difference by experience (>8 vs ≤8 years)
 - Less experienced assessors changed diagnosis slightly more often after AI

References

Esteva A, et al. *Nature*. 2017; Brinker TJ, et al. *J Med Internet Res*. 2018; Haenssle HA, et al. *Ann Oncol*. 2018; Tschandl P, et al. *Lancet Oncol*. 2019; Haenssle HA, et al. *Ann Oncol*. 2020; Takahashi S, et al. *J Med Syst*. 2024; Hekler A, et al. *Eur J Cancer*. 2019; Tschandl P, et al. *Nat Med*. 2020; Winkler JK, et al. *JAMA Dermatol*. 2023; Anriot J, et al. *JAMA Dermatol*. 2026; Han SS, et al. *J Invest Dermatol*. 2022; Yan S, et al. *Nat Med*. 2025; Marchetti MA, et al. *npj Digit Med*. 2023; Smak Gregoor AM, et al. *npj Digit Med*. 2023; Haggennüller S, et al. *Eur J Cancer*. 2021; Thomas L, et al. *Front Med (Lausanne)*. 2023; Menzies SW, et al. *Lancet Digit Health*. 2023.