Advances in Non-Invasive Dermatologic Imaging and Large Language Models for Skin Cancer and Inflammatory Disease Diagnosis

PhD thesis

Mehdi Boostani, MD

Semmelweis University Doctoral College, Rácz Károly

Conservative Medicine Division

Supervisor: Norbert Kiss, MD, PhD

Official reviewers: Imre Lőrinc Szabó, MD, PhD and András Budai, MD, PhD

Head of the Complex Examination Committee: György Reusz, DSc

Members of the Complex Examination Committee: András Bánvölgyi, MD, PhD and Katalin Müller, MD, PhD

Budapest

2026

Introduction

Skin diseases are among the most prevalent human disorders and include both inflammatory conditions and malignant neoplasms. Despite their frequency, many clinically important lesions remain difficult to diagnose accurately on visual inspection alone, especially in early stages or when benign and malignant conditions overlap in appearance.

This diagnostic uncertainty has direct clinical consequences. In oncology, delayed recognition of melanoma or squamous cell carcinoma may postpone life-saving treatment, whereas overdiagnosis leads to unnecessary biopsies, excisions, scarring, and anxiety. In inflammatory disease, misclassification of acne, rosacea, or hidradenitis suppurativa can delay appropriate therapy and contribute to chronic symptoms, progression, and psychosocial burden.

Dermatology is particularly well suited to technological innovation because diagnostic decisions are heavily image-based. Over recent decades, non-invasive imaging modalities such as dermoscopy, high-frequency ultrasound, optical coherence tomography, and reflectance confocal microscopy have expanded the clinician's ability to visualize subsurface morphology and tumor architecture without biopsy. These tools can improve diagnostic confidence, refine surgical planning, and reveal disease extension beyond what is visible to the naked eye.

At the same time, artificial intelligence has evolved from narrow image classifiers toward multimodal large language models capable of integrating clinical images with natural language reasoning. Unlike earlier systems restricted to a single task, these models can generate differential diagnoses, explain their predictions, and offer management suggestions. Because they are accessible through common digital devices, they may extend dermatologic support to patients and clinicians in settings where specialist care is limited.

However, accessibility does not guarantee safety. The real value of both advanced imaging and large language models depends on disease-specific validation, reproducibility, and careful alignment with real clinical workflows. This thesis therefore examines how these technologies perform across several distinct dermatologic problems,

from oncologic margin assessment to inflammatory disease recognition and staging. The overall technology-disease framework addressed in this thesis is summarized in **Table 1**.

Table 1. Technology-disease framework of the studies included in the thesis. Abbreviations: AK: Actinic keratosis, BCC: Basal cell carcinoma, DG-HFUS: Dermoscopy-guided high-frequency ultrasound, HS: Hidradenitis suppurativa, SCC: Squamous cell carcinoma

Technology	**Clinical question**	**Main value tested**
DG-HFUS	Preoperative BCC margin delineation	Detection of subclinical lateral extension before excision
GPT-4o and Gemini	Melanoma vs nevi	Sensitivity and specificity in image-based malignant lesion recognition
GPT-4o and Gemini	AK vs SCC	Ability to distinguish premalignant from invasive keratinocyte lesions
GPT-4o and Gemini	Acne vs rosacea	Recognition of overlapping inflammatory facial dermatoses and their subtypes
GPT-4o, Gemini, Claude	HS staging	Recognition, Hurley staging, and treatment-oriented support

Objectives

- To determine whether dermoscopy-guided high-frequency ultrasound improves preoperative identification of basal cell carcinoma margins and whether it may serve as a practical adjunct to conventional excision planning.
- To evaluate the diagnostic performance of multimodal large language models in distinguishing melanoma from melanocytic nevi using both clinical and dermoscopic photographs.
- To assess whether contemporary large language models can differentiate actinic keratosis from squamous cell carcinoma with clinically meaningful accuracy.
- To investigate the ability of large language models to distinguish acne from rosacea and to classify clinically relevant inflammatory subtypes.
- To examine whether large language models can recognize hidradenitis suppurativa, assign Hurley stage, and provide management suggestions aligned with disease severity.
- To compare different model behaviors across tasks in order to define strengths, limitations, and the potential complementary role of imaging and artificial intelligence in dermatologic practice.

Methods

Dermoscopy-guided high-frequency ultrasound for basal cell carcinoma

A prospective single-center study was performed at Semmelweis University between June 2024 and March 2025. Adults with clinically or dermoscopically suspected basal cell carcinoma were enrolled after informed consent. Lesions were photographed, dermoscopic margins were drawn 2 mm beyond visible tumor borders, and four oriented margins were assessed with a Dermus SkinScanner device. Ultrasound findings at the marked borders were compared with blinded histopathology after excision. Sensitivity, specificity, predictive values, likelihood ratios, and correlation of tumor extension measurements were calculated.

Large language models in melanoma versus nevi classification

A histopathology-verified melanocytic image dataset comprising 807 photographs, including both clinical and dermoscopic views, was evaluated using GPT-4o and Gemini 2.0 Flash. Each image was presented with a standardized research prompt, and ambiguous multi-answer responses were resolved using a follow-up request to choose the single most likely diagnosis. Model predictions were compared with the reference diagnosis to calculate sensitivity, specificity, accuracy, predictive values, and likelihood ratios.

Large language models in actinic keratosis and squamous cell carcinoma

Patients with actinic keratosis or squamous cell carcinoma were included between April 2022 and December 2024. Squamous cell carcinoma required histopathologic confirmation, while actinic keratosis was accepted either by histology or by concordant dermatologist diagnosis with follow-up. Standardized clinical images were analyzed by GPT-4o and Gemini 2.0 Flash using the same structured prompt strategy, with diagnostic performance quantified against the reference diagnosis.

Large language models in acne and rosacea

Clinical photographs of acne and rosacea were obtained from patients recruited between December 2021 and December 2024. Two dermatologists independently classified diagnoses and subtypes, with a third resolving disagreements. GPT-4o and Gemini 2.0 Flash were then asked to identify the most likely diagnosis and, when correct, the most

likely subtype. Overall disease recognition and subtype-level performance were compared with dermatologist consensus.

Large language models in hidradenitis suppurativa staging

Standardized photographs from intertriginous sites affected by hidradenitis suppurativa were reviewed by dermatologists and staged according to the Hurley system. The same images were submitted to multimodal large language models, including GPT-4o, Gemini 2.0 Flash, and Claude Sonnet 3.7, to test disease recognition, stage assignment, response consistency, and the quality of treatment-oriented recommendations.

Conclusions

- Dermoscopy-guided high-frequency ultrasound provided highly accurate preoperative information for basal cell carcinoma margin assessment, with sensitivity of 94.4% and specificity of 93.0% for detecting tumor extension within 2 mm of the excision line. This supports its use as a practical adjunct for surgical planning, particularly when tissue preservation is important.
- Among the tested large language models, GPT-4o generally favored sensitivity, making it more suitable for triage-oriented tasks in which missed malignancy is unacceptable. Gemini 2.0 Flash tended to favor specificity, which reduced false positives but increased the risk of overlooking malignant lesions in certain settings.
- In melanoma recognition, GPT-4o achieved very high sensitivity on both clinical and dermoscopic images, while Gemini showed very high specificity. Their differing error profiles suggest that performance must be interpreted in relation to the clinical purpose of the tool rather than by a single summary metric alone.
- For actinic keratosis versus squamous cell carcinoma, GPT-4o outperformed Gemini overall; however, both models had insufficient negative predictive value. A non-malignant output from current large language models therefore cannot safely exclude squamous cell carcinoma in real clinical practice.
- For acne and rosacea, GPT-4o demonstrated strong overall diagnostic accuracy, indicating real promise for first-line supportive assessment. Nevertheless, subtype classification remained substantially weaker, limiting the immediate reliability of these systems for therapy selection without specialist oversight.
- In hidradenitis suppurativa, GPT-4o showed strong recognition, good staging performance, and high reproducibility, yet important staging errors persisted, especially between Hurley stages II and III. This means that large language models may assist but should not replace clinical severity assessment in complex inflammatory disease.
- Taken together, the studies show two complementary innovation pathways in dermatology. Advanced imaging offers structural precision and objective lesion characterization, while large language models offer accessible, scalable interpretive support. The future clinical value of these technologies lies in their integration,

supported by broader validation datasets, disease-specific benchmarking, and careful implementation in supervised workflows.

Bibliography

Candidate's publications related to the thesis

- Boostani M, Wortsman X, Pellacani G, Cantisani C, Suppa M, Mohos A, Goldust M, Pietkiewicz P, Bánvölgyi A, Holló P, Wikonkál NM, Avanaki K, Paragh G, Lőrincz K, Kiss N. Dermoscopy-guided high-frequency ultrasound for preoperative assessment of basal cell carcinoma lateral margins: a pilot study. British Journal of Dermatology. 2025; ljaf209.
- Boostani M, Lallas A, Goldust M, Nádudvari N, Lőrincz K, Bánvölgyi A, Holló P, Wikonkál NM, Paragh G, Kiss N. Diagnostic performance of multimodal large language models in distinguishing melanoma from nevi in clinical and dermoscopic images. JAAD International. 2025.
- Boostani M, Pellacani G, Goldust M, Nádudvari N, Rátky D, Cantisani C, Lőrincz K, Bánvölgyi A, Wikonkál NM, Paragh G, Kiss N. Diagnosing actinic keratosis and squamous cell carcinoma with large language models from clinical images. International Journal of Dermatology. 2025.
- Boostani M, Bánvölgyi A, Goldust M, Cantisani C, Pietkiewicz P, Lőrincz K, Holló P, Wikonkál NM, Paragh G, Kiss N. Diagnostic performance of GPT-4o and Gemini Flash 2.0 in acne and rosacea. International Journal of Dermatology. 2025.
- Boostani M, Bánvölgyi A, Zouboulis CC, Goldfarb N, Suppa M, Goldust M, Lőrincz K, Kiss T, Nádudvari N, Holló P, Wikonkál NM, Paragh G, Kiss N. Large language models in evaluating hidradenitis suppurativa from clinical images. Journal of the European Academy of Dermatology and Venereology. 2025.

Other candidate's publications

- Kiss N, Boostani M. Ultra-wideband raster-scanning optoacoustic mesoscopy as a histologic-grade tool for melanoma depth assessment. Journal of the European Academy of Dermatology. 2025.
- Cantisani C, Hamilton-Maikle PK, Di Segni M, Banvolgyi A, Paolino G, Gonzales S, Boostani M, Pellacani G, Paragh G, Kiss N. Reddish/bluish benign lesions evaluated by Dermus SkinScanner. Journal of Ultrasound. 2025.
- Boostani M, Pellacani G, Wortsman X, Suppa M, Goldust M, Cantisani C, Pietkiewicz P, Lőrincz K, Bánvölgyi A, Wikonkál NM, Huss WJ, Avanaki K, Paragh G, Kiss N. FDA and EMA-approved noninvasive imaging techniques for basal cell carcinoma subtyping: a systematic review. JAAD International. 2025.
- Bánvölgyi A, Lőrincz K, Boostani M, Bar-Ilan E, Hidvégi B, Medvecz M, Wikonkál N, Kiss N. Anaphylaxis with elevated procalcitonin mimicking sepsis: a literature review and report of two cases. Journal of Clinical Medicine. 2025.
- Czurkó N, Martyin K, Boostani M, Kerner T, Bánvölgyi A, Kuroli E, Medvecz M, Ostorházi E, Kiss N. Tinea incognita unmasked: diagnostic challenge in a male patient with a 3-year history of topical steroid use. International Journal of Dermatology. 2025.
- Boostani M, Bozsányi S, Suppa M, Cantisani C, Lőrincz K, Bánvölgyi A, Holló P, Wikonkál NM, Huss WJ, Brady KL, Paragh G, Kiss N. Novel imaging techniques for tumor margin detection in basal cell carcinoma: a systematic scoping review of FDA and EMA-approved imaging modalities. International Journal of Dermatology. 2024.
- Boostani M, Hamilton-Meikle PK, Bozsányi S, Farkas K, Szabó B, Bánvölgyi A, Lőrincz K, Holló P, Wikonkál N, Huss W, Paragh G, Kiss N. New imaging modalities for determining the tumor margins of basal cell carcinoma. Bőrgyógyászati és Venerológiai Szemle. 2024.
- Czurkó N, Bozsányi S, Meznerics FA, Szlávicz E, Becske M, Boostani M, Farkas K, Kasek R, Fejes E, Bánvölgyi A, Lőrincz K, Takács S, Holló P, Wikonkál N, Lázár I, Kiss N. Pszichodermatológiai faktorok szerepe a psoriasis hátterében. Bőrgyógyászati és Venerológiai Szemle. 2024.

- Bozsányi S, Boostani M, Farkas K, Hamilton-Meikle P, Varga NN, Szabó B, Vasanits F, Kuroli E, Meznerics FA, Lőrincz K, Holló P, Bánvölgyi A, Wikonkál NM, Paragh G, Kiss N. Optically guided high-frequency ultrasound to differentiate high-risk basal cell carcinoma subtypes: a single-centre prospective study. Journal of Clinical Medicine. 2023.
- Varga NN, Boostani M, Farkas K, Bánvölgyi A, Lőrincz K, Posta M, Lihacova I, Lihachev A, Medvecz M, Holló P, Paragh G, Wikonkál NM, Bozsányi S, Kiss N. Optically guided high-frequency ultrasound shows superior efficacy for preoperative estimation of Breslow thickness in comparison with multispectral imaging: a single-center prospective validation study. Cancers (Basel). 2023.
- Cantisani C, Baja AV, Gargano L, Rossi G, Ardigò M, Soda G, Boostani M, Kiss N, Pellacani G. Optical coherence tomography as a valuable tool for the evaluation of cutaneous Kaposi sarcoma treated with imiquimod 5% cream. Diagnostics (Basel). 2023.
- Bozsányi S, Czurkó N, Becske M, Kasek R, Lázár BK, Boostani M, Meznerics FA, Farkas K, Varga NN, Gulyás L, Bánvölgyi A, Fehér BA, Fejes E, Lőrincz K, Kovács A, Gergely H, Takács S, Holló P, Kiss N, Wikonkál N, Lázár I. Assessment of frontal hemispherical lateralization in plaque psoriasis and atopic dermatitis. Journal of Clinical Medicine. 2023.