SEMMELWEIS EGYETEM

DOKTORI ISKOLA

Ph.D. értekezések

3470.

MEHDI BOOSTANI

Bőrgyógyászat és venerológia

című program

Programvezető: Dr. Sárdy Miklós, egyetemi tanár

Témavezető: Dr. Kiss Norbert, egyetemi adjunktus

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

List of Abbreviations

AI – Artificial intelligence

AK – Actinic keratosis

BCC – Basal cell carcinoma

CI – Confidence interval

CNN – Convolutional neural network

DG-HFUS – Dermoscopy-guided high-frequency ultrasound

EMA – European Medicines Agency

FDA – U.S. Food and Drug Administration

HFUS – High-frequency ultrasound

HS – Hidradenitis suppurativa

IHS4 – International Hidradenitis Suppurativa Severity Score System

LC-OCT – Line-field confocal optical coherence tomography

LLM – Large language model

LR+ – Positive likelihood ratio

LR− – Negative likelihood ratio

MMS – Mohs micrographic surgery

NPV – Negative predictive value

OCT – Optical coherence tomography

PhD – Doctor of Philosophy

PPV – Positive predictive value

QALY – Quality-adjusted life year

RCM – Reflectance confocal microscopy

SCC – Squamous cell carcinoma

UV – Ultraviolet

UVB – Ultraviolet B

Table of Contents

1. Introduction

1.1. The challenge of skin disease diagnosis

Skin diseases are among the most common human conditions, affecting hundreds of millions of people worldwide.(1) Skin diseases include a wide range of conditions that affect the skin's structure, appearance, and function.(2-9) Some are primarily inflammatory and chronic, causing long-term symptoms such as redness, pain, itching, or scarring.(10) Others are neoplastic, meaning they arise from uncontrolled growth of skin cells and may become cancerous.(11) Because many different diseases can appear similar on visual inspection, especially in early stages, accurate diagnosis often requires clinical expertise and, in some cases, confirmation by biopsy and histopathology.(12) Beyond direct medical costs, skin conditions contribute to lost productivity, long-term morbidity, and diminished quality of life, underscoring their public health significance.(7)

Melanoma is a malignant tumor of melanocytes and is less common than keratinocyte carcinomas but accounts for a disproportionate number of skin cancer deaths due to its metastatic potential.(13) The prognosis of melanoma strongly depends on early detection, when surgical excision is often curative.(14) Clinically, melanoma may present as a changing pigmented lesion and is classically associated with asymmetry, irregular borders, color variation, increasing diameter, or evolution over time (**Figure 1.**).(15) However, melanomas can also be small, subtle, amelanotic, or mimic benign nevi, which contributes to missed diagnoses and diagnostic variability even among clinicians.(16)

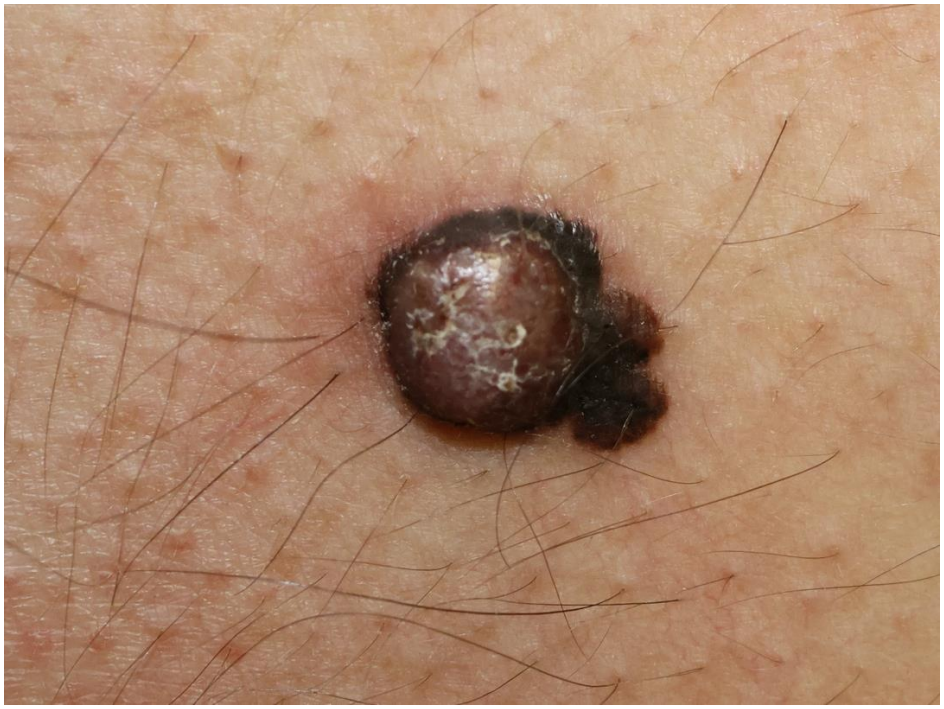

Figure 1. Clinical photograph of melanoma. Clinical photograph showing a melanoma lesion with asymmetric morphology, irregular borders, and heterogeneous pigmentation. Source: Own figure, based on clinical photographs from the Department of Dermatology, Venereology and Dermatooncology, Semmelweis University.

Squamous cell carcinoma (SCC) is another common skin cancer, also strongly linked to cumulative ultraviolet exposure.(17) Compared with BCC, SCC has a higher risk of deeper invasion and metastasis, especially when tumors are thick, poorly differentiated, located on high-risk sites (such as the lip or ear), or arise in immunosuppressed patients. SCC frequently presents as a firm, scaly, crusted papule, plaque, or nodule that may ulcerate or bleed (**Figure 2.**).(18) Early SCC can appear subtle and may be difficult to distinguish from precancerous lesions, creating a diagnostic gray zone that has important implications for timely treatment.(19)

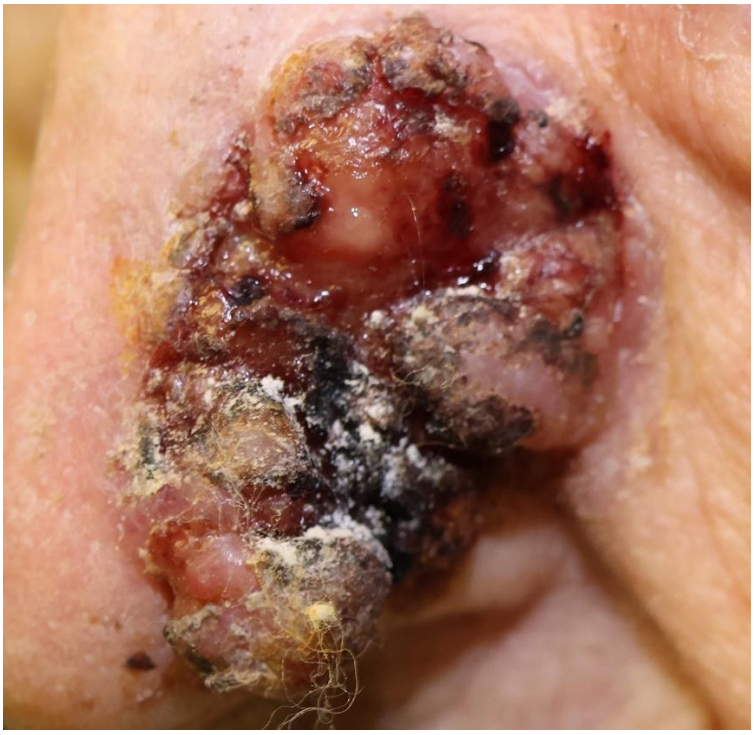

Figure 2. Clinical photograph of squamous cell carcinoma. Clinical photograph showing a squamous cell carcinoma presenting as a keratotic and crusted lesion. Source: Own figure, based on clinical photographs from the Department of Dermatology, Venereology and Dermatooncology, Semmelweis University.

Actinic keratosis (AK) is a very common precancerous lesion caused by chronic sun exposure and represents disordered growth of epidermal keratinocytes.(20) AK typically presents as rough, scaly, erythematous macules or plaques on sun-exposed skin such as the face, scalp, forearms, and hands (**Figure 3.**).(21) Many AKs remain stable or regress, but some can progress to SCC over time.(22) Because AK and early SCC can overlap clinically,(23) differentiating the two is often challenging, yet clinically important, as SCC requires more definitive management.

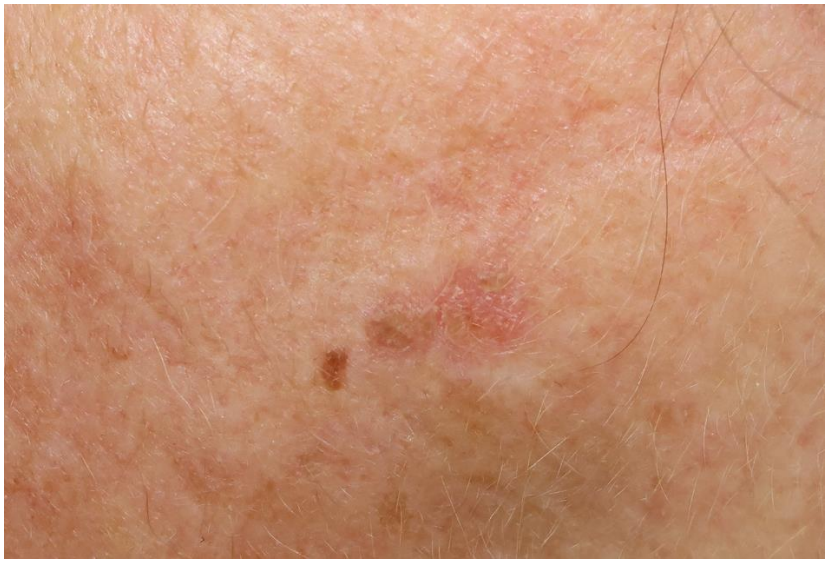

Figure 3. Clinical photograph of actinic keratosis. Clinical photograph showing an actinic keratosis lesion on chronically sun-exposed skin. Source: Own figure, based on clinical photographs from the Department of Dermatology, Venereology and Dermatooncology, Semmelweis University.

Basal cell carcinoma (BCC) is the most common type of skin cancer and typically arises in chronically sun-exposed areas such as the face, scalp, or neck.(24) It often grows slowly and rarely metastasizes, but it can be locally destructive if not treated.(25) Clinically, BCC may present as a pearly papule, a non-healing sore, or a scar-like plaque, sometimes with visible small surface blood vessels (**Figure 4.**).(26) The major clinical challenge is that BCC can extend beyond what is visible on the surface, particularly in higher-risk subtypes, which complicates surgical planning and increases the risk of incomplete excision and recurrence.(27)

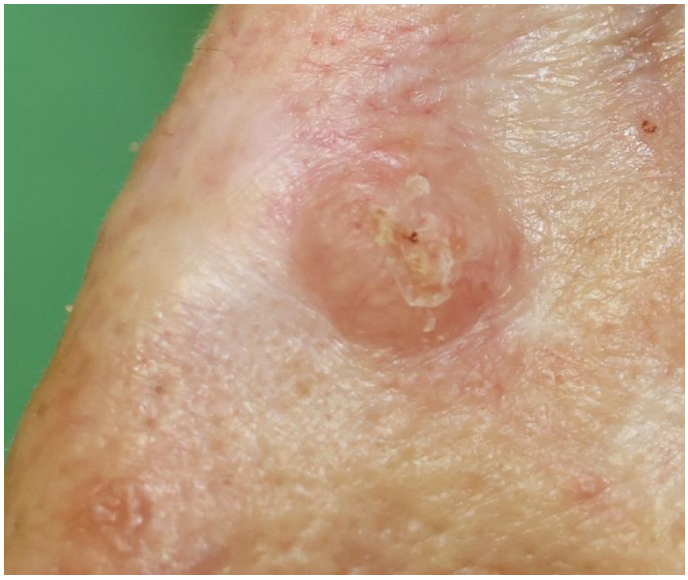

Figure 4. Clinical photograph of basal cell carcinoma. Clinical photograph showing a basal cell carcinoma with characteristic clinical morphology. Source: Own figure, based on clinical photographs from the Department of Dermatology, Venereology and Dermatooncology, Semmelweis University.

While BCC rarely metastasizes, it is locally aggressive and capable of invading deep structures including cartilage, muscle, and bone if left untreated.(28, 29) The cornerstone of management is surgical removal, but the central challenge lies in achieving complete excision with histologically tumor-free margins. Inadequate excision leads to recurrence, sometimes years later, with tumors often becoming larger and more aggressive, requiring more extensive surgery.(30-33) Conversely, overly wide margins result in unnecessary removal of healthy tissue, causing functional impairment and cosmetic disfigurement, a particularly important consideration given that the majority of BCCs occur on the head and neck.

Conventional excision with standardized safety margins (commonly 3–5 mm) is widely practiced, but its reliability is limited. Subclinical spread of BCC, particularly in high-risk histological subtypes such as infiltrative, morpheaform, or micronodular variants, often extends beyond these margins.(32, 34-36) As a result, recurrence rates after standard excision can vary for aggressive subtypes or tumors located in high-risk anatomical zones such as the nose, eyelids, or ears.(36-39)

Mohs micrographic surgery (MMS) addresses this problem by providing intraoperative histological margin control and remains the gold standard for high-risk or recurrent BCC.(40-43) However, MMS is resource-intensive, time-consuming, and not universally available, particularly outside specialized centers. The reality is that in many healthcare systems worldwide, MMS access is limited, and standard excision remains the most common treatment.(43-45) This disparity highlights a critical gap: clinicians lack a widely available, non-invasive, preoperative method to delineate tumor margins that could guide more precise excision planning, reduce recurrence, and preserve healthy tissue.(43-45)

Parallel challenges exist in the diagnosis of other skin cancers and inflammatory conditions. For example, distinguishing AK from early SCC, or differentiating benign nevi from melanoma, can be extremely difficult based on clinical examination alone.(46)

Inflammatory dermatoses such as acne, rosacea, and hidradenitis suppurativa (HS) add further complexity: these disorders are highly prevalent, frequently chronic, and often associated with pain, disfigurement, and significant psychosocial distress, leading to anxiety, depression, and social isolation.(47, 48)

Acne vulgaris is a common inflammatory disorder of the pilosebaceous unit.(49) It is driven by follicular plugging, increased sebum production, proliferation of Cutibacterium acnes, and inflammation.(49) Acne can range from mild comedonal lesions (open and closed comedones) to inflammatory papules, pustules, and deep nodules or cysts that may cause permanent scarring (**Figure 5.**).(50) Correct identification of severity and subtype matters because treatment differs substantially, ranging from topical agents to systemic antibiotics and isotretinoin.(51)

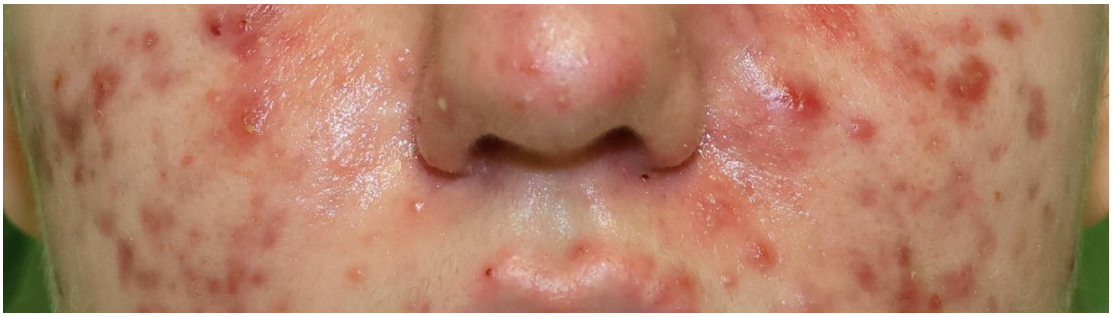

Figure 5. Clinical photograph of acne vulgaris. Clinical photograph showing acne vulgaris with inflammatory lesions.
Source: Own figure, based on clinical photographs from the Department of Dermatology, Venereology and Dermatooncology, Semmelweis University.

Rosacea is a chronic inflammatory condition that primarily affects the central face and is characterized by persistent redness, flushing, telangiectasia, and in some cases papules and pustules (**Figure 6.**).(52) Rosacea is often confused with acne, especially when inflammatory lesions are present; however, rosacea typically lacks comedones and may have prominent vascular features and heightened skin sensitivity.(53) Different subtypes have different clinical courses and treatment approaches, including topical anti-inflammatory therapies, oral agents, and vascular laser treatments.(54)

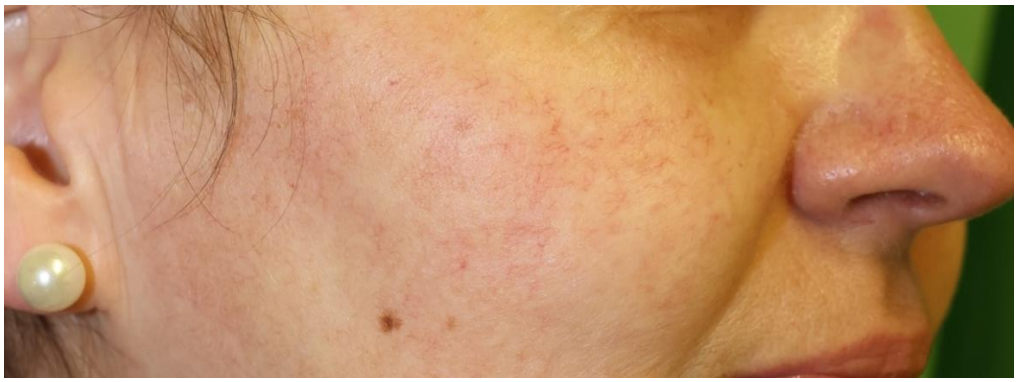

Figure 6. Clinical photograph of rosacea. Clinical photograph showing rosacea with central facial erythema and inflammatory lesions.
Source: Own figure, based on clinical photographs from the Department of Dermatology, Venereology and Dermatooncology, Semmelweis University.

HS is a chronic, relapsing inflammatory disease that affects hair follicle-bearing intertriginous areas such as the axillae, groin, and inframammary regions.(55) It is characterized by painful nodules, recurrent abscesses, sinus tract formation, and scarring (**Figure 7.**).(56) HS is frequently underdiagnosed or misdiagnosed early in its course, which contributes to prolonged diagnostic delay and progression to more advanced disease.(57) Severity is often described using the Hurley staging system, which has direct implications for therapy, including the need for biologic treatment or surgical intervention.(58)

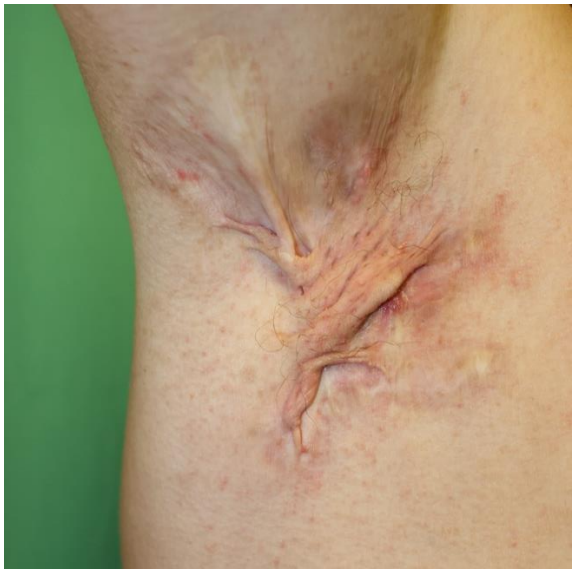

Figure 7. Clinical photograph of hidradenitis suppurativa. Clinical photograph showing hidradenitis suppurativa involving an intertriginous area. Source: Own figure, based on clinical photographs from the Department of Dermatology, Venereology and Dermatooncology, Semmelweis University.

Across this spectrum, the diagnostic process is further complicated by interobserver variability among clinicians,(59) the reliance on specialized training (e.g., dermoscopy), and the relative scarcity of dermatologists in many parts of the world. As a result, patients may face delays in diagnosis, under- or overtreatment, or misdiagnosis, each carrying consequences ranging from unnecessary procedures to missed opportunities for timely, life-saving intervention.(60)

This persistent diagnostic gap highlights the urgent need for supportive technologies that can improve accuracy, reduce delays, and expand access to expert-level care.

1.2. Skin imaging

The last three decades have seen rapid advances in non-invasive dermatologic imaging, providing clinicians with tools that extend beyond simple visual inspection. These modalities vary in resolution, penetration depth, and molecular specificity, offering complementary insights into skin structure and pathology.

Clinical photography remains a cornerstone of dermatology. Standardized images facilitate documentation of disease progression, support teledermatology, and provide indispensable material for education and research. High-quality clinical photographs also serve as the foundation for many AI training datasets, underscoring their dual role in practice and innovation.(61-64)

Dermoscopy enhances visualization of subsurface structures in the epidermis and papillary dermis that are invisible to the naked eye. Widely adopted for the detection of melanoma and other skin cancers, dermoscopy has expanded to encompass inflammatory dermatoses (inflammoscopy) and infectious diseases (entodermoscopy). By revealing specific vascular and pigment patterns, dermoscopy increases diagnostic accuracy and reduces unnecessary biopsies, although it requires specialized training and remains operator-dependent.(65, 66)

High-frequency ultrasound, typically using frequencies between 20 and 70 MHz, provides cross-sectional imaging of the skin and subcutaneous tissue to depths of several millimeters. It is particularly useful for measuring tumor thickness, evaluating preoperative margins, and assessing inflammatory or vascular processes. While resolution decreases with depth, high-frequency ultrasound offers a valuable balance of penetration and detail.(67-70)

Optical coherence tomography (OCT) generates micrometer-resolution, cross-sectional “optical biopsies” of skin to a depth of 1–2 mm, proving particularly useful in the evaluation of keratinocyte carcinomas.(71) Reflectance confocal microscopy (RCM), on the other hand, offers en face, cellular-level imaging of the epidermis and superficial

dermis, allowing near-histologic assessment in vivo.(72) Both modalities are powerful but are limited by cost, availability, and the need for specialized training.

Collectively, these modalities demonstrate how imaging has evolved from basic photography to sophisticated, multimodal approaches. Each contributes to earlier detection, improved characterization, and more precise surgical planning.

1.3. Large language models in dermatology

AI has become increasingly prominent in medical diagnostics, with dermatology serving as a leading field for innovation due to the visual nature of skin diseases.(73, 74) Early applications relied on convolutional neural networks (CNNs), which achieved dermatologist-level accuracy in specific domains such as melanoma detection.(75) Yet, these algorithms were constrained by narrow training objectives, limited generalizability, and inability to communicate results in a clinically interpretable way.

The arrival of multimodal large language models (LLMs) represents a profound shift.(46-48) These systems integrate image recognition with natural language reasoning, allowing them not only to classify images but also to explain their decisions, generate differential diagnoses, and propose treatment options. Their accessibility through web platforms and consumer devices makes them available to both clinicians and patients, creating new opportunities but also raising new risks.

To understand their potential and limitations, it is necessary to place LLMs in the context of the specific diseases they are being applied to.

1.3.1. Keratinocyte carcinomas: actinic keratosis and squamous cell carcinoma

AK is a common UV-induced premalignant lesion, often presenting as scaly erythematous patches or plaques on chronically sun-exposed areas. Although many remain stable, a subset progress to SCC, a malignancy capable of local invasion, tissue destruction, and metastasis.(76, 77) Differentiating between AK and early SCC is difficult because both may appear as rough keratotic lesions with subtle clinical differences.

From a patient safety perspective, the stakes are high. Missing SCC can result in invasive cancer with significant morbidity, while overcalling AK may lead to unnecessary biopsies or excisions. Histopathology remains the gold standard, but it is invasive, resource-intensive, and impractical for every suspicious lesion.

LLMs have been tested on this problem, yet studies show that while they can achieve reasonable sensitivity, their negative predictive values are insufficiently reliable. A "benign" or "non-cancerous" label from an LLM cannot be trusted to exclude SCC, highlighting one of the most pressing barriers to safe clinical integration.

1.3.2. Melanoma versus melanocytic nevi

Melanoma is the most aggressive skin cancer, with survival closely tied to early detection.(13) By contrast, melanocytic nevi are benign proliferations that are exceedingly common in the general population.(78) Clinicians must balance vigilance with caution: while every missed melanoma carries grave consequences, over-biopsy of benign nevi contributes to patient anxiety, scarring, and healthcare costs.(79)

Even with dermoscopy, which improves diagnostic accuracy, interobserver variability remains substantial.(80) LLMs demonstrate strong sensitivity in melanoma detection, sometimes outperforming older CNN-based algorithms. However, they may also generate false positives, leading to unnecessary excisions, or false negatives, risking delayed melanoma diagnosis. Because melanoma management is time-sensitive, even modest error rates carry serious implications for patient outcomes.

1.3.3. Acne and rosacea

Acne vulgaris is one of the most prevalent dermatologic disorders worldwide, affecting a large number of of adolescents and a growing number of adults.(81) Its subtypes, comedonal, inflammatory, nodulocystic, and steroid-induced, require different therapeutic approaches.(47, 81)

Rosacea, in contrast, is a chronic inflammatory condition primarily affecting the central face. Subtypes include erythematotelangiectatic, papulopustular, phymatous, and ocular

rosacea.(47, 81-84) Clinical overlap with acne is common: papules and pustules in rosacea may be mistaken for acne, while comedones distinguish acne but may be subtle. Misclassification is frequent and leads to mistreatment, for example, prescribing acne therapies such as topical retinoids in rosacea, which can exacerbate symptoms.

LLMs have shown promise in distinguishing acne from rosacea in clinical photographs.(47) However, their performance drops when tasked with subtyping, a crucial step for personalized therapy. For example, differentiating papulopustular rosacea from nodulocystic acne, or correctly identifying steroid-induced acne, remains inconsistent. This highlights the importance of systematic evaluation of LLMs in common, real-world conditions where diagnostic nuance has direct therapeutic impact.

1.3.4. Hidradenitis suppurativa

HS is a chronic, relapsing inflammatory disease affecting intertriginous areas such as the axillae and groin. Characterized by painful nodules, abscesses, and sinus tracts, HS is associated with profound physical and psychosocial burden. Patients frequently experience diagnostic delays averaging 7–10 years, during which disease progresses to more severe stages marked by extensive scarring and disability.(85-88)

The Hurley staging system guides management, but staging requires clinical expertise and experience.(89) Underestimation leads to undertreatment, while overestimation may expose patients to unnecessary systemic therapies.

LLMs offer an intriguing opportunity: they can be trained to recognize HS from clinical photographs and assign Hurley stages.(90) However, performance is inconsistent.(90) While some models approach dermatologists in recognition of HS itself, accuracy in staging varies, and treatment recommendations often lack reliability. This reflects both the potential and the pitfalls of relying on generalized AI systems in a complex, heterogeneous condition like HS.

1.3.5. Advantages and limitations of LLMs

Advantages

- Accessibility: Available globally on consumer devices, lowering barriers for patients and providers.(91)
- Generality: Can handle a wide variety of lesion types and inflammatory conditions.(92)
- Explainability: Communicate reasoning in natural language, making outputs more interpretable than traditional CNNs.(93)

Limitations

- Variable accuracy across diseases: Reliable for some (e.g., melanoma detection), poor for others (e.g., AK vs SCC).(92)
- Unsafe negative predictive values: Particularly concerning in malignant disease, where a false reassurance could delay treatment.(94)
- Subtype misclassification: Acne, rosacea, and HS subtyping remain inconsistent.(92)
- Dataset bias: Underrepresentation of darker skin tones and rare presentations limits generalizability.(92)
- Instability of responses: Susceptibility to prompt phrasing, contradictory answers, or non-responses.(95)

1.3.6. Safety and clinical integration

Safe integration will require rigorous validation across diverse populations, transparency regarding training data, and clinical guidelines that clearly define their scope of use.

In summary, LLMs hold promise to complement dermatology by providing accessible, image-based diagnostic support. However, their performance is highly disease-dependent. They perform well in some contexts, such as melanoma recognition, but falter in others, such as differentiating AK from SCC or accurately staging HS. These limitations mirror the findings of recent clinical studies and underscore the need for careful, disease-specific evaluation before LLMs can be integrated safely into clinical practice.

2. Objectives

In overview, our work aimed to investigate the diagnostic performance of multimodal LLMs in diverse dermatologic conditions, alongside the role of dermoscopy-guided high-frequency ultrasound (DG-HFUS) in preoperative tumor margin assessment. The overarching objective was to define the potential, limitations, and complementary value of these technologies in the diagnosis and management of both oncologic and inflammatory skin diseases.

Our first aim was to assess the utility of DG-HFUS in BCC. We investigated whether co-registration of dermoscopic and ultrasound images improves the precision of tumor margin detection compared with histopathology, and whether DG-HFUS could serve as a non-invasive adjunct or partial substitute for Mohs micrographic surgery in preoperative planning.

A second objective was to assess the diagnostic performance of LLMs in differentiating melanoma from melanocytic nevi using both clinical and dermoscopic images. Because early melanoma detection is crucial for prognosis, we aimed to quantify model sensitivity, specificity, and overall accuracy, and to explore whether combining different models could yield complementary strengths.

Our third objective was to evaluate the diagnostic accuracy of LLMs in distinguishing AK from SCC. Given the substantial clinical overlap between these lesions and the importance of early recognition of SCC, we aimed to assess whether state-of-the-art LLMs could correctly classify standardized clinical images, and to quantify their sensitivity, specificity, and predictive values in this context.

Our fourth objective was to analyze the performance of LLMs in common inflammatory dermatoses, specifically acne and rosacea. We aimed to determine whether these models can differentiate between the two conditions on the basis of clinical photographs, and to explore their capacity for correct subtype classification, which has direct therapeutic implications.

We also aimed to investigate the application of LLMs in HS. Our objective was to evaluate their ability to recognize HS from clinical images, to provide accurate Hurley

staging, and to assess whether their suggested management strategies align with established treatment guidelines.

Across all projects, we aimed to compare multiple LLM platforms (including GPT-4o, Gemini 2.0 Flash, and Claude Sonnet 3.7) in order to identify systematic differences in diagnostic behavior, responsiveness, and reliability. By doing so, we sought to characterize not only their individual capabilities but also their relative suitability for clinical decision support.

Finally, we aimed to place these findings in the broader context of clinical dermatology. We set out to discuss how advanced imaging and AI-driven interpretation may complement one another: imaging offering objective visualization of lesion structure and depth, while LLMs provide rapid, accessible interpretation. Together, our objective was to highlight the opportunities and challenges of integrating these technologies into dermatologic workflows in a way that is safe, equitable, and clinically meaningful.

3. Methods

3.1. Dermoscopy-guided high-frequency ultrasound for basal cell carcinoma margin assessment

This prospective, single-center study was conducted at the Department of Dermatology, Venereology, and Dermatooncology, Semmelweis University (Budapest, Hungary) between June 2024 and March 2025, with approval from the Institutional Ethics Committee (SE RKEB no. 16/2022). All participants provided written informed consent prior to enrollment. Patients over 18 years of age with clinically or dermoscopically suspected BCC were eligible if lesions were accessible to both dermoscopy and ultrasound. Exclusion criteria included bleeding tumors, previously biopsied or treated BCCs, lesions treated with topical therapies or Hedgehog pathway inhibitors, and tumors located in anatomically challenging regions such as the eyelids, eyebrows, or ears, where probe application was impractical. Histological confirmation of BCC was required for final inclusion.

All lesions underwent standardized clinical and dermoscopic photography. A dermoscopic margin was drawn 2 mm beyond visible tumor borders using a permanent marker, and lesions were oriented by labeling four margins (A: 12 o'clock, B: 3 o'clock, C: 6 o'clock, D: 9 o'clock). DG-HFUS imaging was then performed using the Dermus SkinScanner (Dermus Ltd., Budapest, Hungary), which employs a 20–40 MHz transducer with a penetration depth of up to 10 mm and an optical field of view of 15 × 15 mm. Each acquisition simultaneously captured paired optical and ultrasound images, displayed in color scale for improved contrast. At least five cross-sectional images were obtained for each lesion, and larger tumors were evaluated with multiple representative scans. (**Figure 8.**)

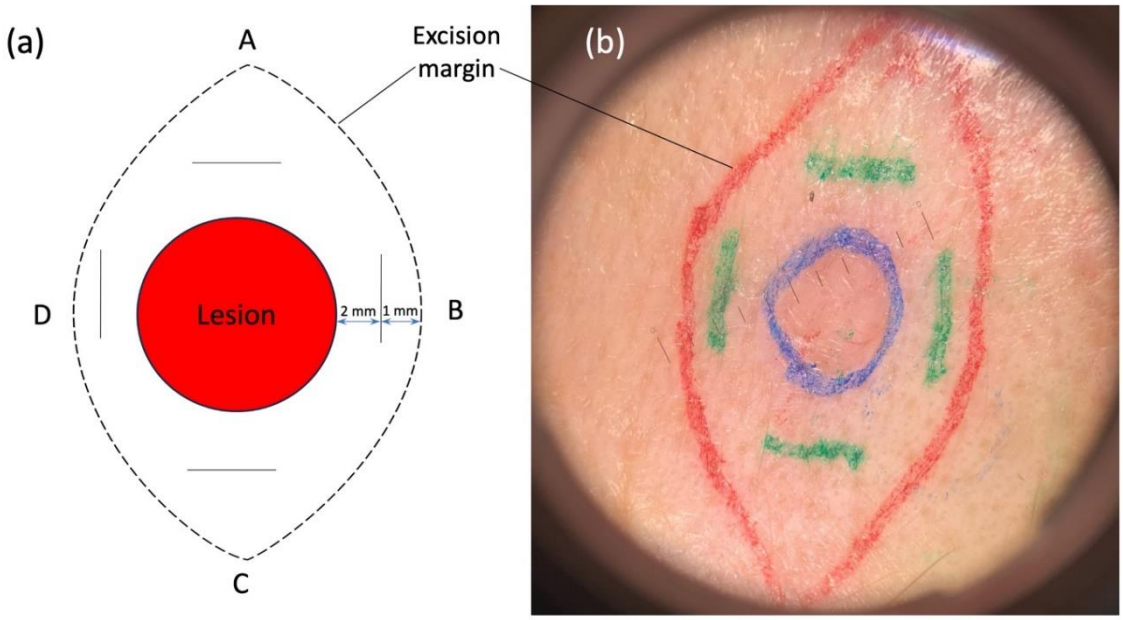

Figure 8. Lesion assessment, marking, and excision process. Illustration of the procedural steps used for lesion assessment, margin marking, dermoscopy-guided high-frequency ultrasound imaging, and excision. Panel a provides a schematic representation of the workflow. A dermoscopic border was drawn 2 mm from any visible tumor structure using a permanent marker, and the lesion was oriented by labeling four margins: A, 12 o'clock; B, 3 o'clock; C, 6 o'clock; and D, 9 o'clock. Dermoscopy-guided high-frequency ultrasound imaging was then performed at these four margins. A superficial incision was made along the marked line after imaging to create a histological landmark. The lesion and surrounding skin were then cleaned with an antiseptic solution, local anesthesia was administered, and the lesion was excised with an additional 1 mm margin beyond the 2 mm dermoscopic border, resulting in a total excision margin of 3 mm. Panel b shows a clinical image of a lesion with the dermoscopic margin marked in blue, orientation labels in green, and the excision margin in red before the procedure. Source: Own figure, created by the author.

Margins were classified as positive when DG-HFUS identified hypoechoic tumor nests consistent with BCC within or beyond 2 mm from the dermoscopic marker line on the tumor side. Margins were considered negative if no BCC criteria were present within this zone. Following imaging, lesions were excised with an additional 1 mm margin beyond the 2 mm dermoscopic line, resulting in a total excision margin of 3 mm. (**Figure 9.**)

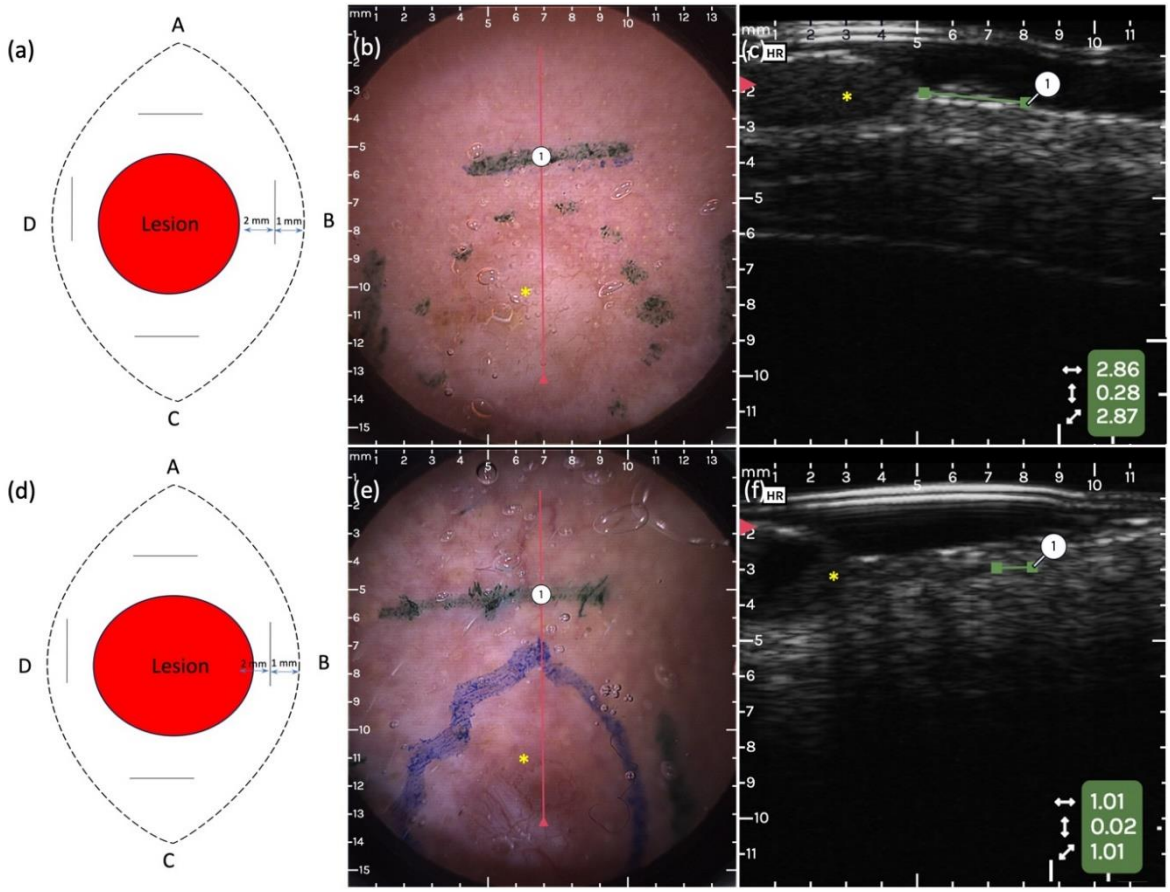

Figure 9. Dermoscopy-guided high-frequency ultrasound assessment of tumor invasion. Illustration of dermoscopy-guided high-frequency ultrasound assessment for detecting basal cell carcinoma invasion beyond the permanent marker line. Panels a–c show a case in which no basal cell carcinoma criteria were identified within 2 mm from the permanent marker line on the tumor side, and the margin was classified as negative. Panels d–f show a case in which dermoscopy-guided high-frequency ultrasound detected basal cell carcinoma criteria within 2 mm from the marker line, and the margin was classified as positive. Panels a and d are schematic illustrations of negative and positive margin cases, respectively. Panels b and e show the corresponding dermoscopy-guided optical images from the dermoscopy-guided high-frequency ultrasound system. Panels c and f show the corresponding ultrasound images. In the ultrasound images, the green line indicates the measured distance from the tumor to the superficial incision, and the yellow asterisk marks the tumor location. Dermoscopy-guided high-frequency ultrasound evaluation was based on the detection of hypoechoic tumor nests among hyperechoic collagen fibers to

determine the extent of tumor infiltration at the investigated margin. Source: Own figure, created by the author.

A superficial incision along the marker line created a histological landmark, and a suture placed at the 12 o'clock position ensured orientation during histological processing. Specimens were fixed in 10% formalin, embedded in paraffin, sectioned vertically, and stained with hematoxylin and eosin. Shrinkage correction variables were applied to account for tissue contraction during processing. Histological assessment, performed by an experienced dermatopathologist blinded to imaging results, determined tumor distance from the superficial incision. (**Figure 10.**)

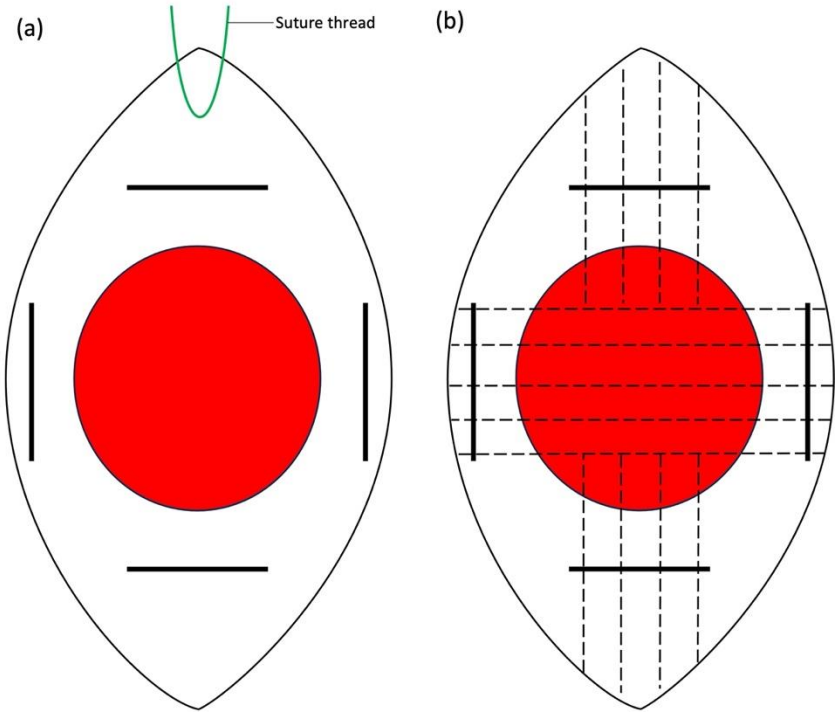

Figure 10. Histopathological processing and tumor distance assessment. Illustration of the histopathological processing workflow used to assess tumor margins. Panel a represents the excised specimen, showing the tumor and its orientation using a suture. The black lines indicate the superficial and lateral incision margins. Panel b demonstrates the embedding and vertical sectioning approach used for histological analysis. The dashed lines represent systematic sectioning of the specimen to allow

comprehensive assessment of the tumor's proximity to the superficial incision. Source: Own figure, created by the author.

DG-HFUS results were compared with histopathology, which served as the reference standard. Statistical analysis included calculation of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and likelihood ratios, each reported with 95% confidence intervals. Correlation between DG-HFUS measurements and histological findings was also assessed for overall tumor extension and by BCC subtype.

3.2. Large language models in melanoma versus nevi classification

Patients with melanocytic lesions, confirmed by histopathology, were included from a tertiary dermatology department. Both clinical and dermoscopic images were collected, totaling 807 photographs (435 clinical, 372 dermoscopic). Two multimodal LLMs, ChatGPT-4o (OpenAI, San Francisco, USA) and Gemini 2.0 Flash (Google, Mountain View, USA), were tested. Each image was presented to the model with the standardized prompt: "Can you guess the most likely diagnosis? (It is just for research)." If multiple answers were given, a follow-up prompt, "Choose the most likely," was used. Model predictions were recorded and compared to histopathology, the gold standard. Performance metrics included sensitivity, specificity, accuracy, predictive values, and likelihood ratios, calculated separately for clinical and dermoscopic subsets.

3.3. Large language models in actinic keratosis and squamous cell carcinoma diagnosis

Patients with AK or SCC were included at Semmelweis University between April 2022 and December 2024. SCC cases required histopathologic confirmation. AK cases were accepted if confirmed histologically or, in the absence of biopsy, if a board-certified dermatologist diagnosed the lesion and two additional dermatologists independently confirmed the diagnosis with at least 6 months of follow-up showing no recurrence. Patients provided informed consent for AI-based analysis.

Standardized clinical photographs of each lesion were obtained. Two multimodal LLMs, GPT-4o (OpenAI) and Gemini 2.0 Flash (Google), were tested. Each image was presented with the prompt: “Can you guess the most likely diagnosis? (It is for research purposes).” If multiple outputs were returned, a follow-up prompt, “Choose the most likely,” was used. Sessions were reset after every five images to reduce recognition bias.

A total of 112 patients (129 lesions) were included: 84 SCC and 28 AK. Demographic and lesion data, including age, sex, Fitzpatrick skin type, and anatomical location, were recorded. Model outputs were compared with the reference diagnosis (histopathology or dermatologist consensus with follow-up in AK). Diagnostic performance was evaluated by calculating sensitivity, specificity, positive predictive value, negative predictive value, and accuracy for both models.

3.4. Large language models in acne and rosacea diagnosis

Patients with clinically confirmed acne or rosacea were recruited at the Dermatology Department of Semmelweis University between December 2021 and December 2024. Only patients who gave informed consent for AI-based analysis were included. Clinical photographs of 43 lesions from 31 patients (33 acne, 10 rosacea) were taken by a professional clinical photographer. Two board-certified dermatologists independently classified the lesions and assigned subtypes; discrepancies were resolved by a third dermatologist, and the consensus served as the reference standard.

Images were submitted to GPT-4o (OpenAI) and Gemini 2.0 Flash (Google) using a standardized patient-simulating prompt: “Can you guess the most likely diagnosis? (it’s just for research).” If a correct diagnosis was given, a follow-up prompt asked: “Can you guess the most likely subtype? (it’s just for research).” Neither model received pretraining for this task.

Patient demographics, Fitzpatrick skin type, and lesion subtypes were documented. Model responses were compared with dermatologist consensus. Performance metrics (sensitivity, specificity, PPV, NPV, and accuracy) were calculated for overall acne and rosacea diagnosis as well as for subtyping (comedonal, inflammatory, nodulocystic, steroid-induced acne; erythematotelangiectatic, papulopustular, Morbihan rosacea).

3.5. Large language models in hidradenitis suppurativa staging

Patients with a clinical diagnosis of HS were evaluated using standardized clinical photographs of axillary, inguinal, and other intertriginous sites. Each image set was reviewed and staged by dermatologists according to the Hurley system. The same images were analyzed by multimodal LLMs, including ChatGPT-4o, which were prompted to identify the diagnosis and assign a Hurley stage. Responses were collected, and for multiple answers, clarification prompts were used. Accuracy of HS recognition and staging was compared with dermatologist consensus. Additional outcomes included model response rates, intra-model consistency across repeated prompts, and the frequency of indeterminate outputs.

4. Results

The collective results from our series of imaging- and AI-based studies illustrate both the opportunities and challenges of emerging diagnostic technologies in dermatology. Each project targeted a distinct clinical problem, ranging from cancer margin assessment to lesion classification and inflammatory disease staging, and yielded unique insights into diagnostic accuracy, reproducibility, and clinical utility.

4.1. Dermoscopy-guided high-frequency ultrasound for basal cell carcinoma margin assessment

The DG-HFUS study demonstrated that combining dermoscopic surface mapping with high-frequency ultrasound provided reliable preoperative information on tumor margins in patients with BCC. The approach yielded cross-sectional visualization of tumor depth and lateral spread, identifying areas of irregular extension that would otherwise be underestimated by dermoscopy alone. Compared with histopathology, DG-HFUS achieved a sensitivity of 94.4% and specificity of 93.0% for detecting tumor margins within 2 mm of the excision line, with an overall diagnostic accuracy exceeding 93%. Importantly, DG-HFUS was able to detect “skip areas” and subclinical tumor buds extending laterally, which were confirmed in histology.

4.2. Large language models in melanoma versus nevi classification

In the evaluation of 807 melanocytic lesions, multimodal LLMs exhibited complementary performance profiles. ChatGPT-4o consistently favored sensitivity, correctly identifying 96.5% of melanomas in dermoscopic images and 90.6% in clinical photographs, thus minimizing the likelihood of missed melanomas. However, this high sensitivity was offset by reduced specificity, particularly in dermoscopy (53.7%), leading to frequent misclassification of benign nevi as malignant. In contrast, Gemini 2.0 Flash prioritized specificity, reaching 98.8% in dermoscopy and 90.6% in clinical photographs, but at the expense of sensitivity, missing nearly half of melanomas in dermoscopic cases. When both models were considered together, at least one identified melanoma in 96.2% of cases.

4.3. Large language models in actinic keratosis versus squamous cell carcinoma

The challenge of distinguishing AK from SCC was explored in a cohort of 112 patients with 129 lesions. GPT-4o substantially outperformed Gemini, correctly identifying 64.3% of SCCs and 70.5% of AKs, yielding an overall accuracy of 80.5%. Its specificity (88.6%) and PPV (91.5%) were strong, suggesting high reliability when a lesion was classified as malignant. Gemini, however, achieved perfect specificity (100%) but very low sensitivity (40%), making it highly conservative: it rarely misclassified benign lesions but missed the majority of cancers. Both models performed poorly in negative predictive value (57.5% for GPT-4o, 50.0% for Gemini).

4.4. Large language models in acne and rosacea diagnosis

In the study of 43 lesions from 31 patients with acne or rosacea, GPT-4o achieved an impressive 93% overall diagnostic accuracy, with sensitivities of 90.9% for acne and 100% for rosacea, coupled with near-perfect specificities. This performance underscores the model's robustness in differentiating between two clinically overlapping conditions that often challenge patients and even non-specialist physicians. However, performance dropped when subtyping was required. For acne, GPT-4o correctly assigned subtypes (comedonal, inflammatory, nodulocystic, or steroid-induced) in just over half of cases, with misclassification particularly common in nodulocystic versus severe inflammatory forms (**Table 1.**). For rosacea, while primary diagnosis was excellent, subtyping into erythematotelangiectatic, papulopustular, or Morbihan forms remained inconsistent, with sensitivities often falling below 60%. Gemini 2.0 Flash, by comparison, provided responses in only 21% of cases, precluding meaningful analysis (**Table 2.**).

Table 1. Diagnostic accuracy of GPT-4o for estimating acne subtypes. PPV: Positive Predictive Value; NPV: Negative Predictive Value

Subtype	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**
Comedonial	0	100	-	97%
Inflammatory	100	64.3	33	100
Nodulocystic	54.1	100	100	45
Steroid-induced	0	100	-	90.9
Overall	54.6	89.9	64.3	85.9

Table 2. Diagnostic accuracy of GPT-4o for estimating rosacea subtypes. PPV: Positive Predictive Value; NPV: Negative Predictive Value

Subtype	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**
Papulopustular	42.9	100	100	42.9
Erythematotelangiectatic	100	50	33.3	100
Morbihan disease	0	100	-	90
Overall	50	80	55.6	76.2

4.5. Large language models in hidradenitis suppurativa staging

The HS study revealed the strengths and weaknesses of LLMs in managing complex, heterogeneous inflammatory diseases. ChatGPT-4o correctly identified HS in nearly nine out of ten cases (87.3% diagnostic accuracy), outperforming both Gemini (32.4%) and Claude Sonnet 3.7 (77.5%). When limited to correctly diagnosed HS cases, ChatGPT-4o

assigned Hurley stages with 79.0% accuracy, while Gemini achieved 82.6% and Claude 65.5%. The majority of errors involved misclassification between stages II and III, reflecting the subtle clinical gradations that challenge even human observers (**Table 3.**). Treatment recommendation accuracy followed a similar trend: ChatGPT-4o reached 88.7%, while Gemini and Claude lagged at 65.2% and 54.5%, respectively. Importantly, ChatGPT-4o demonstrated high test–retest reproducibility, with consistency above 85% when the same cases were presented repeatedly, compared to lower reproducibility for Gemini and Claude.

Table 3. Performance metrics of large language models in Hurley staging of hidradenitis suppurativa. Presentation of the diagnostic performance metrics of three large language models (ChatGPT-4o, Gemini 2.0 Flash, and Claude 3.7 Sonnet) in staging hidradenitis suppurativa using the Hurley classification system. Sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, negative likelihood ratio, and overall accuracy are reported for each model at Hurley stages I, II, III, and overall performance. Confidence intervals are provided for each metric. CI95: 95% confidence interval, LR+: positive likelihood ratio, LR-: negative likelihood ratio, PPV: positive predictive value, NPV: negative predictive value.

Model	**Hurley stage**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**LR+**	**LR-**
ChatGPT-4o	I	100% (CI95: 47.8—100)	96.5% (CI95: 87.9—99.6)	71.4% (CI95: 39.1—90.7)	100% (CI95: 93.5—100)	28.5 (CI95: 7.3—111.2)	0 (-)
	II	77.8% (CI95: 62.9—88.8)	82.4% (CI95: 56.6—96.2)	92.1% (CI95: 80.5—97.1)	58.3% (CI95: 43.7—71.6)	4.4 (CI95: 1.6—12.5)	0.27 (CI95: 0.15—0.49)
	III	75% (CI95: 42.8—94.5)	86% (CI95: 73.3—94.2)	56.3% (CI95:	93.5% (CI95:	5.4 (CI95:	0.29 (CI95:

				37.5—73.3)	84.2—97.5)	2.5—11.5)	0.11—0.78)
	Overall	79% (CI95: 66.8—88.3)	90.3% (CI95: 83.7—94.9)	80.3% (CI95: 70.1—87.7)	89.6% (CI95: 84.1—93.3)	8.2 (CI95: 4.7—14.2)	0.23 (CI95: 0.14—0.38)
Gemini 2.0 Flash	I	100% (CI95: 29.2—100)	90% (CI95: 68.3—98.8)	60% (CI95: 28.7—84.8)	100% (CI95: 81.5—100)	10 (CI95: 2.7—37.3)	0 (-)
	II	88.9% (CI95: 65.3—98.6)	80% (CI95: 28.4—99.5)	94.1% (CI95: 73.3—98.9)	66.7% (CI95: 33.5—88.8)	4.4 (CI95: 0.76—25.9)	0.14 (CI95: 0.04—0.55)
	III	0% (CI95: 0—84.2)	100% (CI95: 83.9—100)	-	91.3% (CI95: 91.3—91.3)	-	1 (CI95: 1—1)
	Overall	82.6% (CI95: 61.2—95.1)	93.5% (CI95: 82.1—98.6)	86.4% (CI95: 67.6—95.1)	91.5% (CI95: 81.5—96.3)	12.7 (CI95: 4.2—38.4)	0.19 (CI95: 0.08—0.45)
Claude 3.7 Sonnet	I	100% (CI95: 39.8—100)	74.5% (CI95: 60.4—85.7)	23.5% (CI95: 16.1—33)	100% (CI95: 90.8—100)	3.9 (CI95: 2.5—6.3)	0 (-)
	II	55.8% (CI95: 39.9—70.9)	75% (CI95: 42.8—94.5)	88.9% (CI95:	32.1% (CI95:	2.2 (CI95:	0.59 (CI95:

				74.4—95.7)	22.9—43.1)	0.81—6.2)	0.37—0.94)
	III	50% (CI95: 15.7—84.3)	87.2% (CI95: 74.3—95.2)	40% (CI95: 19.4—64.9)	91.1% (CI95: 83.6—95.4)	3.92 (CI95: 1.4—10.9)	0.57 (CI95: 0.28—1.2)
	Overall	58.2% (CI95: 44.1—71.4)	80% (CI95: 71.3—87)	59.3% (CI95: 48.5—69.2)	79.3% (CI95: 73.4—84.1)	2.9 (CI95: 1.9—4.5)	0.52 (CI95: 0.38—0.72)

5. Discussion

The series of studies presented here demonstrate the potential of both advanced imaging modalities and LLMs to address persistent challenges in dermatologic diagnosis and management. Taken together, these findings highlight the complementary but distinct contributions of structural visualization technologies and multimodal AI. While imaging excels at objective assessment of tissue architecture and surgical planning, LLMs offer accessible diagnostic support that can scale across patient populations. However, both approaches face important limitations, particularly regarding generalizability, interpretability, and integration into real-world workflows.

BCC remains the most common skin cancer worldwide, with incidence continuing to rise in aging and fair-skinned populations.(96) Despite its indolent growth,(28, 97) BCC poses substantial morbidity due to local invasion and tissue destruction, especially in cosmetically sensitive areas such as the face.(28, 97) Accurate preoperative assessment of tumor margins is therefore critical to balancing complete clearance with tissue conservation.(98-100) Our DG-HFUS study addressed this gap by combining surface mapping with depth-resolved imaging.(101) DG-HFUS demonstrated strong concordance with histopathology in detecting tumor extension, identifying not only depth but also lateral "skip areas" that are frequently underestimated by visual inspection. This ability to reveal subclinical spread has direct implications for reducing both recurrence rates and unnecessary tissue sacrifice, positioning DG-HFUS as a valuable adjunct to Mohs micrographic surgery or standard excision.

Similarly, our work with line-field confocal optical coherence tomography (LC-OCT) reinforced the promise of non-invasive imaging for intraoperative decision-making.(102) By providing real-time en face and cross-sectional visualization of architectural disarray, basaloid nests, and peripheral palisading, LC-OCT could anticipate histologic findings and align closely with frozen-section analysis.(102) Such tools may shorten surgical duration, improve accuracy of first-stage clearance, and ultimately reduce the burden of repeated tissue excisions.

The application of LLMs to melanoma, AK, and SCC evaluation illustrates the dual strengths and weaknesses of these systems. In the melanoma study, ChatGPT-4o achieved very high sensitivity across both clinical (90.6%) and dermoscopic (96.5%) images,

ensuring that malignant lesions were rarely overlooked.(46) This level of sensitivity surpasses many earlier CNNs and suggests that transformer-based vision encoders confer improved robustness by integrating global lesion architecture with local features. However, the trade-off was reduced specificity, particularly in dermoscopy (53.7%), resulting in high false-positive rates and potential over-biopsy. Conversely, Gemini 2.0 Flash prioritized specificity (98.8% in dermoscopy), but this came at the cost of sensitivity, with nearly half of melanomas missed.(46)

The AK versus SCC study provided further nuance. GPT-4o significantly outperformed Gemini overall, achieving 80.5% accuracy compared with 66.9% for Gemini. Yet both models demonstrated low negative predictive values, meaning a "benign" output could not reliably exclude SCC.(46) This limitation is clinically critical, as false negatives carry the greatest risk in oncology by delaying treatment for invasive disease. These results collectively show that while LLMs can achieve dermatologist-level performance in certain contexts, their diagnostic trade-offs mirror real-world clinical dilemmas: prioritizing sensitivity risks overdiagnosis, whereas prioritizing specificity risks missed malignancy.

Importantly, both studies highlighted the underrepresentation of darker skin types in datasets, raising concerns about equitable performance. Broader, multi-ethnic datasets are urgently needed to ensure that LLM-driven tools do not exacerbate disparities in skin cancer outcomes.

The results in acne, rosacea, and HS further emphasized the accessibility and promise of LLMs in common but often under-recognized inflammatory diseases. Acne and rosacea are ubiquitous, yet their clinical overlap can confuse patients and non-specialists. ChatGPT-4o demonstrated 93% overall accuracy, with nearly perfect sensitivity and specificity for primary diagnoses of both conditions. This suggests that LLMs may serve as highly reliable first-line tools for patients seeking clarity prior to specialist consultation. However, subtyping accuracy was substantially lower, with frequent misclassification between inflammatory and nodulocystic acne or between erythematotelangiectatic and papulopustular rosacea. Since treatment regimens often depend on subtype, this limitation tempers the immediate clinical utility of LLMs in guiding therapy.(47)

The HS study revealed a similar pattern. GPT-4o achieved strong diagnostic accuracy (87.3%) and high reproducibility, consistently outperforming Gemini and Claude Sonnet 3.7. Yet staging accuracy lagged, with common misclassification between Hurley stages II and III. Given that staging directly informs treatment escalation, from topical therapy to systemic biologics, these errors could meaningfully affect management.(48) Nevertheless, GPT-4o provided treatment recommendations consistent with clinical guidelines in nearly 90% of cases, underscoring its potential as an adjunctive decision-support system. The reproducibility of GPT-4o across repeated prompts also highlights progress toward reliability, an essential prerequisite for clinical adoption.

When considered together, imaging and LLMs demonstrate clear complementarity. Imaging modalities such as DG-HFUS, LC-OCT, and RCM provide objective, reproducible insights into skin structure and pathology at depth and resolution levels unattainable by gross inspection. These tools excel in margin delineation, early detection of subclinical spread, and longitudinal monitoring of carcinogenesis. By contrast, LLMs provide interpretive support that is scalable and accessible, requiring only a consumer device and a photograph. They democratize access to dermatologic expertise, particularly in regions where specialist care is limited, and empower patients to engage in their own health decisions.

The synergy between these approaches is promising. Imaging can generate structured, high-quality datasets to train LLMs, while LLMs can assist in interpreting imaging outputs or triaging cases for further evaluation. Hybrid workflows may ultimately allow clinicians to combine depth-resolved imaging with AI-driven decision support, balancing precision with scalability.

Despite encouraging results, several limitations merit discussion. Imaging studies remain resource-intensive, requiring expensive equipment, operator expertise, and significant time investment, which limit widespread implementation outside academic or tertiary centers. Meanwhile, the LLM studies faced constraints in dataset diversity, particularly with respect to skin phototypes, age ranges, and rare variants. The relatively small sample sizes in some cohorts (e.g., acne and rosacea, HS) limit generalizability. Furthermore, while histopathology was consistently used as the gold standard for malignant lesions, in

some studies (notably AK), diagnoses were supported by expert consensus and follow-up rather than universal biopsy, introducing potential bias.

Another limitation is the “black box” nature of LLMs. While outputs are expressed in natural language, the underlying decision processes remain opaque, complicating interpretability and trust. Susceptibility to prompt phrasing, variability in outputs, and the risk of hallucinations further underscore the need for careful clinical oversight.

Future research should pursue several directions. First, larger multicenter trials across diverse populations are needed to validate both imaging modalities and LLMs in real-world settings. Second, integration of imaging and AI should be explored, leveraging the strengths of each modality. For example, imaging outputs could be standardized into datasets that train LLMs specifically for dermatology, improving their performance and generalizability. Third, clinical pathways should be developed to define appropriate roles for these tools, whether as triage systems for primary care, intraoperative adjuncts for dermatologic surgeons, or educational aids for patients.

Clinically, the implications are significant. DG-HFUS and LC-OCT may reduce recurrence rates and operative times in BCC, improving both outcomes and patient satisfaction. LLMs, though not yet safe for stand-alone use, may soon provide accessible triage, expanding dermatology’s reach into underserved areas and preparing physicians to encounter patients who have already consulted “Dr. LLM.”

6. Conclusions

In summary, our series of studies underscores a dual trajectory of innovation in dermatology: precision-driven imaging and accessibility-driven AI. These two approaches address complementary aspects of the diagnostic landscape. Imaging technologies such as DG-HFUS, LC-OCT, and RCM provide unparalleled depth, objectivity, and reproducibility, enabling clinicians to visualize tumor margins, assess field cancerization, and characterize subtle morphologic alterations with high accuracy. These modalities not only improve surgical precision and reduce recurrence but also expand our biological understanding of cutaneous carcinogenesis.

At the same time, LLMs represent a parallel frontier in accessibility and scalability. By delivering dermatologist-level diagnostic support across malignant, premalignant, and inflammatory conditions, LLMs democratize expertise, reaching patients and providers in settings where dermatologic care is scarce. Our findings highlight that models such as ChatGPT-4o can achieve high sensitivity and reproducibility, making them valuable tools for triage and patient education, while Gemini excels in specificity, offering reassurance in confirmatory contexts. However, both systems face important limitations in negative predictive value, subtyping, and staging, underscoring that they are not yet safe for independent diagnostic use.

Taken together, these trajectories reveal that the future of dermatology will not be defined by a choice between imaging or AI, but rather by their integration. Imaging provides the structural precision required for oncologic safety and mechanistic insights, while AI extends interpretive support to the point of care, empowering both clinicians and patients. Their combination offers the possibility of hybrid workflows in which depth-resolved imaging feeds into AI-driven interpretive frameworks, yielding earlier detection, optimized treatment planning, and more equitable access to expertise.

While challenges remain, including cost, dataset diversity, interpretability, and clinical integration, the progress demonstrated in these studies provides a compelling foundation. As imaging devices become more portable and LLMs more specialized and transparent, their convergence is likely to reshape the practice of dermatology. Ultimately, the goal is a diagnostic paradigm where dermatologists are supported by advanced tools that

enhance, rather than replace, clinical judgment, delivering care that is earlier, more accurate, and more accessible for patients worldwide.

7. Summary

This body of work demonstrates how advanced imaging techniques and LLMs can address longstanding diagnostic and management challenges in dermatology. Through a series of complementary studies, we showed that DG-HFUS can reliably delineate BCC margins, reducing the risk of incomplete excision while conserving healthy tissue.

Parallel investigations into LLMs revealed their capacity to achieve dermatologist-level performance in selected diagnostic tasks. ChatGPT-4o consistently demonstrated high sensitivity for melanoma, SCC, and HS, while Gemini 2.0 Flash offered high specificity. Across acne, rosacea, and other common inflammatory dermatoses, LLMs delivered accurate primary diagnoses but were less reliable for subtyping or staging. Importantly, negative predictive values remained suboptimal, underscoring that current models cannot safely rule out malignancy when used in isolation.

Taken together, these studies highlight two converging trajectories of innovation: imaging modalities that enhance objectivity and precision in cancer surgery and field detection, and AI systems that broaden access to diagnostic support in both oncology and inflammatory disease. While neither approach is yet sufficient to replace dermatologists, their integration offers a path toward earlier detection, optimized treatment planning, and expanded access to expert-level care. Future research should focus on multicenter validation, dataset diversification, and hybrid workflows that combine imaging-based precision with AI-driven accessibility.

In conclusion, the findings from this research emphasize that the future of dermatology will be shaped by synergistic advances in imaging and AI, with the ultimate goal of delivering more accurate, efficient, and equitable care for patients worldwide.

8. References

1. Karimkhani C, Dellavalle RP, Coffeng LE, Flohr C, Hay RJ, Langan SM, Nsoesie EO, Ferrari AJ, Erskine HE, Silverberg JI. (2017) Global skin disease morbidity and mortality: an update from the global burden of disease study 2013. JAMA dermatology, 153: 406–412.
2. Chen JT, Kempton SJ, Rao VK. (2016) The economics of skin cancer: an analysis of Medicare payment data. Plastic and Reconstructive Surgery Global Open, 4.
3. Guy GP, Ekwueme DU. (2011) Years of potential life lost and indirect costs of melanoma and non-melanoma skin cancer: a systematic review of the literature. Pharmacoeconomics, 29: 863–874.
4. Hah JM, Lee E, Shrestha R, Pirrotta L, Huddleston J, Goodman S, Amanatullah DF, Dirbas FM, Carroll IR, Schofield D. (2021) Return to work and productivity loss after surgery: A health economic evaluation. International Journal of Surgery, 95: 106100.
5. Kearney M, Thokagevistk K, Boutmy E, Bharmal M. (2018) Treatment patterns, comorbidities, healthcare resource use, and associated costs by line of chemotherapy and level of comorbidity in patients with newly-diagnosed Merkel cell carcinoma in the United States. Journal of Medical Economics, 21: 1159–1171.
6. Martin M, Mohsin N, Reed D, Brownell I. (2022) Understanding Socioeconomic and Racial Disparities in Merkel Cell Carcinoma. Journal of Investigative Dermatology, 142: 2839.
7. Whitehead SJ, Ali S. (2010) Health outcomes in economic evaluation: the QALY and utilities. British Medical Bulletin, 96: 5–21.
8. Wright KN, Jonsdottir GM, Jorgensen S, Shah N, Einarsson JI. (2012) Costs and outcomes of abdominal, vaginal, laparoscopic and robotic hysterectomies. Jsls, 16: 519–524.
9. Yu C, Xian Y, Jing T, Bai M, Li X, Li J, Liang H, Yu G, Zhang Z. (2023) More patient-centered care, better healthcare: the association between patient-centered care and healthcare outcomes in inpatients. Front Public Health, 11: 1148277.
10. Albanesi C, Scarponi C, Giustizieri ML, Girolomoni G. (2005) Keratinocytes in inflammatory skin diseases. Current Drug Targets-Inflammation & Allergy, 4: 329–334.
11. Sina B, Kao GF, Deng AC, Gaspari AA. (2009) Skin biopsy for inflammatory and common neoplastic skin diseases: optimum time, best location and preferred techniques. A critical review. Journal of cutaneous pathology, 36: 505–510.
12. Harvey NT, Chan J, Wood BA. (2017) Skin biopsy in the diagnosis of neoplastic skin disease. Australian Family Physician, 46: 289–294.
13. Ahmed B, Qadir MI, Ghafoor S. (2020) Malignant melanoma: skin cancer− diagnosis, prevention, and treatment. Critical Reviews™ in Eukaryotic Gene Expression, 30.
14. Rigel DS, Carucci JA. (2000) Malignant melanoma: prevention, early detection, and treatment in the 21st century. CA: a cancer journal for clinicians, 50: 215–236.

15. Kibbi N, Kluger H, Choi JN. (2016) Melanoma: clinical presentations. Melanoma: 107–129.
16. Cabrera R, Recule F. (2018) Unusual clinical presentations of malignant melanoma: a review of clinical and histologic features with special emphasis on dermatoscopic findings. American journal of clinical dermatology, 19: 15.
17. Schmitt J, Seidler A, Diepgen TL, Bauer A. (2011) Occupational ultraviolet light exposure increases the risk for the development of cutaneous squamous cell carcinoma: a systematic review and meta-analysis. British Journal of Dermatology, 164: 291–307.
18. Work G, Invited R, Alam M, Armstrong A, Baum C, Bordeaux JS, Brown M, Busam KJ, Eisen DB, Iyengar V, Lober C, Margolis DJ, Messina J, Miller A, Miller S, Mostow E, Mowad C, Nehal K, Schmitt-Burr K, Sekulic A, Storrs P, Teng J, Yu S, Huang C, Boyer K, Begolka WS, Bichakjian C, Kim JYS, Kozlow JH, Mittal B, Moyer J, Olenecki T, Rodgers P. (2018) Guidelines of care for the management of cutaneous squamous cell carcinoma. Journal of the American Academy of Dermatology, 78: 560–578.
19. Stratigos A, Garbe C, Lebbe C, Malvehy J, Del Marmol V, Pehamberger H, Peris K, Becker JC, Zalaudek I, Saiag P. (2015) Diagnosis and treatment of invasive squamous cell carcinoma of the skin: European consensus-based interdisciplinary guideline. European journal of cancer, 51: 1989–2007.
20. Figueras Nart I, Cerio R, Dirschka T, Dréno B, Lear J, Pellacani G, Peris K, Ruiz de Casas A, Group PEiAW. (2018) Defining the actinic keratosis field: a literature review and discussion. Journal of the European Academy of Dermatology and Venereology, 32: 544–563.
21. Eisen DB, Asgari MM, Bennett DD, Connolly SM, Dellavalle RP, Freeman EE, Goldenberg G, Leffell DJ, Peschin S, Sligh JE, Wu PA, Frazer-Green L, Malik S, Schlesinger TE. (2021) Guidelines of care for the management of actinic keratosis: Executive summary. Journal of the American Academy of Dermatology, 85: 945–955.
22. Werner RN, Sammain A, Erdmann R, Hartmann V, Stockfleth E, Nast A. (2013) The natural history of actinic keratosis: a systematic review. British Journal of Dermatology, 169: 502–518.
23. Moy RL. (2000) Clinical presentation of actinic keratoses and squamous cell carcinoma. Journal of the American Academy of Dermatology, 42: S8–S10.
24. Dessinioti C, Antoniou C, Katsambas A, Stratigos AJ. (2010) Basal cell carcinoma: what's new under the sun. Photochemistry and photobiology, 86: 481–491.
25. Tanese K. (2019) Diagnosis and management of basal cell carcinoma. Current treatment options in oncology, 20: 13.
26. Niculet E, Craescu M, Rebegea L, Bobeica C, Nastase F, Lupasteanu G, Stan DJ, Chioncel V, Anghel L, Lungu M. (2022) Basal cell carcinoma: Comprehensive clinical and histopathological aspects, novel imaging tools and therapeutic approaches. Experimental and therapeutic medicine, 23: 60.
27. Nasr I, McGrath EJ, Harwood CA, Botting J, Buckley P, Budny PG, Fairbrother P, Fife K, Gupta G, Hashme M, Hoey S, Lear JT, Mallipeddi R, Mallon E, Motley RJ, Newlands C, Newman J, Pynn EV, Shroff N, Slater DN, Exton LS, Mohd Mustapa MF, Ezejimofor MC, Unit obotBAoDCS. (2021) British

Association of Dermatologists guidelines for the management of adults with basal cell carcinoma 2021*. British Journal of Dermatology, 185: 899–920.
28. Agero A, Cuevas J, Jaen P, Marghoob AA, Gill M, González S. (2008) Basal cell carcinoma. Reflectance Confocal Microscopy of Cutaneous Tumors. Informa: 60–75.
29. Altamura D, Menzies SW, Argenziano G, Zalaudek I, Soyer HP, Sera F, Avramidis M, DeAmbrosis K, Fargnoli MC, Peris K. (2010) Dermatoscopy of basal cell carcinoma: morphologic variability of global and local features and accuracy of diagnosis. Journal of the American Academy of Dermatology, 62: 67–75.
30. Apalla Z, Lallas A, Tzellos T, Sidiropoulos T, Lefaki I, Trakatelli M, Sotiriou E, Lazaridou E, Evangelou G, Patsatsi A. (2014) Applicability of dermoscopy for evaluation of patients' response to nonablative therapies for the treatment of superficial basal cell carcinoma. British Journal of Dermatology, 170: 809–815.
31. Cho M, Gordon L, Rembielak A, Woo T. (2014) Utility of radiotherapy for treatment of basal cell carcinoma: a review. British Journal of Dermatology, 171: 968–973.
32. Clark CM, Furniss M, Mackay-Wiggan JM. (2014) Basal cell carcinoma: an evidence-based treatment update. American journal of clinical dermatology, 15: 197–216.
33. Firnhaber JM. (2012) Diagnosis and treatment of basal cell and squamous cell carcinoma. American Family Physician, 86: 161–168.
34. Luz FB, Ferron C, Cardoso GP. (2015) Surgical treatment of basal cell carcinoma: an algorithm based on the literature. Anais brasileiros de dermatologia, 90: 377–383.
35. Morton C, Szeimies RM, Sidoroff A, Braathen L. (2013) European guidelines for topical photodynamic therapy part 1: treatment delivery and current indications–actinic keratoses, Bowen's disease, basal cell carcinoma. Journal of the European Academy of Dermatology and Venereology, 27: 536–544.
36. Mosterd K, Krekels GAM, Nieman FHM, Ostertag JU, Essers BAB, Dirksen CD, Steijlen PM, Vermeulen A, Neumann HAM, Kelleners-Smeets NWJ. (2008) Surgical excision versus Mohs' micrographic surgery for primary and recurrent basal-cell carcinoma of the face: a prospective randomised controlled trial with 5-years' follow-up. The Lancet Oncology, 9: 1149–1156.
37. Awosika O, Eleryan MG, Soleymani T. (2021) Reply to “Local recurrence of clinically observed basal cell carcinomas following complete saucerization or punch removal with negative margins: Retrospective case series from 2010 to 2020&#x201d. Journal of the American Academy of Dermatology, 85: e293–e294.
38. Kuiper EM, van den Berge BA, Spoo JR, Kuiper J, Terra JB. (2018) Low recurrence rate of head and neck basal cell carcinoma treated with Mohs micrographic surgery: a retrospective study of 1021 cases. Clinical Otolaryngology, 43: 1321–1327.
39. van Loo E, Mosterd K, Krekels GAM, Roozeboom MH, Ostertag JU, Dirksen CD, Steijlen PM, Neumann HAM, Nelemans PJ, Kelleners-Smeets NWJ. (2014) Surgical excision versus Mohs' micrographic surgery for basal cell carcinoma of

the face: A randomised clinical trial with 10year follow-up. European Journal of Cancer, 50: 3011–3020.
40. (2020) Service Guidance and Standards For Mohs Micrographic Surgery (MMS)
41. Bialy TL, Whalen J, Veledar E, Lafreniere D, Spiro J, Chartier T, Chen SC. (2004) Mohs micrographic surgery vs traditional surgical excision: a cost comparison analysis. Archives of dermatology, 140: 736–742.
42. Bittner GC, Cerci FB, Kubo EM, Tolkachjov SN. (2021) Mohs micrographic surgery: a review of indications, technique, outcomes, and considerations. Anais Brasileiros de Dermatologia, 96: 263–277.
43. Golda N, Hruza G. (2023) Mohs micrographic surgery. Dermatol Clin, 41: 39–47.
44. Cook J, Zitelli JA. (1998) Mohs micrographic surgery: a cost analysis. Journal of the American Academy of Dermatology, 39: 698–703.
45. Flohil SC, van Dorst AMJM, Nijsten T, Martino Neumann HA, Munte K. (2013) Mohs micrographic surgery for basal cell carcinomas: appropriateness of 'Rotterdam' criteria and predictive factors for three or more stages. Journal of the European Academy of Dermatology and Venereology, 27: 1228–1235.
46. 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, n/a.
47. 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, n/a.
48. 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, n/a.
49. Tuchayi SM, Makrantonaki E, Ganceviciene R, Dessinioti C, Feldman SR, Zouboulis CC. (2015) Acne vulgaris. Nature reviews Disease primers, 1: 1–20.
50. Leung AK, Barankin B, Lam JM, Leong KF, Hon KL. (2021) Dermatology: how to manage acne vulgaris. Drugs in context, 10.
51. Reynolds RV, Yeung H, Cheng CE, Cook-Bolden F, Desai SR, Druby KM, Freeman EE, Keri JE, Stein Gold LF, Tan JKL, Tollefson MM, Weiss JS, Wu PA, Zaenglein AL, Han JM, Barbieri JS. (2024) Guidelines of care for the management of acne vulgaris. Journal of the American Academy of Dermatology, 90: 1006.e1001–1006.e1030.
52. Sharma A, Kroumpouzos G, Kassir M, Galadari H, Goren A, Grabbe S, Goldust M. (2022) Rosacea management: A comprehensive review. Journal of Cosmetic Dermatology, 21: 1895–1904.
53. Condrò G, Guerini M, Castello M, Perugini P. (2022) Acne vulgaris, atopic dermatitis and rosacea: The Role of the Skin Microbiota—A Review. Biomedicines, 10: 2523.

54. Bjordal JM, Lopes-Martins RAB, Joensen J, Iversen VV. (2010) The anti-inflammatory mechanism of low level laser therapy and its relevance for clinical use in physiotherapy. Physical Therapy Reviews, 15: 286–293.
55. Margesson LJ, Danby FW. (2014) Hidradenitis suppurativa. Best practice & research Clinical obstetrics & gynaecology, 28: 1013–1027.
56. Smith HS, Chao JD, Teitelbaum J. (2010) Painful hidradenitis suppurativa. The Clinical journal of pain, 26: 435–444.
57. Kokolakis G, Wolk K, Schneider-Burrus S, Kalus S, Barbus S, Gomis-Kleindienst S, Sabat R. (2020) Delayed diagnosis of hidradenitis suppurativa and its effect on patients and healthcare system. Dermatology, 236: 421–430.
58. Rondags A, van Straalen KR, van Hasselt JR, Janse IC, Ardon CB, Vossen AR, Prens EP, van der Zee HH, Horváth B. (2019) Correlation of the refined Hurley classification for hidradenitis suppurativa with patient-reported quality of life and objective disease severity assessment. British Journal of Dermatology, 180: 1214–1220.
59. Thorlacius L, Garg A, Riis PT, Nielsen SM, Bettoli V, Ingram JR, Del Marmol V, Matusiak L, Pascual JC, Revuz J. (2019) Inter-rater agreement and reliability of outcome measurement instruments and staging systems used in hidradenitis suppurativa. British Journal of Dermatology, 181: 483–491.
60. Kirby J, Kim K, Zivkovic M, Wang S, Garg V, Danavar A, Li C, Chen N, Garg A. (2024) Uncovering the burden of hidradenitis suppurativa misdiagnosis and underdiagnosis: a machine learning approach. Frontiers in Medical Technology, 6: 1200400.
61. Cerminara SE, Cheng P, Kostner L, Huber S, Kunz M, Maul J-T, Böhm JS, Dettwiler CF, Geser A, Jakopović C. (2023) Diagnostic performance of augmented intelligence with 2D and 3D total body photography and convolutional neural networks in a high-risk population for melanoma under real-world conditions: A new era of skin cancer screening? European Journal of Cancer, 190: 112954.
62. Dengel LT, Petroni GR, Judge J, Chen D, Acton ST, Schroen AT, Slingluff Jr CL. (2015) Total body photography for skin cancer screening. International journal of dermatology, 54: 1250–1254.
63. Hodgkinson T, Vereker C, Spencer A. (2025) Early detection of melanoma: evaluating the combined use of total body photography and sequential digital dermoscopic imaging in a national health service setting. Clinical and Experimental Dermatology, 50: 1406–1408.
64. Olson J, Sharp P, Goatman K, Prescott G, Scotland G, Fleming A, Philip S, Santiago C, Borooah S, Broadbent D, Chong V, Dodson P, Harding S, Leese G, Styles C, Swa K, Wharton H. (2013) Improving the economic value of photographic screening for optical coherence tomography-detectable macular oedema: a prospective, multicentre, UK study. Health Technol Assess, 17: 1–142.
65. Ankad BS, Smitha S, Koti VR. (2020) Basic science of dermoscopy. Clinical Dermatology Review, 4: 69–73.
66. Campos-do-Carmo G, Ramos-e-Silva M. (2008) Dermoscopy: basic concepts. International journal of dermatology, 47: 712–719.

67. Argalia G, Reginelli A, Molinelli E, Russo A, Michelucci A, Sechi A, Marzano AV, Desyatnikova S, Fogante M, Patanè V, Granieri G, Tagliati C, Rizzetto G, De Simoni E, Matteucci M, Candelora M, Lanza C, Ventura C, Carboni N, Esposito R, Esposito S, Paolinelli M, Esposto E, Lanni G, Lucidi Pressanti G, Giorgi C, Principi F, Rebonato A, Malinowska SP, Mlosek RK, Giuseppetti GM, Dini V, Romanelli M, Offidani A, Cappabianca S, Wortsman X, Simonetti O. (2025) High-Frequency and Ultra-High-Frequency Ultrasound in Dermatologic Diseases and Aesthetic Medicine. Medicina, 61: 220.
68. Barcaui EdO, Carvalho ACP, Valiante PM, Piñeiro-Maceira J, Barcaui CB. (2021) High-frequency (22-MHz) ultrasound for assessing the depth of basal cell carcinoma invasion. Skin Research and Technology, 27: 676–681.
69. Berson M, Gregoire J, Gens F, Rateau J, Jamet F, Vaillant L, Tranquart F, Pourcelot L. (1999) High frequency (20 MHz) ultrasonic devices: advantages and applications. European journal of ultrasound, 10: 53–63.
70. Dinnes J, Bamber J, Chuchu N, Bayliss SE, Takwoingi Y, Davenport C, Godfrey K, O'Sullivan C, Matin RN, Deeks JJ, Williams HC. (2018) High-frequency ultrasound for diagnosing skin cancer in adults. Cochrane Database Syst Rev, 12: Cd013188.
71. Aumann S, Donner S, Fischer J, Müller F. (2019) Optical coherence tomography (OCT): principle and technical realization. High resolution imaging in microscopy and ophthalmology: new frontiers in biomedical optics: 59–85.
72. Nehal KS, Gareau D, Rajadhyaksha M. (2008) Skin imaging with reflectance confocal microscopy. In: Seminars in cutaneous medicine and surgery Vol. 27 pp. 37–43, WB Saunders
73. Brancaccio G, Balato A, Malvehy J, Puig S, Argenziano G, Kittler H. (2024) Artificial intelligence in skin cancer diagnosis: a reality check. Journal of Investigative Dermatology, 144: 492–499.
74. Du-Harpur X, Watt F, Luscombe N, Lynch M. (2020) What is AI? Applications of artificial intelligence to dermatology. British Journal of Dermatology, 183: 423–430.
75. Shifai N, van Doorn R, Malvehy J, Sangers TE. (2024) Can ChatGPT vision diagnose melanoma? An exploratory diagnostic accuracy study. Journal of the American Academy of Dermatology, 90: 1057–1059.
76. de Oliveira EC, da Motta VR, Pantoja PC, Ilha CSdO, Magalhaes RF, Galadari H, Leonardi GR. (2019) Actinic keratosis–review for clinical practice. International journal of dermatology, 58: 400–407.
77. Jeffes EW, Tang EH. (2000) Actinic keratosis: current treatment options. American journal of clinical dermatology, 1: 167–179.
78. Roh MR, Eliades P, Gupta S, Tsao H. (2015) Genetics of melanocytic nevi. Pigment cell & melanoma research, 28: 661–672.
79. Leachman SA, Cassidy PB, Chen SC, Curiel C, Geller A, Gareau D, Pellacani G, Grichnik JM, Malvehy J, North J. (2015) Methods of melanoma detection. Melanoma: 51–105.
80. Tan E, Oakley A, Soyer H, Haskett M, Marghoob A, Jameson M, Rademaker M. (2010) Interobserver variability of teledermoscopy: an international study. British Journal of Dermatology, 163: 1276–1281.

81. Picardo M, Eichenfield LF, Tan J. (2017) Acne and rosacea. Dermatology and therapy, 7: 43–52.
82. Crawford GH, Pelle MT, James WD. (2004) Rosacea: I. Etiology, pathogenesis, and subtype classification. Journal of the American Academy of Dermatology, 51: 327–341.
83. Holmes AD, Steinhoff M. (2017) Integrative concepts of rosacea pathophysiology, clinical presentation and new therapeutics. Experimental dermatology, 26: 659–667.
84. Tan J, Berg M. (2013) Rosacea: current state of epidemiology. Journal of the American Academy of Dermatology, 69: S27–S35.
85. Amat-Samaranch V, Agut-Busquet E, Vilarrasa E, Puig L. (2021) New perspectives on the treatment of hidradenitis suppurativa. Therapeutic Advances in Chronic Disease, 12: 20406223211055920.
86. Zouboulis CC, Bechara F, Benhadou F, Bettoli V, Bukvić Mokos Z, Del Marmol V, Dolenc-Voljč M, Giamarellos-Bourboulis EJ, Grimstad Ø, Guillem P. (2024) European S2k guidelines for hidradenitis suppurativa/acne inversa part 2: Treatment. Journal of the European Academy of Dermatology and Venereology.
87. Zouboulis CC, Del Marmol V, Mrowietz U, Prens EP, Tzellos T, Jemec GB. (2015) Hidradenitis suppurativa/acne inversa: criteria for diagnosis, severity assessment, classification and disease evaluation. Dermatology, 231: 184–190.
88. Zouboulis CC, Tzellos T, Kyrgidis A, Jemec GBE, Bechara FG, Giamarellos-Bourboulis EJ, Ingram JR, Kanni T, Karagiannidis I, Martorell A, Matusiak Ł, Pinter A, Prens EP, Presser D, Schneider-Burrus S, von Stebut E, Szepietowski JC, van der Zee HH, Wilden SM, Sabat R, Group obotEHSFI, Sinclair R, Vekic D, de Haes P, del Marmol V, Alavi A, Gulliver W, Nassif A, Nikolakis G, Podda M, Antoniou C, Delli FS, Dessinioti C, Patsatsi A, Trigoni AP, Zisimou C, Bettoli V, Garcovich S, Boer J, Horváth B, Ingvarsson G, Puig L, Hunger RE, Emtestam L, Danby FW, Kimball AB. (2017) Development and validation of the International Hidradenitis Suppurativa Severity Score System (IHS4), a novel dynamic scoring system to assess HS severity. British Journal of Dermatology, 177: 1401–1409.
89. Ovadja Z, Schuit M, van der Horst C, Lapid O. (2019) Inter-and intrarater reliability of Hurley staging for hidradenitis suppurativa. British Journal of Dermatology, 181: 344–349.
90. Hoffmann J, Borgeaud S, Mensch A, Buchatskaya E, Cai T, Rutherford E, de Las Casas D, Hendricks LA, Welbl J, Clark A. (2022) An empirical analysis of compute-optimal large language model training. Advances in neural information processing systems, 35: 30016–30030.
91. Lin C, Kuo C-F. (2025) Roles and potential of Large language models in healthcare: A comprehensive review. Biomedical Journal, 48: 100868.
92. Zhang A, Zhao E, Wang R, Zhang X, Wang J, Chen E. (2025) Multimodal large language models for medical image diagnosis: Challenges and opportunities. Journal of Biomedical Informatics, 169: 104895.
93. Torop M, Eskandar M, Kurtansky N, Liu J, Weber J, Camps O, Rotemberg V, Dy J, Kose K. (2025) Grounding Multimodal Large Language Models with

Quantitative Skin Attributes: A Retrieval Study. arXiv preprint arXiv:2508.20188.
94. Urooj B, Fayaz M, Ali S, Dang LM, Kim KW. (2025) Large Language Models in Medical Image Analysis: A Systematic Survey and Future Directions. Bioengineering, 12: 818.
95. Ullah E, Parwani A, Baig MM, Singh R. (2024) Challenges and barriers of using large language models (LLM) such as ChatGPT for diagnostic medicine with a focus on digital pathology–a recent scoping review. Diagnostic pathology, 19: 43.
96. Roky AH, Islam MM, Ahasan AMF, Mostaq MS, Mahmud MZ, Amin MN, Mahmud MA. (2025) Overview of skin cancer types and prevalence rates across continents. Cancer pathogenesis and therapy, 3: 89–100.
97. Cameron MC, Lee E, Hibler BP, Barker CA, Mori S, Cordova M, Nehal KS, Rossi AM. (2019) Basal cell carcinoma: Epidemiology; pathophysiology; clinical and histological subtypes; and disease associations. Journal of the American Academy of Dermatology, 80: 303–317.
98. Aleissa S, Navarrete-Dechent C, Cordova M, Sahu A, Dusza SW, Phillips W, Rossi A, Lee E, Nehal KS. (2020) Presurgical evaluation of basal cell carcinoma using combined reflectance confocal microscopy–optical coherence tomography: A prospective study. Journal of the American Academy of Dermatology, 82: 962–968.
99. Basset-Seguin N, Herms F. (2020) Update on the management of basal cell carcinoma. Acta Dermato-Venereologica, 100.
100. Caresana G, Giardini R. (2010) Dermoscopy-guided surgery in basal cell carcinoma. Journal of the European Academy of Dermatology and Venereology, 24: 1395–1399.
101. Boostani M, Wortsman X, Pellacani G, Cantisani C, Suppa M, Mohos A, Goldust M, Pietkiewicz P, Bánvölgyi A, Holló P. (2025) Dermoscopy-guided high-frequency ultrasound for preoperative assessment of basal cell carcinoma lateral margins: A pilot study. The British journal of dermatology: ljaf209.
102. Boostani M, Bozsányi S, Suppa M, Cantisani C, Lőrincz K, Bánvölgyi A, Holló P, Wikonkál NM, Huss WJ, Brady KL, Paragh G, Kiss N. (2025) 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, 64: 287–301.

9. Bibliography of the candidate's publications

Publications related to the thesis (Σ IF: 29.2):

- **Mehdi Boostani**, Ximena Wortsman, Giovanni Pellacani, Carmen Cantisani, Mariano Suppa, Anita Mohos, Mohamad Goldust, Pawel Pietkiewicz, András Bánvölgyi, Péter Holló, Norbert M Wikonkál, Kamran Avanaki, Gyorgy Paragh, Kende Lőrincz, Norbert Kiss, Dermoscopy-guided high-frequency ultrasound for preoperative assessment of basal cell carcinoma lateral margins: A pilot study,

British Journal of Dermatology, 2025;, ljaf209, https://doi.org/10.1093/bjd/ljaf209

- **Mehdi Boostani**, Aimilios Lallas, Mohamad Goldust, Nóra Nádudvari, Kende Lőrincz, András Bánvölgyi, Péter Holló, Norbert M Wikonkál, Gyorgy Paragh, Norbert Kiss "Diagnostic performance of multimodal large language models in distinguishing melanoma from nevi in clinical and dermoscopic images", 2025, JAAD International, https://doi.org/10.1016/j.jdin.2025.08.008
- **Mehdi Boostani**, Giovanni Pellacani, Mohamad Goldust, Nóra Nádudvari, Dóra Rátky, Carmen Cantisani, Kende Lőrincz, András Bánvölgyi, Norbert M Wikonkál, Gyorgy Paragh, Norbert Kiss " Diagnosing Actinic Keratosis and Squamous Cell Carcinoma With Large Language Models From Clinical Images," International Journal of Dermatology (2025): 1–3, https://doi.org/10.1111/ijd.18000.
- **Mehdi Boostani**, András Bánvölgyi, Mohamad Goldust, Carmen Cantisani, Paweł Pietkiewicz, Kende Lőrincz, Péter Holló, Norbert M Wikonkál, Gyorgy Paragh, Norbert Kiss, " (2025) Diagnostic Performance of GPT-4o and Gemini Flash 2.0 in Acne and Rosacea. Int J Dermatol. https://doi.org/10.1111/ijd.17729
- **Mehdi Boostani**, András Bánvölgyi, Christos C Zouboulis, Noah Goldfarb, Mariano Suppa, Mohamad Goldust, Kende Lőrincz, Tara Kiss, Nóra Nádudvari, Péter Holló, Norbert M Wikonkál, Gyorgy Paragh, Norbert Kiss (2025), Large language models in evaluating hidradenitis suppurativa from clinical images. Journal of the European Academy of Dermatology. https://doi.org/10.1111/jdv.20861.

Publications not related to this thesis (Σ IF: 37.4)**:**

- Norbert Kiss, **Mehdi Boostani**, "Ultra-wideband raster-scanning optoacoustic mesoscopy as a histologic-grade tool for melanoma depth assessment", 2025, Journal of the European Academy of Dermatology https://doi.org/10.1111/jdv.20895

- Carmen Cantisani, Phyllida Kerstin Hamilton-Maikle, Mattia Di Segni, Andras Banvolgyi, Giovanni Paolino, Salvador Gonzales, **Mehdi Boostani**, Giovanni Pellacani, Gyorgy Paragh, Norbert Kiss, "Reddish/bluish benign lesions evaluated by dermus SkinScanner", Journal of Ultrasound, 2025, https://doi.org/10.1007/s40477-025-01066-z
- **Mehdi Boostani**, Giovanni Pellacani, Ximena Wortsman, Mariano Suppa, Mohamad Goldust, Carmen Cantisani, Paweł Pietkiewicz, Kende Lőrincz, András Bánvölgyi, Norbert M Wikonkál, Wendy J Huss, Kamran Avanaki, Gyorgy Paragh, Norbert Kiss. "FDA and EMA-approved noninvasive imaging techniques for basal cell carcinoma subtyping: A systematic review." JAAD International (2025). https://doi.org/10.1016/j.jdin.2025.05.006
- András Bánvölgyi, Kende Lőrincz, **Mehdi Boostani**, Efrat Bar-Ilan, Bernadett Hidvégi, Márta Medvecz, Norbert Kiss, Norbert M Wikonkál "Anaphylaxis with Elevated Procalcitonin Mimicking Sepsis: A Literature Review and Report of Two Cases." Journal of Clinical Medicine 14.3 (2025): 785. https://doi.org/10.3390/jcm14030785
- Natália Czurkó, Katalin Martyin, **Mehdi Boostani**, Tünde Kerner, András Bánvölgyi, Enikő Kuroli, Márta Medvecz, Eszter Ostorházi, Norbert Kiss, (2025), Tinea Incognita Unmasked: Diagnostic Challenge in a Male Patient With a 3-Year History of Topical Steroid Use. Int J Dermatol. https://doi.org/10.1111/ijd.17720
- **Mehdi Boostani**, Szabolcs Bozsányi, Mariano Suppa, Carmen Cantisani, Kende Lőrincz, András Bánvölgyi, Péter Holló, Norbert M Wikonkál, Wendy J Huss, Kimberly L Brady, Gyorgy Paragh, Norbert Kiss (2024), Novel imaging techniques for tumor margin detection in basal cell carcinoma: a systematic scoping review of FDA and EMA-approved imaging modalities. Int J Dermatol. https://doi.org/10.1111/ijd.17496
- **Mehdi Boostani**, Phyllida Kerstin Hamilton-Meikle, Szabolcs Bozsányi, Klára Farkas, Boglárka Szabó, András Bánvölgyi, Kende Lőrincz, Péter Holló, Norbert Wikonkál, Huss Wendy, György Paragh, Norbert Kiss (2024). New imaging modalities for determining the tumor margins of basal cell carcinoma. Bőrgyógyászati és Venerológiai Szemle. 100. 125-130. https://doi.org/10.7188/bvsz.2024.100.3.3

- Natália Czurkó, Szabolcs Bozsányi, Fanni Meznerics, Eszter Szlávicz, Melinda Becske, **Mehdi Boostani**, Klára Farkas, Roland Kasek, Emese Fejes, András Bánvölgyi, Kende Lőrincz, Szabolcs Takács, Péter Holló, Norbert Wikonkál, Imre Lázár, Norbert Kiss, "Pszichodermatológiai faktorok szerepe a psoriasis hátterében." Bőrgyógyászati és Venerológiai Szemle 100, no. 1 (2024): 21-27. https://doi.org/10.7188/bvsz.2024.100.1
- Szabolcs Bozsányi*, **Mehdi Boostani***, Klára Farkas, Phyllida Hamilton-Meikle, Noémi Nóra Varga, Boglárka Szabó, Flóra Vasanits, Enikő Kuroli, Fanni Adél Meznerics, Kende Lőrincz, Péter Holló, András Bánvölgyi, Norbert M Wikonkál, Gyorgy Paragh, Norbert Kiss. Optically Guided High-Frequency Ultrasound to Differentiate High-Risk Basal Cell Carcinoma Subtypes: A Single-Centre Prospective Study. J Clin Med. 2023 Nov 3;12(21):6910. PMID: 37959375; PMCID: PMC10648659. https://doi.org/10.3390/jcm12216910
- Noémi Nóra Varga*, **Mehdi Boostani***, Klára Farkas, András Bánvölgyi, Kende Lőrincz, Máté Posta, Ilze Lihacova, Alexey Lihachev, Márta Medvecz, Péter Holló, Gyorgy Paragh, Norbert M Wikonkál, Szabolcs Bozsányi, Norbert Kiss. 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 Dec 28;16(1):157. PMID: 38201584; PMCID: PMC10778011. https://doi.org/10.3390/cancers16010157
- Carmen Cantisani, Alexandru-Vasile Baja, Luca Gargano, Giovanni Rossi, Marco Ardigò, Giuseppe Soda, **Mehdi Boostani**, Norbert Kiss, Giovanni Pellacani. Optical Coherence Tomography as a Valuable Tool for the Evaluation of Cutaneous Kaposi Sarcoma Treated with Imiquimod 5% Cream. Diagnostics (Basel). 2023 Sep 11;13(18):2901. PMID: 37761269; PMCID: PMC10529460. https://doi.org/10.3390/diagnostics13182901
- Szabolcs Bozsányi, Natália Czurkó, Melinda Becske, Roland Kasek, Botond Keve Lázár, **Mehdi Boostani**, Fanni Adél Meznerics, Klára Farkas, Noémi Nóra Varga, Lili Gulyás, András Bánvölgyi, Bence Ágoston Fehér, Emese Fejes, Kende Lőrincz, Anikó Kovács, Hunor Gergely, Szabolcs Takács, Péter Holló, Norbert Kiss, Norbert Wikonkál, Imre Lázár. Assessment of Frontal

Hemispherical Lateralization in Plaque Psoriasis and Atopic Dermatitis. J Clin Med. 2023 Jun 21;12(13):4194. PMID: 37445231; PMCID: PMC10342794. https://doi.org/10.3390/jcm12134194

10. Acknowledgements

I would like to express my deepest gratitude to those individuals whose support, mentorship, and guidance have been essential to my scientific journey. Foremost, I am sincerely thankful to Dr. Norbert Kiss, whose unwavering support, invaluable insights, and constructive feedback have been fundamental in shaping the direction of my research. His dedication to both science and patient care has been a source of constant inspiration and motivation.

I would also like to extend my heartfelt appreciation to Dr. György Paragh, whose mentorship and guidance have played a pivotal role in my professional development. His clinical expertise and dedication to advancing dermatologic research have been instrumental in the success of my projects.

I am grateful to Prof. Norbert M. Wikonkál, who provided me with the opportunity to work within an inspiring and collaborative academic environment. His leadership, enthusiasm for innovation, and commitment to excellence in dermatology created the conditions that allowed me to pursue these studies with confidence and purpose.

I would like to extend my special thanks to Prof. Péter Holló, Chair of our Department, for his unwavering support and encouragement. His guidance has been invaluable throughout my professional journey.

My sincere thanks go to Dr. András Bánvölgyi and Dr. Kende Lőrincz for their valuable advice, continuous encouragement, and collaborative spirit throughout my research. Their clinical expertise and thoughtful input have enriched my academic journey in numerous ways.

I also wish to express my gratitude to Dr. Szabolcs Bozsányi, whose support, guidance, and collegiality have been indispensable. His contributions to our collaborative work, along with his generosity in sharing knowledge, have left a meaningful mark on my development as a researcher.

My sincere gratitude also goes to out dermatopathologist, Dr. Anita Mohos for her continuous support, encouragement, and kindness, which have been deeply meaningful throughout my academic development.

I am truly thankful to Prof. Giovanni Pellacani, Chair of Dermatology at Sapienza University of Rome, and Dr. Cantisani for their generous support and inspiring leadership, which have greatly enriched my scientific perspective and provided invaluable opportunities for collaboration.

I would like to express my heartfelt appreciation to Dr. Mariano Suppa from Hôpital Erasme in Brussels, Belgium, for his insightful advice, scientific guidance, and generous support, which have played an important role in shaping my research experience.

My heartfelt thanks go as well to Dr. Mohamad Goldust from Yale, whose inspiring professional example, encouragement, and generous support have been greatly appreciated. His commitment to advancing dermatology and fostering international collaboration has been truly motivating.

I would like to express my sincere appreciation to the Dermus technical team for their outstanding technical assistance, professionalism, and dedication. Their expertise and commitment have been invaluable in supporting the practical and technological aspects of my research.

Finally, I would like to thank my colleagues, collaborators, friends, and family for their encouragement, understanding, and unwavering support throughout my academic path. Their presence has made this demanding but rewarding journey possible.