The BRAAFF-Annotated Acral Lesions Dataset (BALD): A Curated Set of Dermatoscopic Images of Acral Melanoma and Nevi from Various Sources

JID Open

Journal of Investigative Dermatology (2025) **145,** 1802−1805; doi:10.1016/j.jid.2024.12.021

TO THE EDITOR

The accuracy and generalizability of machine-learning models for diagnosing neoplastic skin lesions depend heavily on the composition of the training dataset. Although deep learning has achieved accuracy levels comparable with those of human experts, its performance often deteriorates in settings that differ from those of the original training conditions (Combalia et al, 2022). This underperformance is particularly prominent in cases involving rare diseases or atypical presentations of common conditions. For instance, in melanoma detection, underrepresented cases include nonpigmented lesions, pediatric melanomas, melanomas in patients with skin of color (Chang et al, 2024; Groh et al, 2024), and melanomas on specific anatomical sites. Acral melanomas are particularly important because they occur across all ethnicities but are underrepresented in databases such as the International Skin Imaging Collaboration archive (Braun et al, 2013; Saida et al, 2022).

To address this gap, we present a dermatoscopic image dataset that focuses on acral lesions (Müller et al, 2024). Our dataset comprised 666 retrospectively collected dermatoscopic images of acral lesions from 643 patients collected from 7 academic centers in Austria, Greece, France, Italy, and Japan, featuring 144 acral melanomas and 522 acral nevi. The images were obtained from 2 sources: the original BRAAFF checklist development dataset (Lallas et al, 2015), which consisted of 603 lesions (472 nevi and 131 melanomas), and a newer set from a validation study with 74 lesions (61 nevi and 13 melanomas). Images of 11 nevi were excluded because of poor quality (n = 8), incorrect location (n = 2), or postexcision photography (n = 1). All images featured lesions on glabrous skin, specifically on the palms or soles. The majority of the lesions were located on the soles (n = 544), with only 66 on the palms; the exact location was unavailable for the remaining images. The patient were of European (n = 381) and Asian (n = 285) ethnicities, although skin tone was not documented. The mean age of the patients was 43 years (SD = 22 years), and 61.1% of the lesions (n = 407) were from females. All the melanomas were excised and histopathologically examined. Of the 144 melanomas, 33.3% (n = 48) were in situ, and the rest were invasive, with a median invasion depth of 1.65 mm (range = 0.2−14.5 mm). Among the 522 nevi, 193 were excised; the remainder showed typical morphology or were monitored for at least 1 year to confirm their benign status. Because no interventions were conducted on the patients and only anonymized image data were utilized, patient consent was not necessary, as confirmed by the approval of the local ethics board.

The original images were manually cropped and resized to 600 × 450 pixels, saved in JPG format, and pseudonymized, with identifiable features removed if necessary. In some cases, the surrounding skin areas were blurred to obscure unique patterns, such as skin ridges, to prevent reidentification. Metadata included age (in 5-year intervals), sex, and diagnosis. Melanomas have additional data on their in situ status and invasion thickness. Metadata not deemed relevant for research and educational purposes were removed to comply with privacy regulations and data protection requirements.

Annotations were collected through an online reader study designed to re-evaluate the BRAAFF checklist for diagnosing acral melanoma. The raters were recruited through social media announcements made by the International Dermoscopy Society. The study received ethical approval from the Medical University of Vienna's Ethics Committee (EK1220/2022), and all participants provided informed consent for the anonymous use of their annotations in the research by agreeing to the terms of use on the Dermonaut website during registration. The study itself was conducted on the 'Dermachallenge' platform (a subsite of the Dermonaut website) and engaged 409 participants, including board-certified dermatologists (n = 114), dermatology residents (n = 195), general practitioners (n = 58), medical students (n = 9), and other healthcare professionals (n = 30). Dermatoscopy experience varied, with participants categorized as low-experience (n = 20), intermediate-experience (n = 310), or high-experience readers (n = 79). In each round, readers assessed batches of 10 lesions, identifying the presence or absence of the 6 BRAAFF checklist features and providing a melanoma or nevus diagnosis on the basis of pattern recognition.

The BRAAFF checklist includes 6 criteria: 4 malignant features (irregular

Accepted manuscript published online 17 January 2025; corrected proof published online 8 February 2025

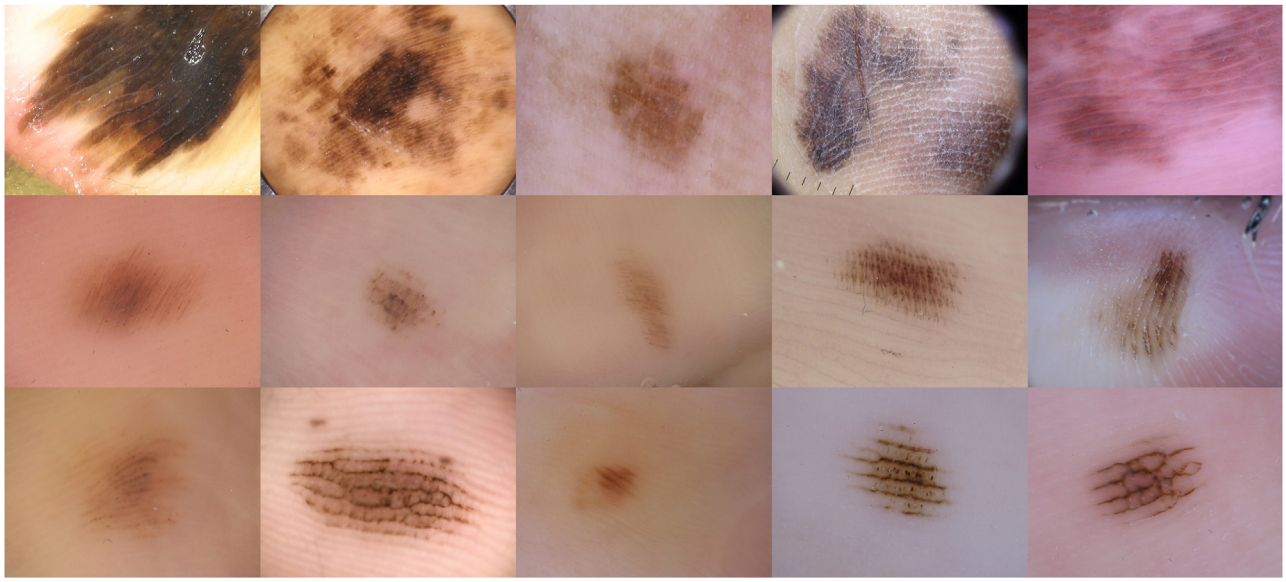

Figure 1. Top 5 examples with highest agreement for parallel lines patterns. Top 5 examples with the highest agreement for the parallel ridge pattern (top row), fibrillar pattern (middle row), and parallel furrow pattern (bottom row) are presented.

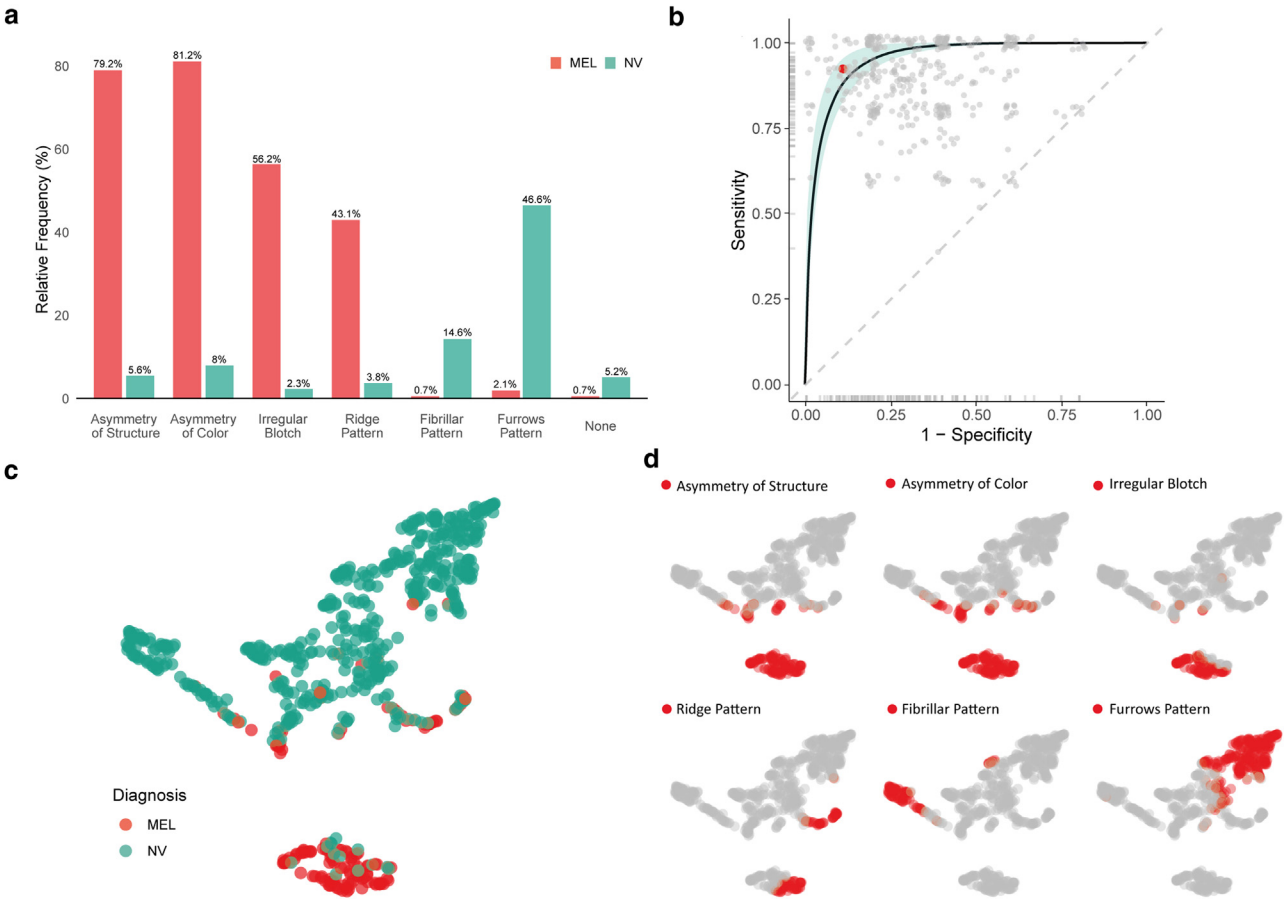

Figure 2. Main characteristics of the BRAAFF-annotated acral lesions dataset. (**a**) Relative frequencies of BRAAFF feature presence stratified by diagnosis type (MEL = melanoma, NV = nevus). Each feature was identified through majority voting among the readers. The bars represent the percentage of lesions exhibiting each BRAAFF feature. (**b**) ROC curve for the consensus BRAAFF score derived by majority voting. The green area represents the 95% confidence interval, and the red dots indicate the optimal cutoff score. Gray dots represent individual sensitivity–specificity pairs for readers (n = 409) based on pattern recognition diagnosis. (**c**) UMAP visualization of 666 acral lesions (144 melanomas and 522 nevi) based on BRAAFF feature annotations. Each point corresponds to a single lesion with spatial proximity, reflecting the similarity in BRAAFF features. Distinct clustering highlights the differential BRAAFF feature profiles of acral nevi and melanomas. (**d**) Distribution of 6 BRAAFF features within the UMAP plot in panel **c**. Each subplot corresponds to a specific BRAAFF feature, with red indicating the presence and gray indicating the absence of the feature, as determined by majority voting. ROC, receiver operating characteristic; UMAP, Uniform Manifold Approximation and Projection.

blotch, parallel ridge pattern, asymmetry of structures, and asymmetry of colors) and 2 benign features (parallel furrow pattern, including lattice-like and fibrillar patterns) (Lallas et al, 2015). The BRAAFF score was calculated by adding points for malignant features (3 points for the parallel ridge pattern and 1 point for each other malignant feature) and subtracting 1 point for each benign feature. A score of 1 or higher suggests melanoma. The BRAAFF score was calculated automatically in the background but was not displayed to the readers. The rounds with incomplete data (n = 240) or below-chance accuracy (n = 16) were excluded. In total, 409 participants contributed 15751 valid readings, with a median of 16 readings per lesion (range = 3−128).

For each lesion, annotations included the number of readings and the absolute and relative counts for the BRAAFF features. Figure 1 highlights the top 5 examples with the highest consensus for the parallel ridge, fibrillar, and parallel furrow patterns. A majority-based annotation was generated for each feature, considering it present if over 50% of the readers indicated it. Figure 2a shows the relative proportions of the majority consensus for the 6 features by diagnosis. The consensus-based BRAAFF score achieved a sensitivity of 92.4% (95% confidence interval = 87.5−96.0%) and a specificity of 89.1% (95% confidence interval = 86.2−91.8%), with an area under the curve of 0.96 (95% confidence interval = 0.94−0.97) (Figure 2b). The Uniform Manifold Approximation and Projection visualization in Figure 2c shows distinct clustering patterns in the feature distribution, where the acral melanomas formed separate regions from the nevi. This demonstrates the utility of consensus annotations and BRAAFF features in distinguishing acral melanomas from nevi. These clusters were based on the relative frequency of the BRAAFF feature, as indicated by multiple reader assessments (Figure 2d).

This annotated dataset offers a benchmark for human performance in differentiating acral melanomas from nevi in dermatoscopic images and provides a resource for training and validating machine learning algorithms (Daneshjou and Kittler, 2024). Notably, minimal overlap exists with the HAM 10000 dataset, which is frequently used in algorithm training and should be considered by users (Tschandl et al, 2018). Limitations include the exclusive use of dermatoscopic images, absence of clinical close-ups, and lack of representation of patients with very dark skin tones, which may limit generalizability (Groh et al, 2024).

DATA AVAILABILITY STATEMENT

The dataset and metadata are available on the International Skin Imaging Collaboration archive (https://doi.org/10.34970/669187), with extended metadata on GitHub (https://github.com/kittler/BALD), under a CC-BY-NC license for research and teaching purposes, with citation required.

KEYWORDS

Acral; Dermoscopy; Melanoma; Nevus

CONFLICT OF INTEREST

IZ is on the advisory board of and received fees from Fotofinder, Heine, Novartis, MSD, BMS, Philogen, Regeneon, and Sanofy Genzyme and is a consortium member of the project iToboS. HKi received speaker fees from Fotofinder, Eli Lilly, Novartis Equipment, Heine, Fotofinder, 3Gen, and Casio. ADS received consulting fees from Janssen-Cilag, Sun Pharma, and Sanofi and speaker fees from Pierre-Fabre, Galderma, Sun Pharma, and Sanofi. HKo received grants/contracts/speaker fee/equipment material from Casio Computer. PT received grants/contracts from Lilly; received speaker fees from Lilly, FotoFinder, and Novartis; is executive board member of International Dermoscopy Society; is the president of Austrian Dermatopathology Society; is a coauthor of Dermatoscopy Textbook (Facultas); and is a coauthor of Dermatopathology Textbook (Springer). The remaining authors state no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization: HKi, CM; Data Curation: CM, PT, CR, AK, HKo, EM, ZA, ADS, KK, EL, CL, AP, TS, ES, MT, LT, IZ, GA, AL, HKi; Formal Analysis: HKi, CM; Investigation: HKi, CM; Methodology: HKi; Project Administration: HKi, CM; Resources: PT, HKi, CR; Software: PT, HKi, CR; Supervision: HKi; Validation: HKi; Visualization: CM, HKi; Writing - Original Draft Preparation: CM, HKi; Writing - Review and Editing: CM, PT, CR, AK, HKo, EM, ZA, ADS, KK, EL, CL, AP, TS, ES, MT, LT, IZ, GA, AL, HKi

Christoph Müller[1,*], Philipp Tschandl[1], Christoph Rinner[2], Athanassios Kyrgidis[3], Hiroshi Koga[4], Elvira Moscarella[5], Zoe Apalla[6], Alessandro Di Stefani[7,8], Ken Kobayashi[9], Elisabeth Lazaridou[6], Caterina Longo[5], Alice Phan[10], Toshiaki Saida[4], Elena Sotiriou[11], Masaru Tanaka[12], Luc Thomas[13], Iris Zalaudek[14], Giuseppe Argenziano[15], Aimilios Lallas[11] and Harald Kittler[1]

[1]*Department of Dermatology, Medical University of Vienna, Vienna, Austria;* [2]*Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Vienna, Austria;* [3]*Department of Maxillofacial Surgery, Papanikolaou General Hospital of Thessaloniki, Aristotle University of Thessaloniki, Thessaloniki, Greece;* [4]*Department of Dermatology, Shinshu University Hospital, Matsumoto, Japan;* [5]*Department of Dermatology, University of Modena and Reggio Emilia, Modena, Italy;* [6]*Second Department of Dermatology, Aristotle University, Thessaloniki, Greece;* [7]*Dermatologia, Dipartimento Universitario di Medicina e Chirurgia Traslazionale, Università Cattolica del Sacro Cuore, Rome, Italy;* [8]*Dermatologia, Dipartimento di Scienze Mediche e Chirurgiche, Institute of Dermatology, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy;* [9]*Kobayashi Skin Care Clinic, Tokyo, Japan;* [10]*Nephrology-Rheumatology-Dermatology Department, Hôpital Femme Mère Enfant, Hospices Civils de Lyon, Bron, France;* [11]*First Department of Dermatology, School of Medicine, Aristotle University, Thessaloniki, Greece;* [12]*Department of Dermatology, Adachi Medical Center, Tokyo Women's Medical University, Adachi, Japan;* [13]*Service de Dermatologie, Centre Hospitalier Lyon Sud, Hospices Civils de Lyon, Lyon, France; Université Claude-Bernard-Lyon Lyon, Lyon, France;* [14]*Department of Dermatology, Medical University of Trieste, Trieste, Italy; and* [15]*Dermatology Unit, Department of Mental and Physical Health and Preventive Medicine, University of Campania Luigi Vanvitelli, Napoli, Italy*

[*]*Corresponding author e-mail: christoph.mueller@meduniwien.ac.at*

REFERENCES

Braun RP, Thomas L, Dusza SW, Gaide O, Menzies S, Dalle S, et al. Dermoscopy of acral melanoma: a multicenter study on behalf of the international dermoscopy society. Dermatology 2013;227:373−80.

Chang CT, Pathmarajah P, Allerup J, Jafry S, Yekrang K, Mitchell DC, et al. DDI-2: a diverse skin condition image dataset representing self-identified Asian patients. J Invest Dermatol 2025;145:1205−8.e4.

Combalia M, Codella N, Rotemberg V, Carrera C, Dusza S, Gutman D, et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge. Lancet Digit Health 2022;4:e330−9.

Daneshjou R, Kittler H. Empowering the next generation of artificial intelligence in dermatology: the datasets and benchmarks track of

the journal of investigative dermatology. J Invest Dermatol 2024;144:437–8.

Groh M, Badri O, Daneshjou R, Koochek A, Harris C, Soenksen LR, et al. Deep learning-aided decision support for diagnosis of skin disease across skin tones. Nat Med 2024;30:573–83.

Lallas A, Kyrgidis A, Koga H, Moscarella E, Tschandl P, Apalla Z, et al. The BRAAFF checklist: a new dermoscopic algorithm for diagnosing acral melanoma. Br J Dermatol 2015;173:1041–9.

Müller C, Kittler H, Tschandl P, Rinner C, Grausenburger ML, Kyrgidis A, et al. Validation of a dermatoscopy-based algorithm for the diagnosis of acral melanoma. Dermatology 2024;240:739–802.

Saida T, Koga H, Uhara H. Dermoscopy for acral melanocytic lesions: revision of the 3-step algorithm and refined definition of the regular and irregular fibrillar pattern. Dermatol Pract Concept 2022;12:e2022123.

Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 2018;5:180161.

Allogeneic Hematopoietic Stem Cell Transplant Recipients and Skin Cancer Risk: A Commercial Claims Cohort Study

Journal of Investigative Dermatology (2025) **145,** 1805–1807; doi:10.1016/j.jid.2024.10.617

TO THE EDITOR

Immunosuppression is a known risk factor for skin cancer and is well-documented for solid organ transplant (SOT) recipients (Euvrard et al, 2003; Hartevelt et al, 1990; Jensen et al, 1999), for whom skin cancer surveillance recommendations and risk prediction tools (ie, SUNTRAC) are well-established (Jambusaria-Pahlajani et al, 2019).

Allogeneic hematopoietic stem cell transplant (HSCT) recipients are also reported to have increased risks of skin cancer, with some studies suggesting risk comparable with that of renal SOT recipients (Herr et al, 2020; Omland et al, 2016; Rizzo et al, 2006; Scott et al, 2020; Wu et al, 2019). However, studies on skin cancer risk in allogeneic HSCT recipients are fewer than in SOT recipients and, in the United States, are typically single institution cohorts and report relative rather than absolute risks (Herr et al, 2020; Scott et al, 2020; Wu et al, 2019). We used a large commercial claims database to report the absolute and relative risks of skin cancers in allogeneic HSCT recipients.

We performed a retrospective cohort study using Optum's deidentified Clinformatics Data Mart Database (Clinformatics) from January 2007 to March 2022, with approximately 7.7 million patients with commercial health insurance. We included patients allogeneic HSCT procedure codes (Supplementary Table S1). As comparator groups, we included patients with SOT procedure codes (larger and more established high-risk group) (Supplementary Table S2) and patients with diagnosis codes for seborrheic keratosis (SK) (general population group with decreased potential for surveillance bias) (Supplementary Table S3). Patients had at least 1 year between dataset entry and the first allogeneic HSCT, SOT, or SK code and were aged ≥18 years. This study used deidentified data and, as such, does not require written informed consent (MD Anderson Institutional Review Board 2019-0966).

Our primary outcome was the first surgically treated skin cancer, including keratinocyte carcinoma (KC) (squamous and basal cell carcinomas) or melanoma (including all melanoma subtypes), identified using diagnosis codes and same-day treatment codes (Supplementary Tables S4 and S5) (Eide et al, 2010). Secondary outcomes included the first surgically treated KC or melanoma separately (if patients had KC, they were still followed for melanoma and vice versa). Follow-up started 31 days after allogeneic HSCT, SOT, or SK code.

We calculated the absolute risks of skin cancer overall and of KC and melanoma in allogeneic HSCT recipients overall and stratified by race (non-Hispanic White and non-White recipients) using cumulative incidence curves. We compared the skin cancer risk in allogeneic HSCT recipients with that in SOT recipients and patients with SK using competing risk regression (competing risk of death), adjusted for age, sex, race, history of actinic keratosis, history of skin cancer, and dermatology access (whether patients had an encounter with a dermatologist on or before the start of follow-up, using National Provider Identifier codes). In a secondary analysis, we evaluated risk factors for skin cancer within HSCT recipients, including factors such as graft-versus-host disease, chronic lymphocytic leukemia, and total body irradiation (Supplementary Tables S6–8).

We identified 3282 patients with allogeneic HSCT; 11,612 patients with SOT; and 3,715,638 patients with SKs (baseline characteristics are presented in Table 1). Patients with SK were more likely than allogeneic HSCT and SOT recipients to be non-Hispanic White, to have history of actinic keratoses, to have history of skin cancer, and to have had dermatology access (chi-square $P < .001$) (2-group comparisons are not shown in Table 1). Allogeneic HSCT

Abbreviations: CI, confidence interval; HR, hazard ratio; HSCT, hematopoietic stem cell transplant; KC, keratinocyte carcinoma; SK, seborrheic keratosis; SOT, solid organ transplant

Accepted manuscript published online 9 January 2025; corrected proof published online 3 February 2025