Research letter

Patients' and dermatologists' preferences in artificial intelligence–driven skin cancer diagnostics: A prospective multicentric survey study

To the Editor: Artificial intelligence (AI) has shown promise for improving diagnostics of skin cancer by matching or surpassing experienced clinicians.[1] However, the successful clinical application depends on acceptance by patients and dermatologists.

In this prospective multicentric survey study with a response rate of 63%, we therefore investigate the criteria required for patients and dermatologists to accept AI-systems and assess their importance on patients' and dermatologists' decision-making when considering the use of such systems. To this end, we perform an adaptive choice-based conjoint analysis and analyze it using hierarchical Bayes estimation.[2] By employing an adaptive choice-based conjoint analysis, we investigate multiple influencing AI-features simultaneously (see Table I) whilst accounting for possible trade-offs (see Fig 1). For details on questionnaire development, participant recruitment, and statistical analysis, see Supplementary Methods, available via Mendeley at https://data.mendeley.com/datasets/2chcwnhpwj/1.

The data of 293 respondents (178 patients and 115 dermatologists) showed a positive general attitude toward AI-systems (see Supplementary Results, available via Mendeley at https://data.mendeley.com/datasets/2chcwnhpwj/1 for participant characteristics). However, AI-systems were considered unacceptable by 42% of patients (95% confidence interval [CI]: 34%-49%) and 48% of dermatologists (95% CI: 38%-57%) if neither the dermatologist nor the patient could trace (ie, understand and follow) the assessment, and AI-systems were systematically ruled out by 37% of patients (95% CI: 29%-44%) and 36% of dermatologists (95% CI: 27%-45%) if they did not provide explanations on a case-by-case basis. Diagnostic accuracy and explainability were the most important AI-features in decision-making with an average importance of 21% (95% CI: 19%-22%) and 27% (95% CI: 26%-27%) for patients, and 33% (31%-35%) and 20% (19%-21%) for dermatologists, respectively.

Participants preferred an increased explainability with display of both decision criteria and relevant image regions. Patients prioritized an AI assessment that is traceable for patients and clinicians, and dermatologists preferred a multiclass differentiation among various disorders (see Supplementary Results, available via Mendeley at https://data.mendeley.com/datasets/2chcwnhpwj/1). Specifically, the differentiation between melanoma and nevi, which has been the primary focus of AI research in dermatology,[3] is considered insufficient. Consequently, there is a need for prospective studies evaluating AI-performance in multiclass assessments.

Current AI research is mainly performance-oriented (eg, International Skin Imaging Collaboration challenges[4]). However, patients and dermatologists require AI-systems that explain the rationale behind their decision-making and are at least somewhat traceable for patients and dermatologists. This growing demand for explainable AI poses a key challenge for future research since state-of-the-art technology does not fully explain the reasoning behind its decisions due to the AI black box phenomenon.[5]

Moreover, it is crucial to acknowledge that a substantial number of respondents in this survey study had a personal history of melanoma and therefore may have different perspectives on AI for skin cancer diagnostics compared to the general population (see Supplementary Fig 4, available via Mendeley at https://data.mendeley.com/datasets/2chcwnhpwj/1). To mitigate this potential bias, future studies should prioritize the recruitment of patients with no or other types of skin cancer.

In conclusion, the prioritization of AI-systems with increased explainability and traceability (ie, making them understandable) along with the call for multiclass decision-making, highlights that AI-systems need to evolve beyond pure performance advancements. Adhering to these criteria will be pivotal for fostering potentially more successful clinical adoption.

COLLABORATORS

Carina Nogueira Garcia, Janis Thamm, Sandra Schuh, Julia Welzel, Lara Valeska Maul-Duwendag, Paul Georg, Laurence Feldmeyer, Falk Bechara, Julian Steininger, Sophia Lehr, Ricarda Rauschenberg, Anna-Lisa Eberle, Sören Manfred Hartmann, Helena Hasler, Sophia Bender-Säbelkampf, Jürgen P. Bauerschmitz, Matthias Kaufmann, Cornelia Erfurt-Berge, Wiebke Katharina Peitsch, Ulrike Wehkamp, Marion Jost, Cindy Franklin, Julia Holzgruber, Maximilian

Table I. Overview of the artificial intelligence features and corresponding options within the adaptive choice-based conjoint design

AI-feature	Options
Integration How should the AI assessment be integrated into routine diagnostics?	• The physician first decides independently and then always obtains a second opinion from the AI. • The physician first decides independently and obtains a second opinion from the AI only in case of doubt. • The AI assessment is always obtained first, and the physician makes his or her decision based on it.
Explainability To what extent should the AI be able to explain its assessment?	• AI shows the criteria (eg, color, color distribution) and image regions used to make the assessment. • AI cannot display the image regions, but it displays which criteria (eg, color, color distribution) were used to make the assessment. • AI cannot display any criteria, but it shows which image regions were used to make the assessment. • AI does not have to explain its assessment on a case-by-case basis. However, it could be shown during the clinical trial that the AI pays attention to biologically relevant structures. • AI does not have to explain its assessment on a case-by-case basis. It could not be shown during the clinical trial that the AI pays attention to biologically relevant structures.
Traceability Who should be able to trace the AI assessment?	• The physician and the patient are able to trace (ie, understand and follow) the AI assessment. • The physician is able to trace (ie, understand and follow) the AI assessment. • Neither the physician nor the patient is able to trace (ie, understand and follow) the AI assessment.
Diagnostic accuracy Beyond what level of diagnostic accuracy should AI be used?	• AI performs worse than the average dermatologist. • AI performs equally well as the average dermatologist. • AI performs better than the average dermatologist.
Decision task (only asked for dermatologists) What should the AI be able to distinguish?	• AI distinguishes between benign and malignant skin lesions but gives no indication of a precise diagnosis. • AI makes recommendations for or against biopsy but gives no indication of a precise diagnosis. • AI distinguishes between melanomas and nevi. • AI distinguishes among melanomas, nevi and 1 category for other skin lesions. • AI distinguishes between melanomas and nonmelanomas. • AI distinguishes among melanomas, 1 category for other types of skin cancer and 1 for benign skin lesions.
Input data (only asked for patients) What data should the AI use for its assessment?	• AI makes a diagnosis based on skin images exclusively. • AI makes a diagnosis based on skin images and additional information about the skin lesion (eg, diameter). • AI makes a diagnosis based on skin images and additional information about the patient (eg, age). • AI makes a diagnosis based on skin images, additional information on the patient and the skin lesion.

Five artificial intelligence features and corresponding options were included in the adaptive choice-based conjoint analysis based on insights from a literature review and semistructured interviews. The decision task feature was included only for the subgroup of dermatologists, and the input data feature was included only for the subgroup of patients.
AI, Artificial intelligence.

Haist, Mario Giulini, Sebastian Wohlfeil, Valentina Faihs, Elke Sattler, Suzan Stürmer, Sebastian Krammer, Benjamin Kendziora, Semra Larissa Akcetin, Mohammed Mitwalli, Pinar Avci, Zeno Fiocco, Daniela Hartmann, Manuel Pedro Fernandes Lobo Pereira, Alexander Thiem, Valentin Aebischer, Stephan Forchhammer, Isabel Wolff.

Sarah Haggenmüller, MSc,[a] *Roman C. Maron, MSc,*[a] *Achim Hekler, MSc,*[a] *Eva Krieghoff-Henning, PhD,*[a] *Jochen S. Utikal, MD,*[b,c,d]

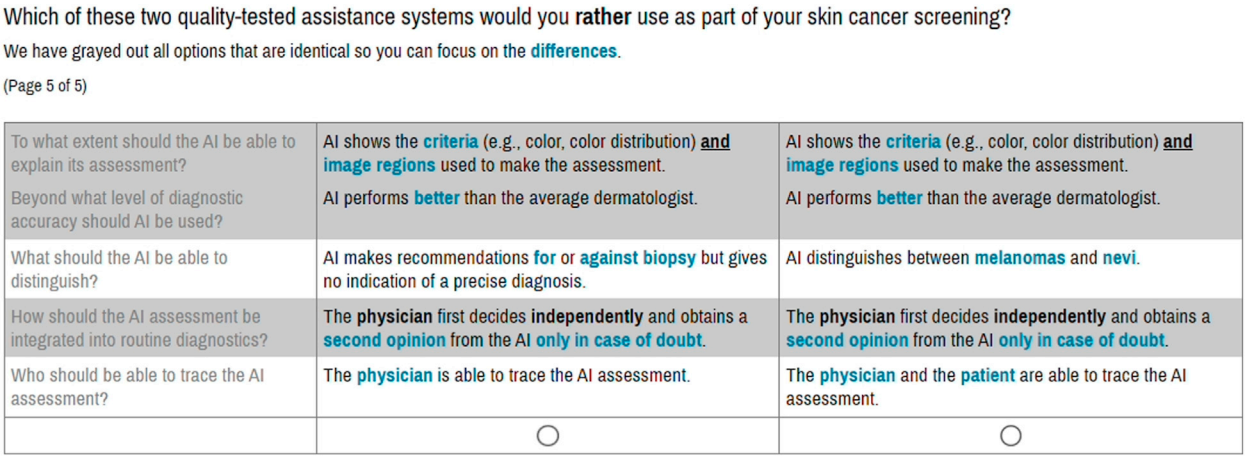

Fig 1. Example choice tournament of the present adaptive choice-based conjoint study design. The survey was conducted in German, and this example choice tournament was translated into English for this illustration. *AI*, Artificial intelligence.

Maria Gaiser, MD,[b,c,d] Verena Müller, MD,[b,c,d] Sascha Fabian, PhD,[e] Friedegund Meier, MD,[f] Sarah Hobelsberger, MD,[f] Frank F. Gellrich, MD,[f] Mildred Sergon, MD,[f] Axel Hauschild, MD,[g] Michael Weichenthal, MD,[g] Lars E. French, MD,[b,i] Lucie Heinzerling, MD,[b,k] Justin G. Schlager, MD,[b] Kamran Ghoreschi, MD,[j] Max Schlaak, MD,[j] Franz J. Hilke, PhD,[j] Gabriela Poch, MD,[j] Sören Korsing, MD,[j] Carola Berking, MD,[k] Markus V. Heppt, MD,[k] Michael Erdmann, MD,[k] Sebastian Haferkamp, MD,[l] Konstantin Drexler, MD,[l] Dirk Schadendorf, MD,[m] Wiebke Sondermann, MD,[m] Matthias Goebeler, MD,[n] Bastian Schilling, MD,[n] Jakob N. Kather, MD,[o] Stefan Fröhling, MD,[p] Katharina Kaminski,[q] Astrid Doppler,[q] Tabea Bucher, MSc,[a] Titus J. Brinker, MD,[a] and Collaborators

From the Digital Biomarkers for Oncology Group, National Center for Tumor Diseases (NCT), German Cancer Research Center (DKFZ), Heidelberg, Germany[a]; Department of Dermatology, Venereology and Allergology, University Medical Center Mannheim, Ruprecht-Karls University of Heidelberg, Mannheim, Germany[b]; Skin Cancer Unit, German Cancer Research Center (DKFZ), Heidelberg, Germany[c]; DKFZ Hector Cancer Institute at the University Medical Center Mannheim, Mannheim, Germany[d]; Department of Economics, University of Applied Science Neu-Ulm, Neu-Ulm, Germany[e]; Department of Dermatology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany and Skin Cancer Center at the University Cancer Centre Dresden and National Center for Tumor Diseases, Dresden, Germany[f]; Department of Dermatology, University Hospital (UKSH), Kiel, Germany[g]; Department of Dermatology and Allergy, University Hospital, LMU Munich, Munich, Germany[h]; Dr Phillip Frost Department of Dermatology and Cutaneous Surgery, University of Miami, Miller School of Medicine, Miami, Florida[i]; Department of Dermatology, Venereology and Allergology, Charité — Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany[j]; Department of Dermatology, University Hospital Erlangen, Comprehensive Cancer Center Erlangen — European Metropolitan Region Nürnberg, CCC Alliance WERA, Erlangen, Germany[k]; Department of Dermatology, University Hospital Regensburg, Regensburg, Germany[l]; Department of Dermatology, University Hospital Essen, Essen and German Cancer Consortium, partner site Essen and National Center for Tumor Diseases (NCT), NCT-West, Campus Essen and University Alliance Ruhr, Research Center One Health, University Duisburg-Essen, Essen, Germany[m]; Department of Dermatology, Venereology and Allergology, University Hospital Würzburg and National Center for Tumor Diseases (NCT) WERA, Würzburg, Germany[n]; Else Kroener Fresenius Center for Digital Health, Medical Faculty Carl Gustav Carus, Technical University Dresden, Dresden, Germany[o]; Department of Translational Medical Oncology, National Center for Tumor Diseases (NCT) Heidelberg and German Cancer Research Center (DKFZ), Heidelberg,

Germany[p]*; and Melanom Info Deutschland — MID e.V., Essen, Germany.*[q]

Funding sources: This study was funded by the Federal Ministry of Health, Berlin, Germany (grant: Skin Classification Project 2 [SCP2]; grant holder: Titus J. Brinker, German Cancer Research Center, Heidelberg, Germany), the Ministry of Social Affairs, Health and Integration of the Federal State Baden-Württemberg, Germany (grant: AI-Translation-Initiative [KTI]; grant holder: Titus J. Brinker, German Cancer Research Center, Heidelberg, Germany), and the Studienstiftung des deutschen Volkes, Bonn, Germany (scholarship holder: Sarah Haggenmüller). The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Dr Utikal has received travel support from Amgen, Bristol Myers Squibb, GSK, Immunocore, LeoPharma, Merck Sharp and Dohme, Novartis, Pierre Fabre, Roche, and Sanofi outside the submitted work. Dr Meier has received travel support from Novartis, Roche, BMS, MSD, and Pierre Fabre and research funding from Novartis and Roche. Dr Hobelsberger reports clinical trial support from Almirall and has received travel support from the following companies: UCB, Janssen Cilag, Almirall, Novartis, Lilly, LEO Pharma, and AbbVie outside the submitted work. Gellrich has received travel support from Sun Pharma, Sanofi, and Merck. Dr Erdmann declares travel support from Bristol-Meyers Squibb, Immunocore, and Novartis outside the submitted work. Dr Hauschild reports clinical trial support from the following companies: Agenus, Amgen, BMS, Dermagnostix, Highlight Therapeutics, Immunocore, Incyte, IO Biotech, MerckPfizer, MSD, NercaCare, Novartis, Philogen, Pierre Fabre, Regeneron, Roche, Sanofi-Genzyme, Seagen, Sun Pharma, and Xenthera, outside the submitted work. Dr Schilling reports research funding from BMS, Pierre Fabre Pharmaceuticals, and MSD and travel support from Novartis, Roche, BMS, Pierre Fabre Pharmaceuticals, and Amgen, outside the submitted work.

Patient consent: Not applicable.

IRB approval status: Not applicable (anonymous survey study).

Key words: acceptance; artificial intelligence; dermatologist view; dermatology; melanoma; patient view; skin cancer; survey study.

Correspondence to: Titus J. Brinker, MD, Digital Biomarkers for Oncology Group, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

E-mail: titus.brinker@dkfz.de

X handle: @TitusBrinker

Conflicts of interest

Dr Utikal is on the advisory board or has received honoraria from Amgen, Bristol Myers Squibb, GSK, Immunocore, LeoPharma, Merck Sharp and Dohme, Novartis, Pierre Fabre, Roche, and Sanofi outside the submitted work. Dr Meier has received speaker's fees or/and advisor's honoraria from Novartis, Roche, BMS, MSD, and Pierre Fabre. Dr Hobelsberger reports speaker's honoraria from Almirall, UCB, and AbbVie. Dr Gellrich has received speaker's fees or/and advisor's honoraria by Sun Pharma, Sanofi, and Merck. Dr Hauschild reports speaker's honoraria or consultancy fees from the following companies: Agenus, Amgen, BMS, Dermagnostix, Highlight Therapeutics, Immunocore, Incyte, IO Biotech, MerckPfizer, MSD, NercaCare, Novartis, Philogen, Pierre Fabre, Regeneron, Roche, Sanofi-Genzyme, Seagen, Sun Pharma, and Xenthera, outside the submitted work. Dr French is on the advisory board or has received consulting/speaker honoraria from for Galderma, Janssen, Leo Pharma, Eli Lilly, Almirall, Union Therapeutics, Regeneron, Novartis, Amgen, Abbvie, UCB, Biotest, and InflaRx. Dr Schlaak has received consultant or speaker fees or travel grants from BMS, MSD, Roche, Kyowa Kirin, Novartis, Sanofi Genzyme, Pierre Fabre, Sun Pharma, and Immunocore. Dr Erdmann declares honoraria from Bristol-Meyers Squibb, Immunocore, and Novartis outside the submitted work. Dr Haferkamp reports advisory roles for or has received honoraria from Pierre Fabre Pharmaceuticals, Novartis, Roche, BMS, Amgen, and MSD outside the submitted work. Dr Drexler has received honoraria from Pierre Fabre Pharmaceuticals and Novartis outside the submitted work. Dr Sondermann reports grants, speaker's honoraria, or consultancy fees from medi GmbH Bayreuth, Abbvie, Almirall, Amgen, Bristol-Myers Squibb, Celgene, GSK, Janssen, LEO Pharma, Lilly, MSD, Novartis, Pfizer, Roche, Sanofi Genzyme, and UCB outside the submitted work. Dr Schilling reports advisory roles for or has received honoraria from Pierre Fabre Pharmaceuticals, Incyte, Novartis, Roche, BMS, and MSD. Dr Goebeler has received speaker's honoraria and/or has served as a consultant and/or member of advisory boards for Almirall, Argenx, Biotest, Eli Lilly, Janssen Cilag, Leo Pharma, Novartis, and UCB, outside the submitted work. Dr Kather reports consulting services for Owkin, France, Panakeia, UK, and DoMore Diagnostics, Norway and has received honoraria for

lectures by MSD, Eisai, and Fresenius. Dr Brinker reports owning a company that develops mobile apps (Smart Health Heidelberg GmbH, Handschuhsheimer Landstr. 9/1, 69120 Heidelberg). Author Haggenmüller, Author Maron, Author Hekler, Dr Krieghoff-Henning, Dr Gaiser, Dr Müller, Dr Fabian, Dr Sergon, Dr Weichenthal, Dr Heinzerling, Dr Schlager, Dr Ghoreschi, Dr Hilke, Dr Pochi, Dr Korsing, Dr Berking, Dr Heppt, Dr Schadendorf, Dr Fröhling, Author Kaminski, Author Doppler, and Author Bucher have no conflicts of interest to declare.

REFERENCES

1. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115-118.
2. *ACBC Technical Paper. Technical Paper Series*. Sawtooth Software, Inc; 2014.
3. Haggenmüller S, Maron RC, Hekler A, et al. Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts. *Eur J Cancer*. 2021;156:202-216.
4. Marchetti MA, Liopyris K, Dusza SW, et al. Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma: results of the International Skin Imaging Collaboration 2017. *J Am Acad Dermatol*. 2020;82:622-627.
5. Hauser K, Kurz A, Haggenmüller S, et al. Explainable artificial intelligence in skin cancer recognition: a systematic review. *Eur J Cancer*. 2022;167:54-69.

https://doi.org/10.1016/j.jaad.2024.04.033