Aus dem Zentrum für Innere Medizin der Universität zu Köln
Klinik und Poliklinik für Innere Medizin I
Direktor: Universitätsprofessor Dr. med. M. Hallek

# Decision support for structural improvement of melanoma tumor boards: using standard cases to optimize workflow

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der Universität zu Köln

vorgelegt von
David Hoier aus Aachen

promoviert am 26. Februar 2025

Gedruckt mit Genehmigung der Medizinischen Fakultät der Universität zu Köln

2025

Dekan: Universitätsprofessor Dr. med. G. R. Fink

1. Gutachter: Privatdozent Dr. med. Th. Elter
2. Gutachterin: Universitätsprofessorin Dr. rer. medic. N. Ernstmann

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertationsschrift ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes habe ich Unterstützungsleistungen von folgenden Personen erhalten: Thomas Elter, Philipp Koll, Carolin Groß-Ophoff-Müller, Cindy Franklin, Michael Hallek, Esther von Stebut, Cornelia Mauch und Nicole Kreuzberg.

Weitere Personen waren an der Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin / eines Promotionsberaters in Anspruch genommen. Zur statistischen Beratung wurde Frau Elena Gilman beauftragt.

Die Dissertationsschrift wurde von mir bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die dieser Arbeit zugrunde liegenden Daten wurden der Tumorboard-Dokumentation der Abteilung für Dermatologie und Venerologie des Universitätsklinikums Köln von den hierzu Berechtigten entnommen und vor Verwendung im Rahmen der Dissertation datenschutzkonform pseudonymisiert. Ein Ethikvotum der Ethikkommission der Uniklinik Köln zur Durchführung dieser Arbeit liegt vor. Im weiteren wurden die Tumorboard Fallberichte von mir selbst ausgewertet und im Anschluss von Dr. med. Philipp Koll und nach Rücksprache mit PD Dr. med. Thomas Elter überprüft und überarbeitet. In eigenständiger Arbeit wurde das komplette Manuskript inkl. Datenkuration und Visualisierung von mir selbst erarbeitet. Die statistische Auswertung erfolgte in Zusammenarbeit mit der Statistikerin Elena Gilman. Deskriptive Statistiken und Datenanalysen wurden mit der IBM Statistiksoftware SPSS Version 28 und Microsoft Excel von Frau Gilman und mir selbst durchgeführt. Das Manuskript wurde anschließend gemeinsam mit Dr. med. Philipp Koll und PD Dr. med. Thomas Elter überarbeitet. Ein finaler Entwurf wurde erneut von Dr. med. Carolin Groß-Ophoff-Müller, Dr. med. Cindy Franklin, Prof. Dr. med. Michael Hallek, Prof. Dr. med. Esther von Stebut, Prof. Dr. med. Cornelia Mauch und Dr. med. Nicole Kreuzberg überarbeitet und anschließend im Journal of Cancer Research and Clinical Oncology publiziert.

Erklärung zur guten wissenschaftlichen Praxis:

Ich erkläre hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten (Amtliche Mitteilung der Universität zu Köln AM 132/2020) der Universität zu Köln gelesen habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Zürich, den 25.08.2024

# Danksagung

Mein herzlicher Dank geht an Thomas Elter und Phillip Koll. Beide haben mich bei der Erstellung der Dissertation unterstützt und im höchsten Maß motiviert.

Den Krebspatienten gewidmet

Für einen besseren Lesefluss wird in dieser Arbeit das generische Maskulinum verwendet, wobei alle Personenbezeichnungen gleichermaßen für alle Geschlechter gelten.

# Inhaltsverzeichnis

## Abkürzungsverzeichnis

| | |
|---|---|
| AI | artificial intelligence |
| AJCC | American Joint Committee on Cancer |
| APP | application |
| BRAF | rapidly accelerated fibrosarcoma isoform B |
| CE | Conformité Européenne |
| cKIT | tyrosine protein kinase |
| CLND | complete lymph node dissection |
| CT | computed tomography |
| DKG | Deutsche Krebsgesellschaft |
| DSS | decision support system |
| ECOG | Eastern Cooperative Oncology Group |
| EO | EasyOncology |
| G-BA | Gemeinsamer Bundesausschuss |
| IBM | International Business Machines Corporation |
| KI | künstliche Intelligenz |
| MDT | multidisciplinary tumor boards |
| n | number (of participants) |
| NA | not available |
| nRAS | neuroblastoma rat sarcoma |
| p-value | probability value |
| pT-stage | pathological stage |
| SLNB | sentinel lymph node biopsy |
| T-VEC | talimogene laherparepvec |
| WFO | Watson for Oncology |

# 1. Zusammenfassung

Die Behandlung von Krebspatienten steht vor einer Versorgungs-Herausforderung. Das Grundverständnis von Krebserkrankungen und somit auch neuen Behandlungsoptionen wächst rasant. Eine scheinbar unübersichtliche Vielzahl innovativer und effektiver Krebsmedikamente bietet neue Therapieansätze und eine bessere Versorgung von Erkrankten.

In der täglichen klinischen Routine sehen Onkologen üblicherweise viele Patienten mit unterschiedlichen Tumoren und müssen umfangreiche Befund- und Therapiegespräche führen. Im medizinischen Alltag fehlt oft die Zeit, einfühlsame Gespräche mit Patienten zu führen, deren optimale Versorgung zu gewährleisten und sich über neue Behandlungsmöglichkeiten mit Kollegen auszutauschen.

In der Gesamtheit stehen Ärzte somit vor der Herausforderung, die rasanten Entwicklungen in Diagnostik und Therapie zu erfassen und diese unter ständigem Zeitdruck umzusetzen. Daher stellt sich die entscheidende Frage, wie sich unter diesen schwierigen Umständen die bestmöglichen Therapieentscheidungen treffen lassen.

Zur Entscheidungsfindung orientieren sich Ärzte insbesondere an den Leitlinien der Fachgesellschaften und an Empfehlungen von multidisziplinären Tumorboards (MDT). Bei Letzteren handelt es sich um eine Expertenrunde erfahrener Mediziner unterschiedlicher Fachrichtungen, die sich regelmäßig treffen und die Therapiemöglichkeiten jedes einzelnen Krebspatienten diskutieren. Im Anschluss wird gemeinsam ein Behandlungsplan erstellt. Diese Vorgehensweise bildet nicht zwangsläufig den aktuellen Wissensstand ab, da auch in aktuellen Leitlinien beispielsweise neu zugelassene Medikamente fehlen können. Außerdem ist die Qualität der Entscheidungen von multidisziplinären Tumorboards hochgradig von der Erfahrung, der Qualifikation und dem aktuellen Wissen der teilnehmenden Ärzte abhängig.[1] Diese Gründe können zur Abweichung von der als besten klinischen Praxis betrachteten Behandlung führen und es kann davon ausgegangen werden, dass ein relevanter Anteil der Krebspatienten keine optimale Behandlung erhält.[2-4] Eine Auswertung der Nationalen Krebsdatenbank (USA) zur Behandlung von Melanomen zeigte zum Beispiel, dass etwa 20 % der Patienten mit T2/3-Melanomen nicht gemäß den Leitlinien behandelt wurden.[5]

Um die Qualität onkologischer Behandlungen zu gewährleisten, wurde die Zertifizierung von Kliniken als Krebszentrum durch die Deutsche Krebsgesellschaft etabliert.

Zertifizierte Krebszentren sind unter anderem verpflichtet, alle Krebspatienten in multidisziplinären Tumorboards zu besprechen, um die Zertifizierungsanforderungen

fortlaufend zu erfüllen. Gemäß den Zertifizierungskriterien müssen alle Behandlungsfälle ab Stadium IIC aufwärts in multidisziplinären Tumorboards vorgestellt und die gegebenen Empfehlungen dokumentiert werden. Dabei ist zu beachten, dass Fälle, deren Behandlungskonzepte ohne umfangreiche Diskussion entschieden werden können, bereits bei der Anmeldung zur Konferenz als Standardfälle definiert und damit im Konferenzprotokoll gekennzeichnet werden können. Als Standardfälle gelten in diesem Zusammenhang eindeutige Leitlinienfälle, z. B. Frühstadien, die keiner ausführlichen interdisziplinären Diskussion bedürfen. In der klinischen Realität führte die Zertifizierungsvorgabe, auch Standardfälle zu diskutieren, somit zu einer erheblichen Zunahme der Fallbesprechungen pro Tumorboard mit absehbar weniger Zeit und Aufmerksamkeit zur Besprechung komplexerer Tumorfälle.

Paradoxerweise gibt es kaum Untersuchungen zur Objektivierung der Qualität von Behandlungsentscheidungen multidisziplinärer Tumorboards, obgleich die Notwendigkeit hierzu offensichtlich ist. Insbesondere, das Fehlen einer Qualitätsevaluierung individueller Therapieentscheidungen in der evidenzbasierten Krebsbehandlung stellt eine Lücke im derzeitigen Stand der Forschung dar.[1, 6]

Ziel dieser Arbeit war es, ein Entscheidungsunterstützungssystem zu validieren, das Behandlungsempfehlungen bereits bei Konferenzanmeldung für jene Fälle erteilt, die sich durch eindeutige und aktuelle Vorgaben der Leitlinien herleiten lassen und diese als "Standardfälle" klassifiziert. Laut DKG-Vorgaben können diese bereits bei Anmeldung mit einer Therapieempfehlung versehenen Standardfälle nur noch im Konferenzprotokoll aufgeführt werden. Eine zeitaufwändige Diskussion erfolgt somit nur noch optional. Dies ermöglicht den Teilnehmern der Tumorkonferenz mehr Zeit für die Diskussion komplexer Fälle. Das Bestreben der vorliegenden Arbeit bestand also nicht darin, digitale Empfehlungen für alle Fälle zu erteilen, die in multidisziplinären Tumorboards vorgestellt werden oder die klinische Erfahrung der Ärzte zu ersetzen. Sie zielte stattdessen auf die Notwendigkeit ab, diese kritischen Konferenzen zeitlich zu entlasten, damit die Teilnehmer ihre Expertise auf die komplexeren Tumorfälle konzentrieren können.

## 2. Einleitung

Die Präzisionsmedizin entwickelt sich rasant und DSS bieten ein immenses Potenzial zur Unterstützung der klinischen Entscheidungsfindung. Zukünftig werden Ärzte vermehrt im Alltag auf Entscheidungsunterstützungssysteme zurückgreifen können, einschließlich auf künstlicher Intelligenz (KI) basierende Systeme. Die KI soll in diesem Kontext die Lücke

zwischen einer unübersichtlichen Vielzahl an Therapiemöglichkeiten und einer damit einhergehenden erschwerten klinischen Entscheidungsfindung schließen.[6-8]

Auf den ersten Blick scheinen KI-basierte Systeme gut geeignet zu sein, um die immensen Mengen medizinischer Daten nach Korrelationen und Mustern zu analysieren und daraus Therapieempfehlungen abzuleiten. Im Einklang mit dieser Annahme haben beispielsweise KI-basierte DSS bei der Analyse radiologischer Bilder für unterschiedliche Krebserkrankungen eine hervorragende Genauigkeit und Sensitivität gezeigt.[9-12]

Im Gegensatz zur Analyse von Bilddaten bieten KI basierende Systeme bislang keine verlässliche Unterstützung bei onkologischen Therapieempfehlungen. Hier gelingt es selbst bei vermeintlich einfachen Standardfragen zur Erstlinientherapie nicht, eine ausreichende Empfehlungsicherheit zu gewährleisten.[13, 14]

Ein Beispiel hierfür ist der KI-basierte Algorithmus Watson for Oncology (WFO), der entwickelt wurde, um Behandlungsempfehlungen für Krebspatienten zu geben. Die WFO-Empfehlungen stimmten zu 12 % (Magenkrebs), 80 % (Dickdarm- und Brustkrebs) und 93 % (Eierstockkrebs) mit den Behandlungsempfehlungen des MDT überein.[15-20] Darüber hinaus wurden für andere Krebsarten relativ niedrige Konkordanzraten ermittelt.[18, 21-23]

Die Gründe hierfür sind vielfältig. Eine wichtige Ursache für die negative Performance von bisherigen KI-Systemen ist der Mangel an hochwertigen Trainingsdatensätzen. Bei KI-Systemen zur Bilderkennung sind diese massenhaft und normiert verfügbar, nicht aber als reguläre Fälle aus der onkologischen Routineversorgung.

Neben dem Mangel an ordnungsgemäß organisierten und geprüften Trainingsdaten ist als weiteres Problem der erhebliche technische, personelle und finanzielle Ressourcenaufwand zu nennen, der zur Integration der Systeme in die teils sehr schwachen IT-Strukturen vieler Kliniken erforderlich ist.

Nicht zu unterschätzen ist zudem, dass in der Realität Deutscher Kliniken und Arztpraxen viele Befunde auswärtiger Institute weiterhin per Post oder Fax zugestellt und erst deutlich zeitverzögert in die Klinikinformationssysteme (KIS) oder die Praxissoftware eingescannt werden. In der Planung KI-basierter DSS muss also bedacht werden, dass selbst im Falle einer perfekten und maschinenlesbaren Dateninfrastruktur in den Kliniken und Praxen die essentiellen Daten zum frühen Zeitpunkt der Therapieentscheidungen noch nicht verfügbar sein können.

Trotz der zunehmenden Bedeutung von KI-basierten DSS in der Gesundheitsversorgung ist ihre Umsetzung in der onkologischen Routineversorgung also noch nicht angekommen und bleibt eine Herausforderung. Es bedarf also eines Systems, das unabhängig von vorhandener oder nicht-vorhandener Systeminformation die Informationen dort erhält, wo sie erstmals den Ärzten vollständig verfügbar sein müssen - bei der Anmeldung eines Patienten zur Vorstellung im Tumorboard.

Einen überlegenen und praxisnahen Ansatz für den Einsatz in der klinischen Routineversorgung können von Experten kuratierte DSS-Algorithmen bieten, die auf einer von onkologischen Fachleuten bereitgestellten Wissensbasis entwickelt wurden.

Wie von Bungartz et al. (2018) beschrieben, bieten diese von Experten kuratierten DSS im Vergleich zu KI-basierten Systemen noch weitere Vorteile.[24] Insbesondere scheinen Experten-kuratierte DSS besser in der Lage zu sein, die klinische Realität abzubilden als die bislang vorgestellten, KI-basierten Systeme. In dieser Arbeit wird erstmals ein Experten-kuratiertes DSS zur Entscheidungsunterstützung in der Therapie des malignen Melanoms vorgestellt.

## 2.1. Fragestellungen und Ziel der Arbeit

Ziel dieser Arbeit war es, ein Entscheidungsunterstützungssystem für maligne Melanome zu entwickeln, das bereits bei der Konferenzanmeldung Behandlungsempfehlungen für scheinbar einfache Fälle gibt und diese als "Standardfälle" klassifiziert. Gemäß den Zertifizierungsanforderungen ist die interdisziplinäre Diskussion von Standardfällen im Tumorboard fakultativ, sodass mehr Zeit für komplexe Fälle zur Verfügung stehen würde. In dieser Hinsicht wird das maligne Melanom als eine geeignete Modellerkrankung für die Entwicklung eines DSS für multidisziplinäre Tumorboards betrachtet, da der Melanom-Behandlungsalgorithmus als vergleichsweise komplex gilt.

Die Fragestellung der vorliegenden Arbeit zielte also darauf ab, den potenziellen Nutzen eines DSS zur Optimierung des Arbeitsablaufs von Tumorkonferenzen zu bewerten. Hierzu wurden die Konkordanzraten der Diagnose- und Therapieempfehlungen für neu diagnostizierte Fälle ermittelt, die von EO und einem zertifizierten MDT für Patienten mit kutanem malignem Melanom vorgeschlagen wurden.

# 3. Material und Methoden – Die Publikation

Journal of Cancer Research and Clinical Oncology (2024) 150:115
https://doi.org/10.1007/s00432-024-05627-3

RESEARCH


## Digital decision support for structural improvement of melanoma tumor boards: using standard cases to optimize workflow

David Hoier[1] · Carolin Groß-Ophoff-Müller[3] · Cindy Franklin[2] · Michael Hallek[1] · Esther von Stebut[2] · Thomas Elter[1] · Cornelia Mauch[2,4] · Nicole Kreuzberg[2] · Philipp Koll[2]

Received: 31 May 2023 / Accepted: 18 January 2024


**Abstract**

**Purpose** Choosing optimal cancer treatment is challenging, and certified cancer centers must present all patients in multidisciplinary tumor boards (MDT). Our aim was to develop a decision support system (DSS) to provide treatment recommendations for apparently simple cases already at conference registration and to classify these as "standard cases". According to certification requirements, discussion of standard cases is optional and would thus allow more time for complex cases.

**Methods** We created a smartphone query that simulated a tumor conference registration and requested all information needed to provide a recommendation. In total, 111 out of 705 malignant melanoma cases discussed at a skin cancer center from 2017 to 2020 were identified as potential standard cases, for which a digital twin recommendation was then generated by DSS.

**Results** The system provided reliable advice in all 111 cases and showed 97% concordance of MDT and DSS for therapeutic recommendations, regardless of tumor stage. Discrepancies included two cases (2%) where DSS advised discussions at MDT and one case (1%) with deviating recommendation due to advanced patient age.

**Conclusions** Our work aimed not to replace clinical expertise but to alleviate MDT workload and enhance focus on complex cases. Overall, our DSS proved to be a suitable tool for identifying standard cases as such, providing correct treatment recommendations, and thus reducing the time burden of tumor conferences in favor for the comprehensive discussion of complex cases. The aim is to implement the DSS in routine tumor board software for further qualitative assessment of its impact on oncological care.

**Keywords** Malignant melanoma · Algorithm · Multidisciplinary tumor board (MDT) · Tumor board evaluation · Digital recommendations · Documentation burden · Digital health · Mobile application · Expert-curated decision support system · Oncology

## Introduction

Everyday dermato-oncologists face the challenge to ensure evidence-based best clinical practice treatment for their cancer patients. Continuously updated treatment recommendations lead to an almost incomprehensible number of treatment options (Iqvia 2022). In malignant melanoma, we are faced with a rapidly changing therapeutic landscape in which immunotherapeutics and targeted treatment options are the gold standard in palliative and curative treatment (DKG 2020). Other potentially new emerging fields are neoadjuvant treatment approaches and adapted surgical procedures at earlier cancer stages (Amaria 2022; Luke 2022).

However, there are often deviations from the treatment considered best clinical practice and it must be assumed that a significant proportion of cancer patients do not receive optimal care (Bierbaum 2020; Bierbaum 2023; Heins 2017). An evaluation of the National Cancer Database (USA) on melanoma treatment showed, for example, that approx. 20% of patients with T2/3 melanoma were not treated according to the guidelines (Narang 2021).

In order to improve the quality of oncological care in Germany, the certification of cancer centers was implemented and has since become increasingly mandatory. Current data shows that treatment in certified cancer centers leads to higher treatment quality and better survival

David Hoier, Philipp Koll have contributed equally to this work.

Extended author information available on the last page of the article



Published online: 08 March 2024

rates (Beckmann 2011; Dkg 2019; Kowalski 2022; Modabber 2021; Schmitt 2022; Wolff 2017).

A central feature for certification as a cancer center by the German Cancer Society (DKG) is the presentation of all treated melanoma patients from Stage IIC in a multidisciplinary tumor board (MDT) (Kowalski 2017). At the skin cancer center Cologne all melanoma cases from stage IIB have to be presented to MDT.

However, the MDT presentation of all tumor patients due to certification requirements led to a significant quantitative increase in case discussions. Studies have shown that this does not automatically translate to an increase in quality (Soukup 2016; Walraven 2019).

A large number of cases to be discussed pose a challenge to the participants in these boards (Soukup 2016; Walraven 2019). Data evaluating the extent to which MDT improves outcomes remain mixed and some studies have found a survival benefit, while others have found no such advantage. A 2019 analysis indicated that the 5 year survival rate was 15.6% higher among cases in well-organized MDT but almost 20% lower in disorganized MDT compared with no MDT (Keating 2013; Kesson 2012; Lu 2019; Stone 2020; Wong 2022).

In the end, the quality of MDT's depends on the expertise and motivation of the participants and the time available to discuss the individual cases (Jalil 2018).

Decision support systems (DSS) could provide useful support here and appear to offer remarkable potential (Chen 2016; Letzen 2019; Soukup 2019). Surprisingly, however, artificial intelligence (AI)-based DSS have so far failed to gain acceptance in the field of clinical oncology (Bungartz 2018). This is even the case for standard oncological questions regarding first-line therapy (Schmidt 2017).

Expert-curated DSS algorithms, developed on knowledge base provided by oncological professionals, might be a superior approach for adequate decision support. As described in Nature Biotechnology in 2018, these expert-curated DSS provide multiple advantages when compared with AI-based systems (Bungartz 2018). Most importantly, expert-curated systems seem to represent clinical reality better than the AI-based systems used so far.

The aim of the present study was to evaluate the potential benefit of a DSS to optimize the workflow of tumor conferences.

## Materials and methods

We first created a query that simulated a melanoma tumor conference registration and requested all the information needed to provide a treatment recommendation. This algorithm was then implemented in the established oncology smartphone application "EasyOncology" (EO). This certified medical product is aimed at specialized personnel and provides the treatment concepts for the most common tumor entities. Thus, the basis was provided to easily match the smartphone DSS recommendations with the real decisions of the MDT and to determine the quality of the concordance.

To evaluate the reliability of each DSS recommendation, we followed the same approach that was used to assess the accuracy of AI-based DSS and benchmarked the digital treatment recommendations against real-world MDT decisions (Choi 2019; Kim 2019; Lee 2018; Somashekhar 2018; Yu 2021; Zhou 2019). For this purpose, we determined the concordance rates of diagnostic and therapeutic recommendations for newly diagnosed cases proposed by our DSS and a certified MDT for patients with cutaneous malignant melanoma.

### Smartphone application

The smartphone application EasyOncology (EO) was developed by clinically experienced oncologists and is intended to provide evidence-based diagnostic and therapeutic recommendations for common solid cancer entities. EO's oncological treatment algorithm "*therapy finder*" is based on a decision tree, which was developed through a systematic process with clinical experts in oncology across diverse institutions. More precisely, EO's platform is based on current oncological guidelines, (e.g., S3-guidelines (Dkg 2020) and NCCN guidelines (Swetter 2021)), drug approval status, current publications of relevant studies, and best clinical practice from leading German cancer centers.

Frequent testing and challenging of the algorithm with real-world test cases enables identification of practice changing medical standards with subsequent corresponding adjustment of the query. Finally, frequent version updates ensure to display the latest advancements in the field of dermato-oncology.

EO was ranked top three in a worldwide comparison of 157 oncological applications in 2017 and was certified as a medical device in 2020 (Calero 2017). The software is a CE marked medical device and subject to the according regulations to ensure its security and reliability. The software relies on anonymous input without exchange of identifiable patient information via hospital intranet.

In the present work, EO's *therapy finder (version 5.06)* was used to generate first-line diagnostic and therapeutic recommendations for patients with all stages of newly diagnosed cutaneous malignant melanoma. This software version displayed the 8th edition of the AJCC classification for melanoma. A graphic illustration of the App interface of EasyOncology's *therapy finder* is depicted in Fig. 1.

The DSS query algorithm of EO's *therapy finder* requests clinicopathologic data to generate treatment recommendations in a stepwise fashion.

The variables requested by the DSS included Breslow thickness; the presence of an ulceration (pT-stage); the histopathologic results of the sentinel lymph node biopsy (SLNB); radiologic staging information; the histopathologic results of the complete lymph node dissection (CLND); the postresection residual cancer status (R0, R1, and R2); the clinical and pathological evaluation of in-transit and satellite metastases; and the presence of important driver gene mutations (i.e., BRAF, NRAS, and cKIT).

The number of input variables necessary to generate a treatment recommendation depended on the complexity of each case. For simple cases (i.e., early-stage localized malignant melanoma) merely two variables are needed for DSS output, whereas more complex cases required up to five clinicopathologic variables. A simplified graphic illustration of how EO's *therapy finder*-based DSS generates treatment recommendations is depicted in Fig. 2.

## Definition of standard cases and study inclusion criteria

According to certification criteria, all treatment cases from stage IIC upward must be presented in MDT and given recommendations are to be documented. Of note, those cases whose treatment concepts can be decided without extensive discussion can already be defined as "standard cases" when registering for the conference and thus flagged in the conference protocol. In this context, clear guideline cases, e.g., early stages that do not require extensive interdisciplinary discussion, are considered as standard cases. Since a discussion of these standard cases is not mandatory, they do not consume any time of the actual conference.

In clinical routine, these possible standard cases are often not recognized at the time of registration, for example because the registering clinician does not have sufficient clinical experience. Thus, an increase in the proportion of standard cases that were pre-answered by a DSS could relieve the tumor conference accordingly.

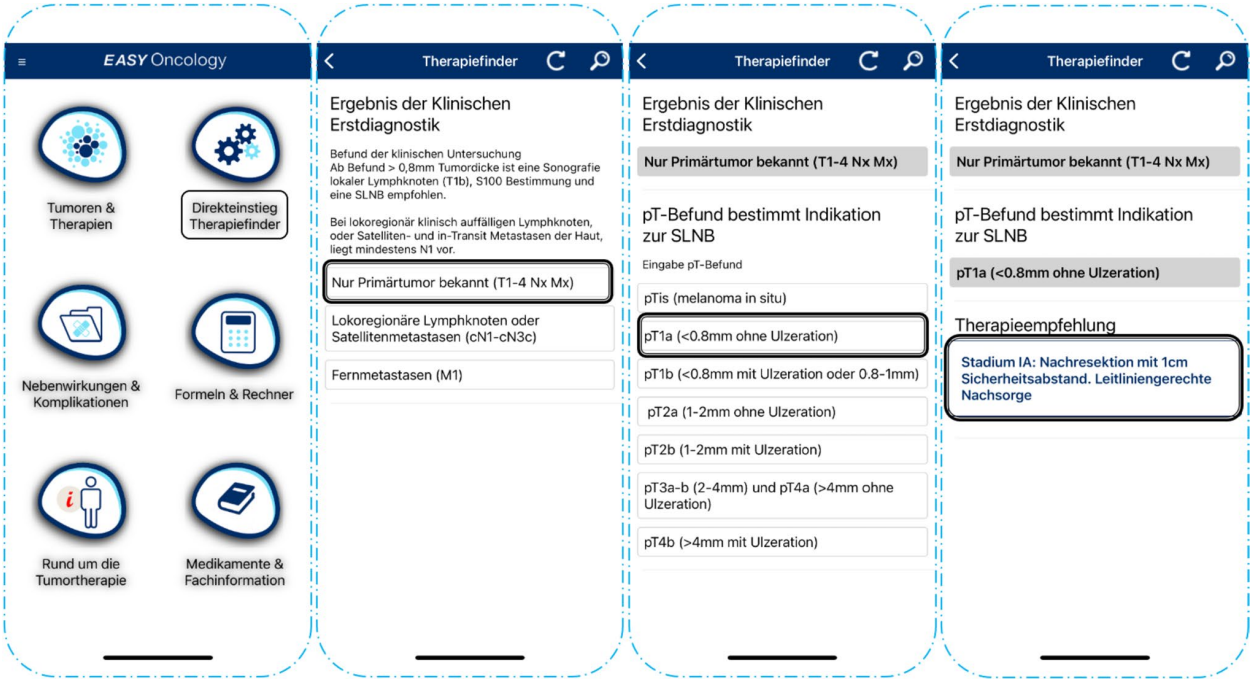


**Fig. 1** App interface of EasyOncology's therapy finder: Stepwise diagnostic query for malignant melanoma treatment recommendation. Relevant information is requested by EO's query algorithm to enable a treatment recommendation. In this example of primary disease, the pT-stage is at first requested. Subsequently, the algorithm query requires the Breslow thickness, which, in this example, leads to the treatment recommendation for stage IA melanoma

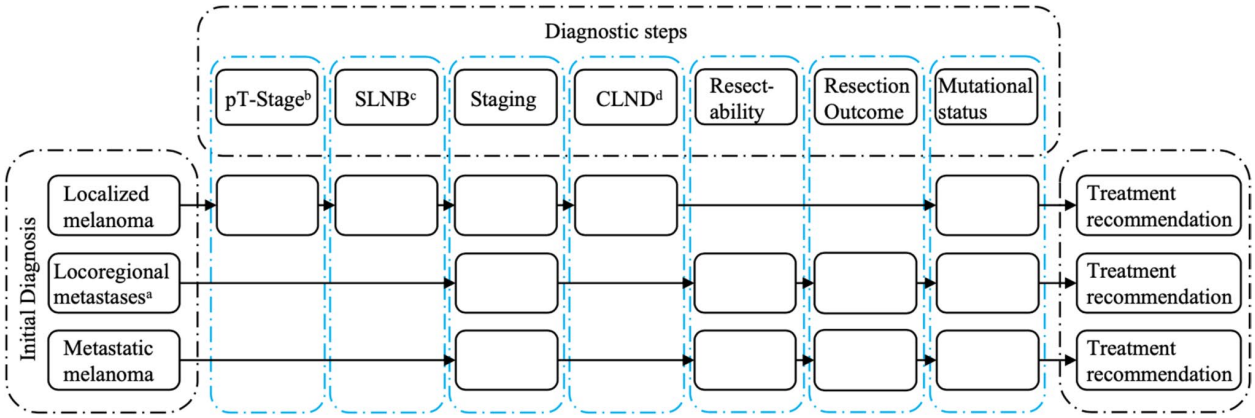


**Fig. 2** Query algorithm of EasyOncology's therapy finder-based DSS. Depending on the selected initial diagnosis relevant diagnostic steps are requested by EO's query algorithm until a treatment recommendation is given. Abbreviations: a: including satellite and in-transit metastases; b: determined by Breslow thickness (in mm) and presence of an ulceration; c: sentinel lymph node biopsy; d: complete lymph node dissection

In accordance with the specifications for automatic definition as a standard case, only cases without complicating factors were included in the analysis (Fig 3.). This procedure also corresponds to our approach that complicated cases should of course be discussed in the tumor board and by no means just decided digitally.

Cases with non-cutaneous melanoma (i.e., mucosal melanoma, uveal melanoma, or melanoma of unknown origin), as well as patients with relapsed disease, with secondary cancers, severe comorbidities, patients treated in clinical trials, or patients who explicitly declined diagnostic procedures (i.e., SLNB) were excluded from evaluation.

By intention, cases with brain metastases were also excluded from analysis, as stereo tactical or neurosurgical procedures should not be declared as standard cases. Lastly, we excluded those cases that lacked relevant clinicopathologic data needed as input variables by EO's *therapy finder*.

## Patient selection and study design

Ethical approval to conduct this work was granted by the Ethics Committee of the Medical Faculty of the University of Cologne (#20-1116).

The retrospective MDT dataset evaluation initially included 2399 cases with malignant diseases of the skin who received treatment at the Department of Dermatology and Venereology, University of Cologne, between January 2017 and December 2020.

As depicted in the study inclusion flowchart (Fig 3), we first excluded all non-melanoma cases. In total, 705 melanoma cases remained that were screened for eligibility, of which 594 patient cases presented to the MDT were not suitable for analysis with EO.

Finally, MDT treatment recommendations of 111 cases that fulfilled our pre-selection criteria for standard cases were included for comparison. Hereafter, the clinical information that remained was used to generate a DSS treatment recommendation.

Treatment recommendations for each case given by MDT or DSS were compiled as response pairs. Subsequently, after blinding of decision origin, each response pair was assessed for obvious discrepancies in recommendations and accordingly classified as "concordant" or "incorrect recommendations."

In a second independent review, an experienced dermato-oncologist analyzed each non-concordant decision pair for their quality of decisions and sub-grouped them in three categories, similar to previous publications evaluating the DSS Watson for Oncology by IBM (Choi 2019; Kim 2019; Lee 2018; Somashekhar 2018; Yu 2021; Zhou 2019):

1. If both decisions of DSS and MDT were identical, recommendations were classified as "concordant."
2. If both decisions of "non-concordant" cases were different, but correct clinical alternatives, classification changed to "correct alternative recommendation."
3. Case pairs were classified as "incorrect recommendation" if one of the digital or real-world recommendations was either incongruent with current best clinical practice guidelines or did not provide any recommendation, at all.

In fact, if using comparative cases from the past, the DSS will provide an updated treatment recommendation, thus leading to a concordance rate that only reflects the

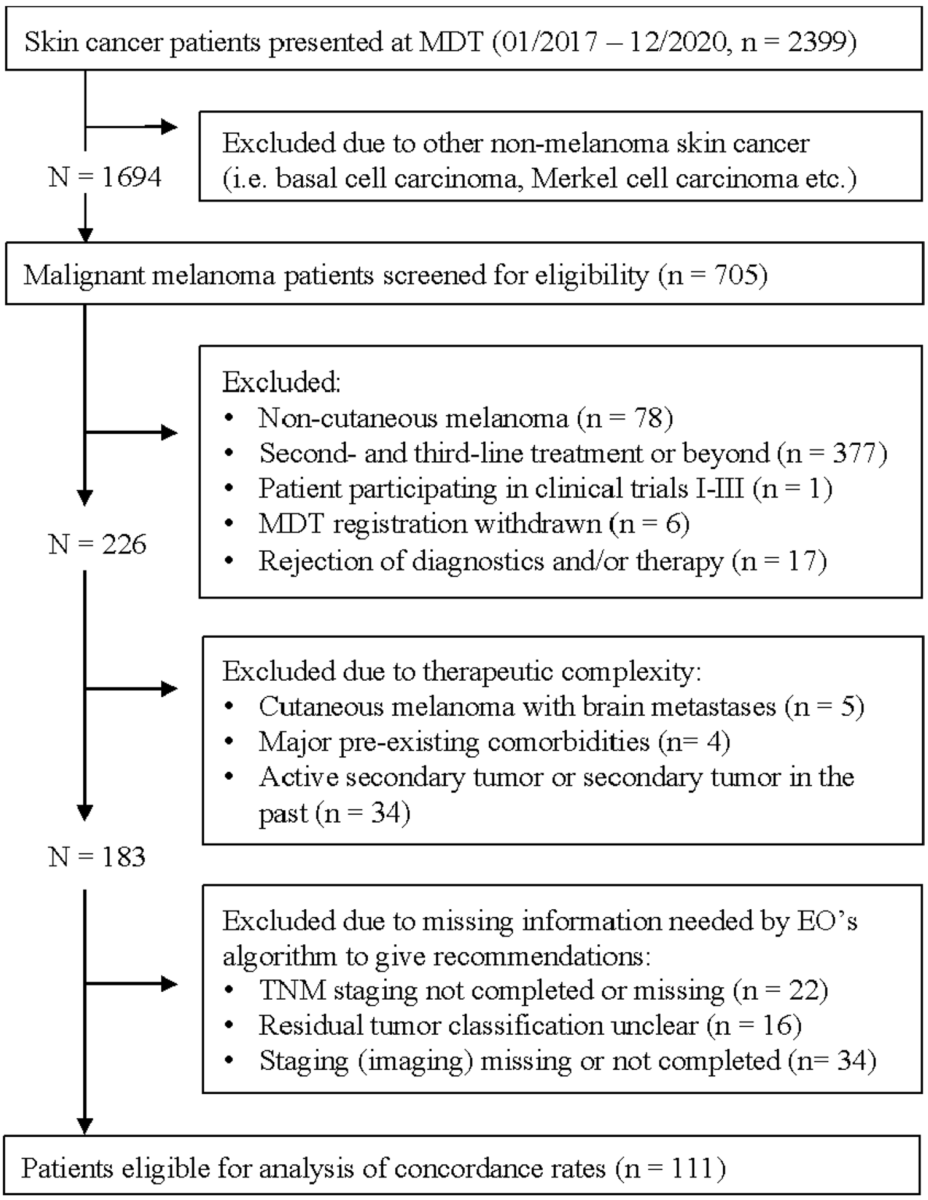


**Fig. 3** Flowchart of patient case exclusion process

change of treatment concepts between two time points. Therefore, we assumed historic MDT treatment decisions to be correct and decided to classify this time-dependent deviation in the same group as those more actual cases with correct alternative recommendation according to best clinical standard.

As an example, for the group of "correct alternative recommendation" cases, our DSS recommended adjuvant therapy based on the mutational status, whereas, prior to 2018, MDT issued a recommendation for adjuvant interferon.

Finally, "incorrect recommendations" cases were analyzed in detail to identify potential algorithm query errors. The evaluation process is depicted in Fig. 4.

## Data analysis and statistics

Descriptive statistics and data analysis were carried out using IBM's statistics software SPSS version 28 and Microsoft Excel. Descriptive statistics were depicted as numbers, percentages, or median. In line with previous publications, the concordance rate was presented as a percentage agreement between DSS and MDT, i.e.,

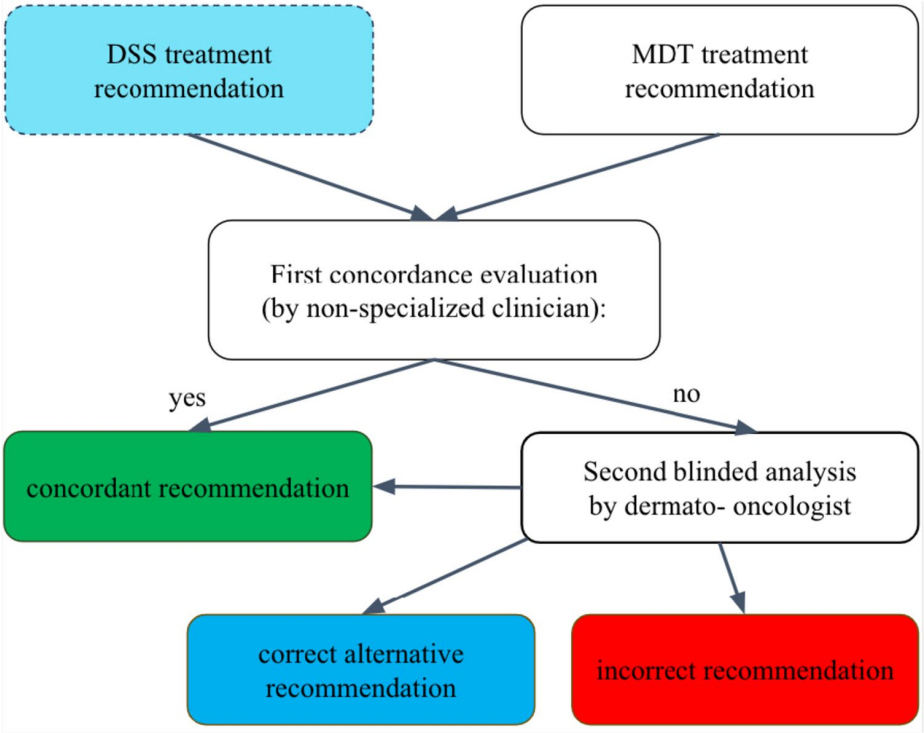


**Fig. 4** Flowchart for evaluating decision concordance of recommendations given by MDT or DSS. A first evaluation compared DSS and MDT treatment recommendations for concordance. Discordant recommendations were blinded to their origin and analyzed in detail by an dermato-oncologist, who categorized each recommendation either as "concordant recommendation," "correct alternative recommendation" or as "incorrect recommendation"

$$overallconcordance = \frac{concordantrecommendations+correctalternativerecommendations}{allrecommendations}$$

(Choi 2019; Kim 2019; Lee 2018; Somashekhar 2018; Yu 2021; Zhou 2019). After assigning patients to the concordant or the non-concordant group, a Chi-squared test was used to compare categorical variables and the Mann–Whitney U test was applied to compare ordinal and numerical variables between the groups. Statistical significance was assumed if the p-value was $< 0.05$ for all statistical analysis. Graphics, charts, and tables were generated using SPSS, Microsoft Excel, and PowerPoint.

## Results

### Descriptive statistics

Clinicopathologic characteristics of 111 malignant melanoma patients fulfilling our predefined standard case inclusion criteria for the determination of concordance are depicted in Table 1.

Median age of these patients was 62 years [interquartile range 52–72], and most patients ($n = 93$, 83.8%) had a good performance status (Eastern Cooperative Oncology Group (ECOG): 0) (<.001). At initial diagnosis, the majority of patients ($n = 99$, 89.2%) presented with localized melanoma, 7 (6.3%) presented with "regional lymph node metastases," and 5 (4.5%) of them had "metastatic melanoma." Ninety (81.1%) patients underwent a "sentinel lymph node biopsy" (SLNB) of whom 76 (68.5%) had at least one "positive" SLNB and 14 (12.6%) had a "negative" SLNB. A "complete lymph node dissection" was performed in 1 (0.9%) case. The number of patients with stage I, II, III, and IV melanoma was 3 (2.7%), 19 (17.1%), 85 (76.6%), and 4 (3.6%), respectively.

### Concordance rates

In decisions regarding the optimal first-line treatment for patients with malignant melanoma, the overall concordance rate between recommendations proposed by our DSS and those given by MDT was 97%. This includes 87 (78%) "concordant" cases and 21 (19%) "correct alternative recommendation" cases (Fig. 5a). Treatment concordance rates according to malignant melanoma stages, i.e., I, II, III, and IV, were 100%, 95%, 98%, and 100%, respectively (Fig. 5b). Quality of concordance was independent of age, melanoma stage, histologic subtype, gene mutation status, and complete lymph node dissection.

### Non-concordant cases

As requested by protocol, the 3 "incorrect" cases were analyzed to identify potential systematic errors caused by our DSS decision algorithm. This independent review process was performed by an experienced dermato-oncologist. Two of these non-concordant cases with high-risk melanoma

**Table 1** Clinicopathologic characteristics of malignant melanoma patients selected for concordance analysis

| Clinicopathologic characteristics | n (%) | concordant | non-concordant | p-value |
|---|---|---|---|---|
| Malignant melanoma patients | 111 | 108 | 3 | |
| Median age (range) | 62 (52–72) | 59 (51–70) | 57 (57) | .476 |
| ≤ 45 | 16 (14.4%) | 16 (14.8%) | 0 (0%) | .553 |
| 45–65 | 43 (38.7%) | 41 (38.0%) | 2 (66.7%) | |
| ≥ 65 | 52 (46.8%) | 51 (42.7%) | 1 (33.3%) | |
| ECOG performance status, n (%) | | | | **.011*** |
| ECOG 0 | 93 (83.8%) | 91 (84.3%) | 2 (66.7%) | |
| ECOG 1 | 9 (8.1%) | 9 (8.3%) | 0 (0%) | |
| ECOG 2 | 6 (5.4%) | 6 (5.6%) | 0 (0%) | |
| ECOG 3 | 3 (2.7%) | 2 (1.9%) | 1 (33.3%) | |
| Status at initial diagnosis, n (%) | | | | **<.001**** |
| Localized melanoma | 99 (89.2%) | 98 (90.7%) | 1 (33.3%) | |
| Regional lymph node metastases | 7 (6.3%) | 5 (4.6%) | 2 (66.7%) | |
| Metastatic melanoma | 5 (4.5%) | 5 (4.6%) | 0 (0%) | |
| Melanoma stage, n (%) | | | | .871 |
| I | 3 (2.7%) | 3 (2.8%) | 0 (0%) | |
| II | 19 (17.1%) | 18 (16.7%) | 1 (33.3%) | |
| III | 85 (76.6%) | 83 (76.9%) | 2 (66.7%) | |
| IV | 4 (3.6%) | 4 (3.7%) | 0 (0%) | |
| Histogenetic type, n (%) | | | | .906 |
| Superficial spreading melanoma | 26 (23.4%) | 26 (24.1%) | 0 (0%) | |
| Nodular melanoma | 47 (42.3%) | 45 (41.7%) | 2 (66.7%) | |
| Lentigo maligna melanoma | 1 (0.9%) | 1 (0.9%) | 0 (0%) | |
| Acral lentiginous melanoma | 5 (4.5%) | 5 (4.6%) | 0 (0%) | |
| Amelanotic melanoma | 5 (4.5%) | 5 (4.6%) | 0 (0%) | |
| NA, other type | 27 (24.3%) | 26 (24.1%) | 1 (33.3%) | |
| Sentinel lymph node biopsy, n (%) | | | | **.001**** |
| Positive | 76 (68.5%) | 76 (70.4%) | 0 (0%) | |
| Negative | 14 (12.6%) | 14 (13.0%) | 0 (0%) | |
| NA | 21 (18.9%) | 18 (16.7%) | 3 (100%) | |
| Gene mutation status, n (%) | | | | .943 |
| BRAF | 9 (8.1%) | 9 (8.3%) | 0 (0%) | |
| NRAS | 6 (5.4%) | 6 (5.6%) | 0 (0%) | |
| cKIT | 1 (0.9%) | 1 (0.9%) | 0 (0%) | |
| wild type | 6 (5.4%) | 6 (5.6%) | 0 (0%) | |
| NA | 89 (80.2%) | 86 (79.6%) | 3(100%) | |
| Resection | | | | **.002**** |
| Performed | 10 (9.0%) | 8 (7.4%) | 2 (66.7%) | |
| Not performed | 2 (1.8%) | 2 (1.9%) | 0 (0%) | |
| NA | 99 (89.2%) | 98 (90.7%) | 1 (33.3%) | |
| Postresection residual cancer status, n (%) | | | | **<.001**** |
| R0 | 8 (7.2%) | 7 (6.5%) | 1 (33.3%) | |
| R1 | 2 (1.8%) | 1 (0.9%) | 1 (33.3%) | |
| NA | 101 (91.0%) | 100 (92.6%) | 1 (33.3%) | |
| Complete lymph node dissection, n (%) | | | | >.999 |
| Performed | 1 (0.9%) | 1 (0.9%) | 0 (0%) | |
| Not performed | 110 (99.1%) | 107 (99.1%) | 3 (100%) | |

Values are presented as median or number (%). Concordant cases include "concordant" cases and cases defined as "correct alternative recommendation" cases

*ECOG* Eastern Cooperative Oncology Group; *NA* not available (not relevant or" pending" for EO's query); *R0* no residual cancer; *R1* macroscopic residual cancer removed, while margins remain positive for microscopic residual cancer. Bold type numbers indicate statistical significance

**denotes that the p-value is significant at 1% level, and * that the p-value is significant at 5% level

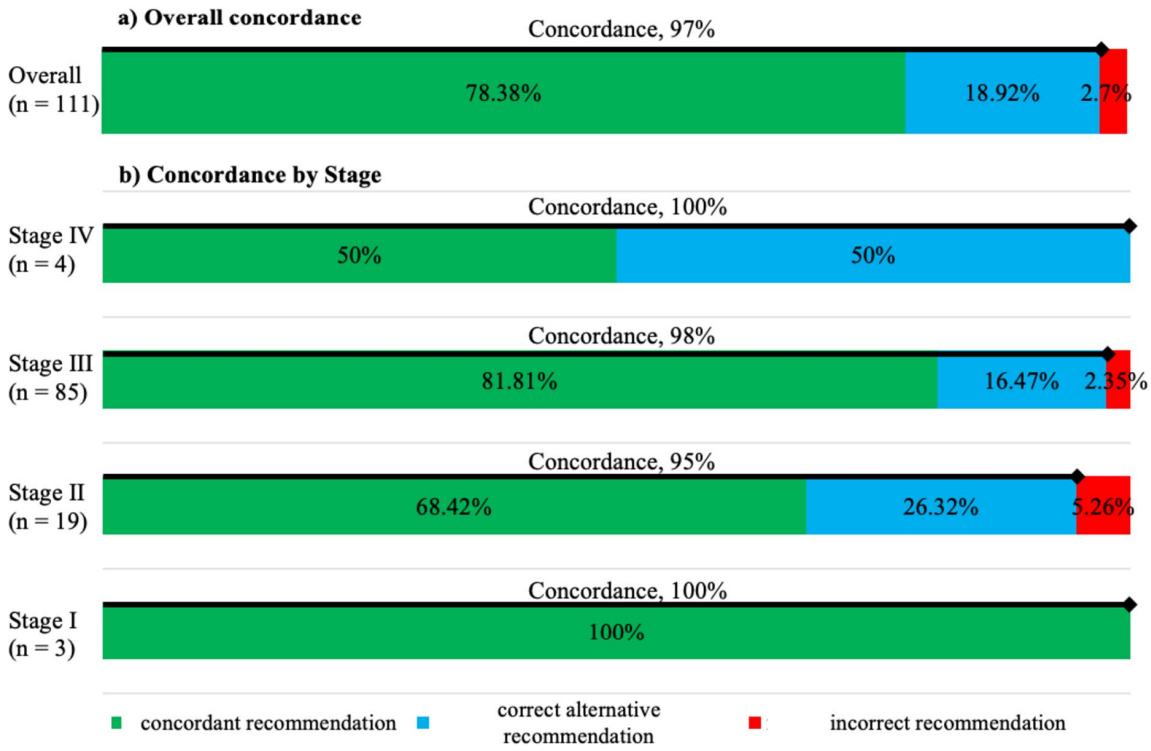


**Fig. 5** Treatment concordance between DSS and MDT. **a** Overall treatment concordance rates between the therapeutic recommendation given by the MDT and the treatment recommendation given by DSS for malignant melanoma. Overall concordance was 97%. **b** Treatment concordance rates according to malignant melanoma stages I, II, III, and IV were 100%, 95%, 98%, and 100%, respectively

(pT4b) showed either suspicious cervical lymph nodes or a solitary pulmonary nodule in CT imaging after primary resection. Because these findings remained unclear from a diagnostic perspective, the DSS advised that both cases need to be discussed in the MDT to find an individualized solution and no therapeutic recommendation was provided. The third "incorrect" case was a 80-year-old patient who presented with cutaneous satellite metastases. In accordance with the S3 guidelines for malignant melanoma (Dkg 2020), the DSS recommended a complete surgical resection of satellite metastases and adjuvant therapy. In contrast, MDT recommended an oncolytic viral immunotherapy with talimogene laherparepvec (T-VEC). According to the reviewing dermato-oncologist, this therapy was correctly recommended by MDT due to the advanced age of the patient and the difficult resection site of the melanoma. Even though our DSS provided the correct stage-specific therapy recommendation, its recommendation was not checked for its applicability due to patient-specific factors.

## Discussion

Treatment in certified tumor centers undeniably improves the quality of oncological care, and MDTs are one of the most important quality key features. In fact, however, it is relatively unclear to what extent the tumor boards have contributed to the improvement in survival rates in the certified centers (Devitt 2013; Keating 2013; Krasna 2013; Soukup 2019; Specchia 2020).

The certification requirement to present the majority of tumor cases in MDTs has led to a noticeable increase in the number of cases to be discussed in the conferences, which could have an unfavorable effect on the quality of recommendations (Soukup 2016; Walraven 2019).

Data evaluating the extent to which MDT improve outcomes remain mixed and some studies have found a survival benefit, while others have found no such advantage. A 2019 analysis indicated that the 5 year survival rate was 15.6% higher among cases in well-organized MDT but almost 20% lower in disorganized MDT compared with no MDT (Keating 2013; Kesson 2012; Lu 2019; Stone 2020; Wong 2022).

It seems almost surprising that AI-based systems still have not become established to support decision making

of MDT in routine clinical oncology. However, previous attempts to provide standardized treatment recommendations for the first-line treatment of tumor diseases using AI-based applications have shown too much uncertainty compared to the expertise of experienced oncologists. In several studies, for example, the use of Watson for Oncology could only achieve agreement rates of 12% to 93% in direct comparison with real MDT decisions (Zhou 2019; Choi 2019; Kim 2019; Lee 2018; Somashekhar 2018; Seidman 2015).

These previous AI approaches have been tested in many different tumor entities, but experience in decision support for melanoma therapy has not yet been published.

One important reason for the poor performance of AI systems is the lack of high-quality training datasets. They are available en masse as standardized data for AI systems used in image recognition, but not as regular cases in oncology care (McKinney et al. 2020; Ardila et al. 2019; Rubin 2019; Esteva et al. 2017).

In addition to the lack of well-organized and verified training data, another problem is the limited resource of experts who are initially required for the human interpretation and evaluation of the AI results.

An optimized workflow of the tumor conference is another crucial criterion for effectiveness. As requirements for an optimally structured conference, all information that is necessary for a therapy decision should first be available. In addition, those cases that require greater concentration due to their complexity should ideally be in the focus of the conference. Those medically simple cases for which clear recommendations can easily be derived from the guidelines should ideally be noted in the protocol as standard cases and only discussed optionally.

Here we see digital decision support systems as a suitable tool for structural improvement. By issuing guideline-compliant therapy recommendations at the time of conference registration, standard cases could be defined, and the conference could be unburdened accordingly.

In this respect, we see malignant melanoma as a suitable model disease for the development of decision support for tumor boards, as the melanoma treatment algorithm is considered comparatively complex.

Our dataset included 2399 cases with malignant diseases of the skin, of which 705 melanoma cases were discussed in the tumor board of the Skin Cancer Center of the University Hospital Cologne in the period from 2017 to 2020. This fits in well with the DKG annual reports of the certified skin cancer centers, which for example shows a melanoma proportion of 22.2% of all presented skin cancers for the year 2022 (Dkg 2022).

The fact that only 111 of the 705 cases with malignant melanoma could be included in the evaluation of our work is mainly due to the high proportion of recurrent diseases (377), non-cutaneous melanomas (78), and very complex disease patterns (43) that obviously cannot be declared as standard cases (see flowchart of patient case exclusion process, Fig 3). In addition, no therapy recommendation could be automatically derived in 72 cases due to insufficient information relevant for decision making.

As a result, we found three cases with divergent standard recommendations. Of these, two cases were rated as non-compliant in accordance with the protocol, as no automated recommendation was made at all, but rather a presentation to the tumor board was correctly recommended.

An important finding was the identification of a case in which the recommendation deviated due to the patient's advanced age. This implies to adapt the query algorithm to assess age in these clinical constellations so that systemic treatment can be critically discussed with the patient, especially in adjuvant treatment.

The present work has several limitations that should be considered.

As the first limitation, we only selected first-line cases and thus missed a large number of discussions on relapsed cases with first metastatic disease. In addition, certification requirements only request the presentation of melanoma cases in stages IIC and higher, leading to a low number of stage I melanoma cases.

As we included cases that were presented to MDT for the first time during the clinical course, there were few cases with initial metastatic disease. Thus, concordance rates for stage I and stage IV were calculated based on a small number of patient cases (3 and 4 patient cases, respectively). For these stages, no general conclusions can be drawn and further validation is necessary.

The second limitation derives from our strict inclusion criteria, which intended a preselection of cases that were most likely standard cases. As a consequence only 16 % of all cases were selected for detailed comparison, most of them stage II and III. If the high proportion of up to 50% stage II and III melanoma cases is taken into account, which displays the reality of MDT in the annual DKG reports, the potential to unburden MDT by DSS becomes more evident. Our evaluation also indicates 10% of cases ($n = 72$) for which the DSS was unable to issue a recommendation due to missing information. If treatment-relevant information would be obtained more systematically by a DSS at the time of conference registration, a further reduction in the burden could thus be easily derived.

As a third limitation, S3 guidelines for malignant melanoma were updated four times (DKG 2020) and numerous new therapeutics received approval during study period between 2017 and 2020. These changes in the guidelines led to deviations in the recommendations, for example on the role of interferon therapy. Of note, the presented data are based on a software (EO version 5.06) that displayed the 8th edition of the AJCC classification for

melanoma and remained unchanged during the study period. Changes from 7th to 8th version in 2018 did not affect cases included in this study.

The development of the DSS in collaboration with the skin cancer center of the University Hospital Cologne could indicate a performance bias and the risk of overfitting. In further studies validation has to be performed by a multicentric approach. However, it is important to emphasize that the decision logic is based on the S3 guideline for malignant melanoma (Dkg 2020) and the approval status of new therapeutics that are used in best clinical practice. The continuous comparison of real world to digital decisions ensures that new treatment concepts are quickly detected, and thus, enables EO's expert curators to integrate the latest medical standard and to appropriately adjust the query algorithm.

As a statistical limitation, the research methodology used to determine concordance rates is only of descriptive nature and describes the degree of agreement between DSS and MDT. At this time, no conclusion of the clinical benefit using DSS, such as overall survival or progression free survival of the patients, can be drawn from this approach.

The advantages we show with this work could not only provide substantial support in everyday clinical practice, but also provide a basis for the later integration of AI-based systems.

As a first advantage, the proposed principle allows an automatic structuring of the tumor conferences according to complexity and enables a quality assurance of the recommendations given by automatic comparison with the guidelines of the medical societies.

Second, the system ensures that all information necessary for a therapeutic decision is already requested at the time of registration for the conference. Otherwise, missing information often leads to postponement or only very vague recommendations such as "indication for systemic therapy."

The repetitive questioning of decision-relevant information that occurs with repeated use leads to the third advantage, the teaching effect. This could also be reinforced by the sense of satisfaction that comes from having received a (digital) treatment recommendation.

Not least, another advantage is that the type of recommendation matching described above generates training data for the future integration of AI systems. As said before, a machine learning tool can only ever be as good as the data available for training and the trainers who evaluate the AI results.

## Summary

It must be underlined again that the aim of our work is not to provide digital recommendations for all questions addressed to multidisciplinary tumor boards or to replace clinical experience. However, we see the need to relieve the time burden of these critical conferences so that the participants can focus their expertise on the more complex tumor cases. Our results suggest that this automated approach would allow a more concentrated and detailed discussion of complex tumor cases, on which the valuable expertise of the board members should be focused. In addition, the implementation of the presented algorithm in the routine software of tumor boards could provide the basis for transparent and comparable quality management.

## Perspective

The principle of digital decision support described in this paper, whose algorithmic query is based on a decision tree, may seem unspectacular at first glance against the background of current AI development. However, it is currently the most suitable way to make recommendations for clinical situations that are already defined by clear therapeutic standards and guidelines. These commonly accepted therapeutic recommendations are based, among other things, on the approval status and availability of the therapeutic agents, as well as some clinical and structural aspects which are not evidence-based per se. It should therefore come as no surprise that the recommendations made by AI-based systems for well-defined standard clinical situations often do not correspond to standard practice. At this point, it should be considered that the AI recommendations could well be a better therapeutic choice, though this cannot be proven due to the AI black box effect described above.

Accordingly, it is to be expected that AI applications will initially realize their full potential in complex clinical constellations of advanced cancer diseases in which clinical standards do not exist.

However, this requires the availability of sufficient training data generated under the requirements of a defined healthcare system in routine care. This is precisely the data that are currently lacking, however, training data from other countries and healthcare systems cannot be used without hesitation.

The intended integration of query algorithms into the organizational software of tumor boards offers an important cornerstone here. The data collected on standard situations in oncological care can serve as the necessary training data for the planned integration of suitable AI models.

However, there is still a long way to go before AI delivers such impressive results in decision making as we are currently seeing with image-supported AI systems.

**Acknowledgements** We would like to thank Andreas Greeske for the programming of the software and Jonas Heidelbach for the development and implementation of the medical algorithms used and the critical quality assurance.

**Author contributions** D.H. and P.K. contributed with data curation, formal analysis, investigation, methodology, resources, validation, visualization, writing—original draft. T.E. contributed with project administration and software CM., NK., and C.F. contributed to data curation, supervision, and writing—review & editing. C.G., E.S., contributed to writing—review and editing. M.H. contributed with resources, and supervision.

**Funding** The authors received no specific funding for this work.

**Data availability** All data contained in the manuscript as well as the primary metadata are based on the tumor board documentation of the University Hospital Cologne. The data were evaluated in a pseudonymized form in compliance with data protection guidelines and can be assigned to the real cases by the treating physicians who are permitted to inquire about them.

## Declarations

**Conflict of interest** T.E. is the founder and chief scientific officer of the company Onqo Health GmbH who developed the application "Easy-Oncology." D.H. is an employee of Onqo Health GmbH. C.G. is an employee of Onqo Health GmbH. All other authors declare no conflicts of interest regarding the submitted work.

**Ethical approval** Ethical approval to conduct this work was granted by the Ethics Committee of the Medical Faculty of the University of Cologne (#20-1116).



## References

Amaria RN, Postow M, Burton EM et al (2022) Neoadjuvant relatlimab and nivolumab in resectable melanoma. Nature 611(7934):155–60. https://doi.org/10.1038/s41586-022-05368-8

Ardila D, Kiraly AP, Bharadwaj S et al (2019) End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med 25(6):954–61. https://doi.org/10.1038/s41591-019-0447-x

Beckmann MW, Brucker C, Hanf V et al (2011) Quality assured health care in certified breast centers and improvement of the prognosis of breast cancer patients. Onkologie 34(7):362–7. https://doi.org/10.1159/000329601

Bierbaum M, Rapport F, Arnolda G et al (2020) Clinicians' attitudes and perceived barriers and facilitators to cancer treatment clinical practice guideline adherence: a systematic review of qualitative and quantitative literature. Implement Sci 15(1):39. https://doi.org/10.1186/s13012-020-00991-3

Bierbaum M, Rapport F, Arnolda G et al (2023) Rates of adherence to cancer treatment guidelines in Australia and the factors associated with adherence A systematic review. Asia Pac J Clin Oncol. https://doi.org/10.1111/ajco.13948.10.1111/ajco.13948

Bungartz KD, Lalowski K, Elkin SK (2018) Making the right calls in precision oncology. Nat Biotechnol 36(8):692–6. https://doi.org/10.1038/nbt.4214

Calero JJ, Oton LF, Oton CA (2017) Apps for Radiation Oncology A Comprehensive Review. Transl Oncol 10(1):108–14. https://doi.org/10.1016/j.tranon.2016.08.008

Chen Y, Elenee Argentinis JD, Weber G (2016) IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. Clin Ther 38(4):688–701. https://doi.org/10.1016/j.clinthera.2015.12.001

Choi YI, Chung JW, Kim KO et al (2019) Concordance Rate between Clinicians and Watson for Oncology among Patients with Advanced Gastric Cancer: Early, Real-World Experience in Korea. Can J Gastroenterol Hepatol 2019:8072928. https://doi.org/10.1155/2019/8072928

Devitt B, Philip J, McLachlan SA (2013) Re: Tumor boards and the quality of cancer care. J Natl Cancer Inst 105(23):1838. https://doi.org/10.1093/jnci/djt311

DKG (2020) Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF): Diagnostik, Therapie und Nachsorge des Melanoms, Langversion 3.3, 2020, AWMF Registernummer: 032/024OL http://www.leitlinienprogramm-onkologie.de/leitlinien/melanom/. „Accessed"14.09.2022

DKG (2019) Nationales Zertifizierungsprogramm Krebs. Erhebungsbogen für Onkologische Spitzenzentren und Onkologische Zentren Deutsche Krebsgesellschaft und Deutsche Krebshilfe https://www.krebsgesellschaft.de/zertdokumente.html. „Accessed" 14.09.2022

DKG (2022) DKG Jahresbericht der zertifizierten Hautkrebszentren. Kennzahlenauswertung 2022 /Auditjahr 2021 / Kennzahlenjahr 2020.

Esteva A, Kuprel B, Novoa RA et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–8. https://doi.org/10.1038/nature21056

Heins MJ, de Jong JD, Spronk I et al (2017) Adherence to cancer treatment guidelines: influence of general and cancer-specific guideline characteristics. Eur J Public Health 27(4):616–20. https://doi.org/10.1093/eurpub/ckw234

IQVIA (2022). Institute I https://www.iqvia.com/insights/the-iqvia-institute/reports/global-oncology-trends-2022 „Accessed" 10.06.2022

Jalil R, Soukup T, Akhter W et al (2018) Quality of leadership in multidisciplinary cancer tumor boards: development and evaluation of a leadership assessment instrument (ATLAS). World J Urol 36(7):1031–8. https://doi.org/10.1007/s00345-018-2255-1

Keating NL, Landrum MB, Lamont EB et al (2013) Tumor boards and the quality of cancer care. J Natl Cancer Inst 105(2):113–21. https://doi.org/10.1093/jnci/djs502

Kesson EM, Allardice GM, George WD et al (2012) Effects of multidisciplinary team working on breast cancer survival: retrospective, comparative, interventional cohort study of 13 722 women. Bmj 344:e2718. https://doi.org/10.1136/bmj.e2718

Kim EJ, Woo HS, Cho JH et al (2019) Early experience with Watson for oncology in Korean patients with colorectal cancer. PLoS One 14(3):e0213640. https://doi.org/10.1371/journal.pone.0213640

Kim M, Kim BH, Kim JM et al (2019) Concordance in postsurgical radioactive iodine therapy recommendations between Watson for Oncology and clinical practice in patients with differentiated thyroid carcinoma. Cancer 125(16):2803–9. https://doi.org/10.1002/cncr.32166

Kowalski C, Graeven U, von Kalle C et al (2017) Shifting cancer care towards Multidisciplinarity: the cancer center certification program of the German cancer society. BMC Cancer 17(1):850. https://doi.org/10.1186/s12885-017-3824-1

Kowalski C, Sibert NT, Breidenbach C et al (2022) Outcome Quality After Colorectal Cancer Resection in Certified Colorectal Cancer Centers—Patient-Reported and Short-Term Clinical Outcomes. Dtsch Arztebl Int. https://doi.org/10.3238/arztebl.m2022.0325

Krasna M, Freeman RK, Petrelli NJ (2013) Re: Tumor boards and the quality of cancer care. J Natl Cancer Inst 105(23):1839–40. https://doi.org/10.1093/jnci/djt313

Lee WS, Ahn SM, Chung JW et al (2018) Assessing Concordance With Watson for Oncology, a Cognitive Computing Decision Support System for Colon Cancer Treatment in Korea. JCO Clin Cancer Inform 2:1–8. https://doi.org/10.1200/cci.17.00109

Letzen B, Wang CJ, Chapiro J (2019) The Role of Artificial Intelligence in Interventional Oncology: A Primer. J Vasc Interv Radiol 30(1):38-41.e1. https://doi.org/10.1016/j.jvir.2018.08.032

Lu J, Jiang Y, Qian M et al (2019) The Improved Effects of a Multidisciplinary Team on the Survival of Breast Cancer Patients: Experiences from China. Int J Environ Res Public Health. https://doi.org/10.3390/ijerph17010277

Luke JJ, Rutkowski P, Queirolo P et al (2022) Pembrolizumab versus placebo as adjuvant therapy in completely resected stage IIB or IIC melanoma (KEYNOTE-716): a randomised, double-blind, phase 3 trial. Lancet 399(10336):1718–29. https://doi.org/10.1016/s0140-6736(22)00562-1

McKinney SM, Sieniek M, Godbole V et al (2020) International evaluation of an AI system for breast cancer screening. Nature 577(7788):89–94. https://doi.org/10.1038/s41586-019-1799-6

Modabber A, Schick D, Goloborodko E et al (2021) Impact of quality certification of multidisciplinary head and neck tumor centers. Cost Eff Resour Alloc 19(1):20. https://doi.org/10.1186/s12962-021-00273-9

Narang J, Hue JJ, Bingmer K et al (2021) Sentinel lymph node biopsy guideline concordance in melanoma: Analysis of the National Cancer Database. J Surg Oncol 124(4):669–78. https://doi.org/10.1002/jso.26565

Rubin R (2019) Artificial Intelligence for Cervical Precancer Screening. Jama 321(8):734. https://doi.org/10.1001/jama.2019.0888

Schmidt C (2017) MD Anderson Breaks With IBM Watson Raising Questions About Artificial Intelligence in Oncology. J Natl Cancer Inst. https://doi.org/10.1093/jnci/djx113

Schmitt J, Schoffer O, Klinkhammer-Schalke M, Bobeth C, Roessler M, et al (2022) Wirksamkeit der Versorgung in onkologischen Zentren(WiZen)–Erkenntnisse zur Ergebnisqualität und Erfolg des Datenlinkage AOK „Accessed“ 20.08.23

Seidman AD, Pilewskie ML, Robson ME et al (2015) Integration of multi-modality treatment planning for early stage breast cancer (BC) into Watson for Oncology, a Decision Support System: Seeing the forest and the trees. J Clin Oncol 33(15):e12042. https://doi.org/10.1200/jco.2015.33.15_suppl.e12042

Somashekhar SP, Sepúlveda MJ, Puglielli S et al (2018) Watson for Oncology and breast cancer treatment recommendations: agreement with an expert multidisciplinary tumor board. Ann Oncol 29(2):418–23. https://doi.org/10.1093/annonc/mdx781

Soukup T, Petrides KV, Lamb BW et al (2016) The anatomy of clinical decision-making in multidisciplinary cancer meetings: A cross-sectional observational study of teams in a natural context. Medicine (Baltimore) 95(24):e3885. https://doi.org/10.1097/md.0000000000003885

Soukup T, Lamb BW, Weigl M et al (2019) An Integrated Literature Review of Time-on-Task Effects With a Pragmatic Framework for Understanding and Improving Decision-Making in Multidisciplinary Oncology Team Meetings. Front Psychol 10:1245. https://doi.org/10.3389/fpsyg.2019.01245

Specchia ML, Frisicale EM, Carini E et al (2020) The impact of tumor board on cancer care: evidence from an umbrella review. BMC Health Serv Res 20(1):73. https://doi.org/10.1186/s12913-020-4930-3

Stone E, Rankin N, Currow D et al (2020) Optimizing lung cancer MDT data for maximum clinical impact-a scoping literature review. Transl Lung Cancer Res 9(4):1629–38. https://doi.org/10.21037/tlcr.2020.01.02

Swetter SM, Thompson JA, Albertini MR et al (2021) NCCN Guidelines® Insights Melanoma Cutaneous Version 22021. J Natl Compr Canc Netw 19(4):364–76. https://doi.org/10.6004/jnccn.2021.0018

Walraven JEW, Desar IME, van der Hoeven JJM et al (2019) Analysis of 105.000 patients with cancer: have they been discussed in oncologic multidisciplinary team meetings? A nationwide population-based study in the Netherlands. Eur J Cancer 121:85–93. https://doi.org/10.1016/j.ejca.2019.08.007

Wolff KD, Rau A, Ferencz J et al (2017) Effect of an evidence-based guideline on the treatment of maxillofacial cancer: A prospective analysis. J Craniomaxillofac Surg 45(3):427–31. https://doi.org/10.1016/j.jcms.2016.12.013

Wong BO, Blythe JA, Wu A et al (2022) Exploration of Clinician Perspectives on Multidisciplinary Tumor Board Function Beyond Clinical Decision-making. JAMA Oncol 8(8):1210–2. https://doi.org/10.1001/jamaoncol.2022.1763

Yu SH, Kim MS, Chung HS et al (2021) Early experience with Watson for Oncology: a clinical decision-support system for prostate cancer treatment recommendations. World J Urol 39(2):407–13. https://doi.org/10.1007/s00345-020-03214-y

Zhou N, Zhang CT, Lv HY et al (2019) Concordance Study Between IBM Watson for Oncology and Clinical Practice for Patients with Cancer in China. Oncologist 24(6):812–9. https://doi.org/10.1634/theoncologist.2018-0255

## Authors and Affiliations

**David Hoier[1] · Carolin Groß-Ophoff-Müller[3] · Cindy Franklin[2] · Michael Hallek[1] · Esther von Stebut[2] · Thomas Elter[1] · Cornelia Mauch[2,4] · Nicole Kreuzberg[2] · Philipp Koll[2]**

✉ David Hoier
davidhoier@hotmail.com

Carolin Groß-Ophoff-Müller
Grossophoffmuellerc@kliniken-koeln.de

Cindy Franklin
cindy.franklin@uk-koeln.de

Michael Hallek
michael.hallek@uk-koeln.de

Esther von Stebut
esther.von-stebut@uk-koeln.de

Thomas Elter
thomas.elter@uk-koeln.de

Cornelia Mauch
cornelia.mauch@klinikum-bochum.de

Nicole Kreuzberg
nicole.kreuzberg@uk-koeln.de

Philipp Koll
philipp.koll@uk-koeln.de

[1] Department I of Internal Medicine, Center for Integrated Oncology Aachen Bonn Cologne Duesseldorf, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany

[2] Department of Dermatology and Venereology, Center for Integrated Oncology Aachen Bonn Cologne Duesseldorf, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany

[3] Kliniken der Stadt Köln gGmbH, Lung Clinic, Cologne, Germany

[4] Department of Dermatology, Venereology and Allergology, Ruhr-University Bochum, Bochum, Germany

# 4. Erweiterte Diskussion – DSS in multidisziplinären Tumorboards

Die Zertifizierung von Tumorzentren erfolgt mit dem Ziel, die optimale und leitliniengerechte Versorgung von Krebspatienten sicherzustellen. Der Gemeinsame Bundesausschuss (G-BA) hat vorgegeben, dass zukünftig alle Kliniken, die Krebspatienten behandeln, zur Zertifizierung durch die Deutsche Krebsgesellschaft (DKG) verpflichtet werden. Die Zertifizierung gilt hierbei als Voraussetzung für die Kostenerstattung durch die Krankenkassen. Die Durchführung interdisziplinärer Tumorkonferenzen für jeden Behandlungsfall kann hierbei als zentrales Qualitätsmerkmal zur Zertifizierung benannt werden.

Eine strukturierte Erfassung und detaillierte Dokumentation von Krankheitsverläufen und der im Tumorboard erteilten Therapieempfehlungen gehören hierbei zu den zentralen Zertifizierungsanforderungen. Bei einem beträchtlichen Anteil der aktuell 1.960 zertifizierten Zentren erfolgt die Tumorkonferenz-Dokumentation durch heterogene Software-Eigenentwicklungen.[25] Die Erfassung der diskutierten Fälle erfolgt in diesen nicht standardisierten Systemen oft unstrukturiert, teils während oder erst nach Abschluss der Tumorkonferenz. Dieses Vorgehen birgt die Gefahr, dass wertvolle Informationen zum Erkrankungsverlauf oder der Entscheidungsfindung verloren bzw. nicht ausreichend genau erfasst werden.

Die Übermittlung dieser in den Kliniken erfassten Daten in die Register der Deutschen Krebsgesellschaft erfolgt durch Dokumentationssoftware, die wiederum eine Lizenzierung durch das unabhängige Institut OnkoZert erfordert.

Vor diesem Hintergrund stellt sich die naheliegende Frage, ob und in welchem Maße Entscheidungsunterstützungssysteme neben der Qualitätssicherung von Therapieempfehlungen dazu beitragen können, die zur Zertifizierung erforderlichen Dokumentationspflichten zu entlasten. Die Vorgabe, alle erstdiagnostizierten Tumorfälle im MDT vorzustellen, führt zu einem deutlichen Anstieg der Fallbesprechungen. Dass die hierdurch limitierte Zeit pro Falldiskussion von gewöhnlich nur zwei bis sechs Minuten sich nachteilig auf die Qualität der Therapieempfehlungen auswirkt, lässt sich erahnen.[26, 27] Zeitlicher Druck kann zudem auch durch fehlende, Behandlungs-relevante Informationen entstehen, was zu zeitraubenden und oftmals erfolglosen Falldiskussionen mit der Empfehlung zur Wiedervorstellung führt.[26, 28] Neben der steigenden Anzahl der zu diskutierenden Tumorfälle und teils fehlenden Informationen, ist unter dem Aspekt der zunehmenden Komplexität der onkologischen Diagnostik und der rasanten Entwicklung zielgerichteter und personalisierter Therapieoptionen auch die potenziell nicht ausreichende Expertise der Tumorboard-Teilnehmer zu berücksichtigen.[6, 28-31] Es ist daher kaum verwunderlich, dass diese Faktoren in der klinischen Praxis nicht konsequent zu einer verbesserten Qualität der erteilten Empfehlung führen.[27, 32-34]

Die steigende Bedeutung evidenzbasierter Therapieempfehlungen in der onkologischen Versorgung ist offensichtlich und unterstreicht die Notwendigkeit, Qualitätsstandards in multidisziplinären Tumorboards zu etablieren. Gegenwärtig existiert bedauerlicherweise kein geeignetes Instrument, um die Qualität dieser entscheidenden Konferenzen zu erfassen und zu verbessern. Die derzeit etablierten Tumorboard Software-Lösungen dienen lediglich der strukturellen Organisation und Dokumentation, bieten aber keine digitale Unterstützung bei der Entscheidungsfindung.

Bisherige Bestrebungen, flächendeckend KI-unterstützte Systeme einzuführen, auch von namhaften Herstellern wie Roche, Siemens Healthineers und Phillips, sind an unterschiedlichen Herausforderungen gescheitert. Dazu gehören hohe Kosten, eine fragmentierte Systemlandschaft, Datenschutzprobleme und begrenzte IT-Ressourcen in der medizinischen Versorgung. Zudem besteht eine Intransparenz der Therapieempfehlungen dieser KI-Systeme aufgrund des Black-Box-Effekts, was potenziell zu erheblichen qualitativen Schwankungen bei deren Anwendung in der Praxis führen kann.[35-37] Nicht zuletzt scheitern die bisherigen Systeme aufgrund ungenügender Daten für Therapieempfehlungen, die entweder nicht in maschinenlesbarer Form vorliegen oder zum Zeitpunkt der Tumorkonferenz ausschließlich der behandelnden Ärzte bekannt sind.

In dieser Arbeit wurde ein Lösungsansatz basierend auf einem qualitätsgesicherten, von Experten kuratierten DSS vorgestellt, welches letztendlich auf einem Entscheidungsbaum basiert. Dies könnte angesichts aktueller KI-Entwicklungen unspektakulär erscheinen. Es ist jedoch die derzeit am besten geeignete Methode, um digitale Therapieempfehlungen für Melanompatienten zu geben, deren Behandlungsempfehlungen sich aus eindeutigen therapeutischen Standards und Leitlinien mit größter Sicherheit herleiten lassen. Bemerkenswerterweise kann ein Großteil der primären Melanom-Patienten in eben diese Kategorie eingeordnet werden. Von 15.838 behandelten primären Melanom-Patienten, die im Jahr 2022 in zertifizierten Tumorzentren behandelt wurden, waren 12.901 Fälle entweder dem Stadium I oder II zuzuordnen.[38] Zukünftig könnten die gewonnenen digitalen Empfehlungen zu “Standardfällen” als Trainingsdaten zur Entwicklung verlässlicher KI-Systemen dienen.

Die Herausforderung besteht nun darin, die in dieser Arbeit validierte DSS in der flächendeckenden onkologischen Versorgung zu etablieren. Hierzu erfolgt eine Zusammenarbeit mit Herstellern von bereits etablierter Tumorboard-Software, um den DSS-Algorithmus in die Anmeldemasken der Tumorkonferenzen zu integrieren. Diese Herangehensweise wird zukünftig weitere Erkenntnisse über die Qualität und Praktikabilität von Entscheidungsunterstützungssystemen in der klinischen Praxis liefern und möglicherweise die Versorgung onkologischer Patienten verbessern.

# 5. Literaturverzeichnis


1. Keating NL, Landrum MB, Lamont EB, Bozeman SR, Shulman LN, McNeil BJ. Tumor boards and the quality of cancer care. *J Natl Cancer Inst*. Jan 16 2013;105(2):113-21. doi:10.1093/jnci/djs502
2. Bierbaum M, Rapport F, Arnolda G, Nic Giolla Easpaig B, Lamprell K, Hutchinson K, Delaney GP, Liauw W, Kefford R, Olver I, Braithwaite J. Clinicians' attitudes and perceived barriers and facilitators to cancer treatment clinical practice guideline adherence: a systematic review of qualitative and quantitative literature. *Implement Sci*. May 27 2020;15(1):39. doi:10.1186/s13012-020-00991-3
3. Bierbaum M, Rapport F, Arnolda G, Tran Y, Nic Giolla Easpaig B, Ludlow K, Clay-Williams R, Austin E, Laginha B, Lo CY, Churruca K, van Baar L, Hutchinson K, Chittajallu R, Owais SS, Nullwala R, Hibbert P, Fajardo Pulido D, Braithwaite J. Rates of adherence to cancer treatment guidelines in Australia and the factors associated with adherence: A systematic review. *Asia Pac J Clin Oncol*. Dec 2023;19(6):618-644. doi:10.1111/ajco.13948
4. Heins MJ, de Jong JD, Spronk I, Ho VKY, Brink M, Korevaar JC. Adherence to cancer treatment guidelines: influence of general and cancer-specific guideline characteristics. *Eur J Public Health*. Aug 1 2017;27(4):616-620. doi:10.1093/eurpub/ckw234
5. Narang J, Hue JJ, Bingmer K, Hardacre JM, Winter JM, Ocuin LM, Ammori JB, Mangla A, Bordeaux J, Rothermel LD. Sentinel lymph node biopsy guideline concordance in melanoma: Analysis of the National Cancer Database. *J Surg Oncol*. Sep 2021;124(4):669-678. doi:10.1002/jso.26565
6. Soukup T, Lamb BW, Weigl M, Green JSA, Sevdalis N. An Integrated Literature Review of Time-on-Task Effects With a Pragmatic Framework for Understanding and Improving Decision-Making in Multidisciplinary Oncology Team Meetings. *Front Psychol*. 2019;10:1245. doi:10.3389/fpsyg.2019.01245
7. Chen Y, Elenee Argentinis JD, Weber G. IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. *Clin Ther*. Apr 2016;38(4):688-701. doi:10.1016/j.clinthera.2015.12.001
8. Letzen B, Wang CJ, Chapiro J. The Role of Artificial Intelligence in Interventional Oncology: A Primer. *J Vasc Interv Radiol*. Jan 2019;30(1):38-41.e1. doi:10.1016/j.jvir.2018.08.032
9. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. Feb 2 2017;542(7639):115-118. doi:10.1038/nature21056

10. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G, Naidich DP, Shetty S. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. Jun 2019;25(6):954-961. doi:10.1038/s41591-019-0447-x
11. Rubin R. Artificial Intelligence for Cervical Precancer Screening. *Jama*. Feb 26 2019;321(8):734. doi:10.1001/jama.2019.0888
12. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesus M, Corrado GS, Darzi A, Etemadi M, Garcia-Vicente F, Gilbert FJ, Halling-Brown M, Hassabis D, Jansen S, Karthikesalingam A, Kelly CJ, King D, Ledsam JR, Melnick D, Mostofi H, Peng L, Reicher JJ, Romera-Paredes B, Sidebottom R, Suleyman M, Tse D, Young KC, De Fauw J, Shetty S. International evaluation of an AI system for breast cancer screening. *Nature*. Jan 2020;577(7788):89-94. doi:10.1038/s41586-019-1799-6
13. Schmidt C. M. D. Anderson Breaks With IBM Watson, Raising Questions About Artificial Intelligence in Oncology. *J Natl Cancer Inst*. May 1 2017;109(5)doi:10.1093/jnci/djx113
14. Oehring R, Ramasetti N, Ng S, Roller R, Thomas P, Winter A, Maurer M, Moosburner S, Raschzok N, Kamali C, Pratschke J, Benzing C, Krenzien F. Use and accuracy of decision support systems using artificial intelligence for tumor diseases: a systematic review and meta-analysis. *Front Oncol*. 2023;13:1224347. doi:10.3389/fonc.2023.1224347
15. Somashekhar SP, Sepúlveda MJ, Puglielli S, Norden AD, Shortliffe EH, Rohit Kumar C, Rauthan A, Arun Kumar N, Patil P, Rhee K, Ramya Y. Watson for Oncology and breast cancer treatment recommendations: agreement with an expert multidisciplinary tumor board. *Ann Oncol*. Feb 1 2018;29(2):418-423. doi:10.1093/annonc/mdx781
16. Kim EJ, Woo HS, Cho JH, Sym SJ, Baek JH, Lee WS, Kwon KA, Kim KO, Chung JW, Park DK, Kim YJ. Early experience with Watson for oncology in Korean patients with colorectal cancer. *PLoS One*. 2019;14(3):e0213640. doi:10.1371/journal.pone.0213640
17. Lee WS, Ahn SM, Chung JW, Kim KO, Kwon KA, Kim Y, Sym S, Shin D, Park I, Lee U, Baek JH. Assessing Concordance With Watson for Oncology, a Cognitive Computing Decision Support System for Colon Cancer Treatment in Korea. *JCO Clin Cancer Inform*. Dec 2018;2:1-8. doi:10.1200/cci.17.00109
18. Zhou N, Zhang CT, Lv HY, Hao CX, Li TJ, Zhu JJ, Zhu H, Jiang M, Liu KW, Hou HL, Liu D, Li AQ, Zhang GQ, Tian ZB, Zhang XC. Concordance Study Between IBM Watson for Oncology and Clinical Practice for Patients with Cancer in China. *Oncologist*. Jun 2019;24(6):812-819. doi:10.1634/theoncologist.2018-0255

19. Choi YI, Chung JW, Kim KO, Kwon KA, Kim YJ, Park DK, Ahn SM, Park SH, Sym SJ, Shin DB, Kim YS, Sung KH, Baek JH, Lee U. Concordance Rate between Clinicians and Watson for Oncology among Patients with Advanced Gastric Cancer: Early, Real-World Experience in Korea. *Can J Gastroenterol Hepatol*. 2019;2019:8072928. doi:10.1155/2019/8072928
20. Seidman AD, Pilewskie ML, Robson ME, Kelvin JF, Zauderer MG, Epstein AS, Kris MG, Fu J, Keesing J, Caroline A, Megerian M, Eggebraaten T, DeLima R, Setnes M, Barker K, Gucalp A. Integration of multi-modality treatment planning for early stage breast cancer (BC) into Watson for Oncology, a Decision Support System: Seeing the forest and the trees. *Journal of Clinical Oncology*. 2015;33(15_suppl):e12042-e12042. doi:10.1200/jco.2015.33.15_suppl.e12042
21. Tian Y, Liu X, Wang Z, Cao S, Liu Z, Ji Q, Li Z, Sun Y, Zhou X, Wang D, Zhou Y. Concordance Between Watson for Oncology and a Multidisciplinary Clinical Decision-Making Team for Gastric Cancer and the Prognostic Implications: Retrospective Study. *J Med Internet Res*. Feb 20 2020;22(2):e14122. doi:10.2196/14122
22. Liu C, Liu X, Wu F, Xie M, Feng Y, Hu C. Using Artificial Intelligence (Watson for Oncology) for Treatment Recommendations Amongst Chinese Patients with Lung Cancer: Feasibility Study. *J Med Internet Res*. Sep 25 2018;20(9):e11087. doi:10.2196/11087
23. Kim M, Kim BH, Kim JM, Kim EH, Kim K, Pak K, Jeon YK, Kim SS, Park H, Kang T, Lee BJ, Kim IJ. Concordance in postsurgical radioactive iodine therapy recommendations between Watson for Oncology and clinical practice in patients with differentiated thyroid carcinoma. *Cancer*. Aug 15 2019;125(16):2803-2809. doi:10.1002/cncr.32166
24. Bungartz KD, Lalowski K, Elkin SK. Making the right calls in precision oncology. *Nature Biotechnology*. 2018/09/01 2018;36(8):692-696. doi:10.1038/nbt.4214
25. DKG. OncoMap. Accessed 27.02.2024, https://www.oncomap.de/centers?selectedCountries=[Deutschland]&selectedCerttypes=[DKG]&showMap=1
26. Lamb BW, Brown KF, Nagpal K, Vincent C, Green JS, Sevdalis N. Quality of care management decisions by multidisciplinary cancer teams: a systematic review. *Ann Surg Oncol*. Aug 2011;18(8):2116-25. doi:10.1245/s10434-011-1675-6
27. Hammer RD, Fowler D, Sheets LR, Siadimas A, Guo C, Prime MS. A digital tumor board solution impacts case discussion time and postponement of cases in tumor boards. *Health and Technology*. 2021/05/01 2021;11(3):525-533. doi:10.1007/s12553-021-00533-x

28. Jalil R, Soukup T, Akhter W, Sevdalis N, Green JSA. Quality of leadership in multidisciplinary cancer tumor boards: development and evaluation of a leadership assessment instrument (ATLAS). *World J Urol*. Jul 2018;36(7):1031-1038. doi:10.1007/s00345-018-2255-1
29. Walraven JEW, Desar IME, Hoeven van der JJM, Aben KKH, Hillegersberg van R, Rasch CRN, Lemmens V, Verhoeven RHA. Analysis of 105.000 patients with cancer: have they been discussed in oncologic multidisciplinary team meetings? A nationwide population-based study in the Netherlands. *Eur J Cancer*. Nov 2019;121:85-93. doi:10.1016/j.ejca.2019.08.007
30. Merry D, Schickhardt C, Mehlis K, Winkler EC. Trust and responsibility in molecular tumour boards. *Bioethics*. Sep 2018;32(7):464-472. doi:10.1111/bioe.12464
31. Haier J. Aufgaben und Grenzen von Tumorkonferenzen. *Der Onkologe*. 2016/03/01 2016;22(3):184-191. doi:10.1007/s00761-015-2939-8
32. Hammer RD, Fowler D, Sheets LR, Siadimas A, Guo C, Prime MS. Digital Tumor Board Solutions Have Significant Impact on Case Preparation. *JCO Clin Cancer Inform*. Aug 2020;4:757-768. doi:10.1200/cci.20.00029
33. Roessler M, Schmitt J, Bobeth C, Gerken M, Kleihues-van Tol K, Reissfelder C, Rau BM, Distler M, Piso P, Günster C, Klinkhammer-Schalke M, Schoffer O, Bierbaum V. Is treatment in certified cancer centers related to better survival in patients with pancreatic cancer? Evidence from a large German cohort study. *BMC Cancer*. Jun 7 2022;22(1):621. doi:10.1186/s12885-022-09731-w
34. Soukup T, Lamb BW, Morbi A, Shah NJ, Bali A, Asher V, Gandamihardja T, Giordano P, Darzi A, Sa Green J, Sevdalis N. A multicentre cross-sectional observational study of cancer multidisciplinary teams: Analysis of team decision making. *Cancer Med*. Oct 2020;9(19):7083-7099. doi:10.1002/cam4.3366
35. Jie Z, Zhiying Z, Li L. A meta-analysis of Watson for Oncology in clinical application. *Sci Rep*. Mar 11 2021;11(1):5792. doi:10.1038/s41598-021-84973-5
36. Suwanvecho S, Suwanrusme H, Jirakulaporn T, Issarachai S, Taechakraichana N, Lungchukiet P, Decha W, Boonpakdee W, Thanakarn N, Wongrattananon P, Preininger AM, Solomon M, Wang S, Hekmat R, Dankwa-Mullan I, Shortliffe E, Patel VL, Arriaga Y, Jackson GP, Kiatikajornthada N. Comparison of an oncology clinical decision-support system's recommendations with actual treatment decisions. *J Am Med Inform Assoc*. Mar 18 2021;28(4):832-838. doi:10.1093/jamia/ocaa334
37. Atieh Graham DM, McNamara DM, Waintraub SE, Goldberg SL, Norden AD, Hervey J, Pecora AL, Landstrom C, Snowdon JL, Francis PM, Jungbluth N, Wang CK, Latts L. Are treatment recommendations provided by cognitive computing supported by real

world data (Watson for Oncology with Cota RWE) concordant with expert opinions? *Annals of Oncology*. 2018;29:viii571. doi:10.1093/annonc/mdy297.031

38. DKG. Jahresbericht der zertifizierten Hautkrebszentren. 2023, Deutsche Krebsgesellschaft: Berlin. Accessed 01.08.2024, https://www.krebsgesellschaft.de/jahresberichte.html?file=files/dkg/deutsche-krebsgesellschaft/content/pdf/Zertifizierung/Jahresberichte%20mit%20DOI%20und%20ISBN/qualitaetsindikatoren_hautkrebs_2023-A1_230420.pdf&cid=111983

## 6. Vorabveröffentlichungen von Ergebnissen

Im Rahmen des Deutschen Hautkrebskongresses (ADO) am 14.09.2022 wurde ein E-Poster mit der Kennung eP075 präsentiert. Das Poster trug den Titel: “Development and validation of a decision support system for tumor boards – Experiences for malignant melanoma from a Certified Skin Cancer Center”