



# An overview and a roadmap for artificial intelligence in hematology and oncology

Wiebke Rösler<sup>1</sup> · Michael Altenbuchinger<sup>2</sup> · Bettina Baeßler<sup>3</sup> · Tim Beissbarth<sup>2</sup> · Gernot Beutel<sup>4</sup> · Robert Bock<sup>5</sup> · Nikolas von Bubnoff<sup>6</sup> · Jan-Niklas Eckardt<sup>7,8</sup> · Sebastian Foersch<sup>9</sup> · Chiara M. L. Loeffler<sup>7,8</sup> · Jan Moritz Middeke<sup>7,8</sup> · Martha-Lena Mueller<sup>10</sup> · Thomas Oellerich<sup>11</sup> · Benjamin Risse<sup>12</sup> · André Scherag<sup>13</sup> · Christoph Schliemann<sup>14</sup> · Markus Scholz<sup>15</sup> · Rainer Spang<sup>16</sup> · Christian Thielscher<sup>17</sup> · Ioannis Tsoukakis<sup>18</sup> · Jakob Nikolas Kather<sup>7,8,19</sup>

Received: 9 February 2023 / Accepted: 23 February 2023  
© The Author(s) 2023

## Abstract

**Background** Artificial intelligence (AI) is influencing our society on many levels and has broad implications for the future practice of hematology and oncology. However, for many medical professionals and researchers, it often remains unclear what AI can and cannot do, and what are promising areas for a sensible application of AI in hematology and oncology. Finally, the limits and perils of using AI in oncology are not obvious to many healthcare professionals.

**Methods** In this article, we provide an expert-based consensus statement by the joint Working Group on “Artificial Intelligence in Hematology and Oncology” by the German Society of Hematology and Oncology (DGHO), the German Association for Medical Informatics, Biometry and Epidemiology (GMDS), and the Special Interest Group Digital Health of the German Informatics Society (GI). We provide a conceptual framework for AI in hematology and oncology.

**Results** First, we propose a technological definition, which we deliberately set in a narrow frame to mainly include the technical developments of the last ten years. Second, we present a taxonomy of clinically relevant AI systems, structured according to the type of clinical data they are used to analyze. Third, we show an overview of potential applications, including clinical, research, and educational environments with a focus on hematology and oncology.

**Conclusion** Thus, this article provides a point of reference for hematologists and oncologists, and at the same time sets forth a framework for the further development and clinical deployment of AI in hematology and oncology in the future.

**Keywords** Artificial intelligence · Machine learning · Digital health · Large language models · Computer vision

## Introduction: The need for AI in hematology and oncology

Although Artificial intelligence (AI) is not a new research field (Schmidhuber 2022), recent developments driven by the availability of data sets and computing power, in particular, have led to the view that this rapidly evolving field may have the potential to transform our society (Rajpurkar et al. 2022). In the last ten years, AI has made significant advances in many different areas of society, including medicine. Hematology and oncology are a data-intensive and innovative medical specialty with a high clinical need for improved workflows and advanced methods for diagnosis

and treatment guidance. Due to aging populations, cancer will become more prevalent in the next decades. At the same time, our capabilities to diagnose and treat cancer have multiplied in the recent past, and will continue to do so in the future. This creates a massively growing amount of data and an increasing complexity of clinical workflows. The complexity is further increased by advances in all medical specialties involved in treating cancer patients, including hematology and oncology, radiology, pathology, surgery, human genetics, nuclear medicine, and others. Patients' heterogeneity in these regards require individualized solutions for which new scientific approaches need to be developed.

AI requires data to be available in a digital format. Digital data can be structured (such as data in a spreadsheet table or a database with pre-defined fields) or unstructured (such as unsegmented and/or non-annotated images or free text data). Many of these data types are routinely

✉ Jakob Nikolas Kather  
jakob\_nikolas.kather@tu-dresden.de

Extended author information available on the last page of the article

generated when diagnosing and treating cancer. Image data such as radiological or nuclear medicine imaging, as well as cytology or pathology images, are used to diagnose and stage tumors. Cancer subtyping is routinely performed using molecular and genetic testing, which can generate image data (for example, from fluorescence in situ hybridization or immunohistochemistry), genetic sequencing data (for example, genome, methylome, transcriptome data), metabolome data, proteome data, or other types. Aside from these data types, determining an optimal treatment strategy for a patient entails integrating a large number of highly variable pieces of information obtained from clinical examination, screening of previous medical records, and patient preferences. Moreover, general recommendations as derived, e.g., from randomized clinical trials are often not straight-forwardly transferable to every patient but require individual adaptations in the light of the individual patient characteristics.

In this setting, AI can be used for three purposes: (individualized) clinical care, research, and education. First, AI has the potential to be integrated into clinical routines and be used as a tool to aid humans in daily clinical practice. In this article, we mainly focus on this practical application of AI, since it can provide direct benefits to patients and physicians in hematology and oncology. For example, AI approaches and methods can be used to identify patterns in past cases that may help to predict how well a particular patient will respond to a specific treatment. Also, AI can assist to make treatment recommendations based on specific characteristics of each individual patient's tumor and monitor patients over time. In addition to this practical clinical use, there are two other aims. AI can be used as a research tool, allowing us to draw new scientific insights from clinical data such as new disease entities or pathomechanisms. In the future, it might be possible to better understand changes in the molecular-genetic and cellular composition of cancers, to derive new applications for existing drugs, to identify hidden patterns in oncological image data, to identify new therapy targets or pathologic processes, or to identify new biomarkers (Kleppe et al. 2021; Shmatko et al. 2022; Cifci et al. 2022). Finally, AI can be used as a tool for medical education, for example through the synthesis of data for educational purposes (Dolezal, et al. 2022; Krause et al. 2021; Chen et al. 2021). As populations age and cancer become more prevalent, more trained personnel are required to care for cancer patients. AI can potentially help to train these experts, although this aspect is still an emerging field in hematology and oncology.

To address, shape, and guide these advancements, the German Society of Hematology and Oncology (DGHO), the German Association for Medical Informatics, Biometry and Epidemiology (GMDS) and the Special Interest Group Digital Health of the German Informatics Society together established a joint working group "AI in Hematology and

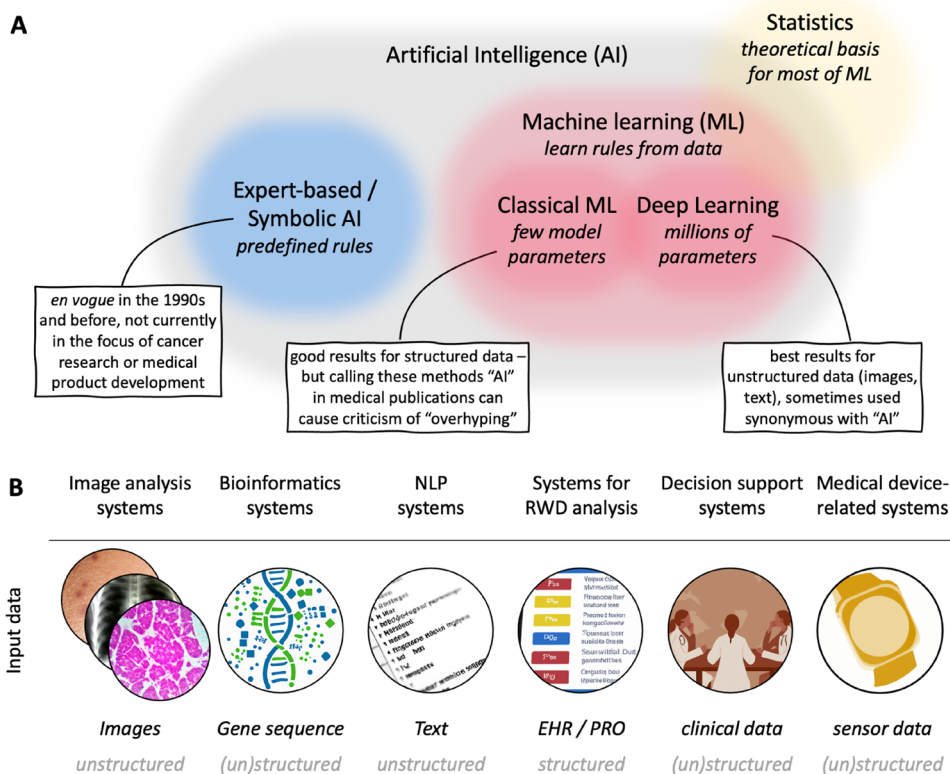
Oncology", entrusted with serving as a central hub for all AI-related activity in the field of hematology, oncology and cancer research. The aim of this article, which represents the result of a collaborative effort within this group, is to provide a consensus definition of AI applications in hematology and oncology and to map out the most promising sub-fields for the near future.

## Definitions and terminology: what is AI in biomedicine?

In the last decades, the field of AI has co-evolved with and drawn ideas from multiple adjacent research disciplines, and the terminology can be confusing (Fig. 1A). The terms machine learning (ML), deep learning (DL) and artificial intelligence (AI) have fuzzy boundaries which are heavily debated (Bzdok et al. 2018). In this article, we aim to provide a pragmatic definition of these methods, aiming to reflect the status quo in the biomedical research literature. Intuitively, a good example is as follows: The term "artificial intelligence" focuses on the word "intelligence", implying that a machine performs some kind of intelligent service. The term "machine learning" focuses on the word "learning", i.e., a machine learns something based on data with a certain aim, for example a disease model or a classifier. In a more formal, yet simplified view, two major sub-fields of AI have alternated and co-existed over the last decades: on the one hand, rule-based approaches, in which a human expert defines a fixed set of rules to classify data. This works well for very simple tasks based on structured data, but invariably fails for unstructured data and is also often not optimal for large quantities of structured data. On the other hand, ML, which does not encode any fixed rules, learns patterns from data. In the last five years, the ML paradigm within AI has made striking breakthroughs. In medicine, several ML-based algorithms have achieved regulatory approval in the last five years, while rule-based systems within AI are becoming less relevant to the practitioner (Shmatko et al. 2022; Benjamens et al. 2020). Hence, in this article, we will not discuss rule-based systems.

Within ML, there are three major classes of training strategies: reinforcement learning, supervised and unsupervised approaches (Shmatko et al. 2022). Reinforcement learning can help computer programs learn procedures, such as playing games, or navigating agents in a virtual world (Vinyals et al. 2019). Recently, reinforcement learning has been applied for medical applications, but it remains a niche approach in cancer research right now (Yala et al. 2022). Supervised techniques use labeled data to train a model, which then learns to predict the output based on the input. This is accomplished by feeding the model a series of input–output pairings known as training data, after which

**Fig. 1** Artificial intelligence in oncology. **A** A simplified visualization of key methodological areas in oncological data science with example methods. **B** An expert-based definition of key fields of application of AI in oncology. Abbreviation: *NLP* natural language processing, *RWD* real-world data, *HER* electronic health record, *PRO* patient-reported outcomes



the model learns to maps the inputs to the matching outputs. This method is often used in problems including classification, regression, and prediction. Supervised techniques may be utilized in medical applications for tasks such as diagnosis, prognosis, and therapy planning. Unsupervised learning methods, on the other hand, do not use labeled data. Instead, the model is given a set of inputs and is expected to discover patterns or structure in the data on its own. Clustering, dimensionality reduction, and anomaly detection are examples of such tasks. Unsupervised techniques in medical applications can be used to identify subgroups within a patient population, find hidden patterns in medical imaging data, and detect aberrant patterns in datasets in general. Concerning clinical routine, supervised methods are more common, while unsupervised methods are sometimes used in medical research. In this article, we focus on supervised methods.

Within supervised ML, many different methods exist, and one way to distinguish them is by their model complexity, i.e., the number of trainable parameters. Most traditional ML methods which were used up until 2012 had dozens, hundreds or thousands of parameters per model. This includes a range of methods such as k-means clustering, support vector machines (SVMs), shallow neural networks, decision trees and random forests. Empirically, these methods are useful for structured data, such as tabular data, and to some degree for unstructured data, for example to classify image features

which are extracted from radiology images according to predefined rules. Here, we refer to these methods as “Classical” ML. Since 2012, deep neural networks have emerged as a new tool with broad applications in medicine. Today’s deep neural networks build on mathematical theory, computer algorithms and hardware developed over many decades. Deep neural networks are different from “classical” ML methods in that they have millions or billions of parameters, and hence a much larger capacity to learn complex patterns. The use of deep neural networks is called “Deep Learning” (DL) and has become the dominant approach to process image, text, and other types of unstructured data in medicine. In the recent biomedical research literature, most “AI” studies are based on DL (Topol 2020, 2019). Medical applications of DL include diagnosis of disease, prediction of therapy outcomes, side effects or long-term prognosis.

## Application areas: How to use AI in hematology and oncology?

### Overview

AI methods can be used to address a broad range of clinical problems in hematology and oncology. In a consensus agreement of the AI working group, six classes of AI methods with established medical applications or potential clinical

**Table 1** An overview of AI systems in hematology and oncology with example applications

AI methods and applications	Source data (examples)	Clinical application in hematology and oncology (examples)
Image analysis systems	Radiology and nuclear medicine imaging data, histopathology image data, endoscopy, dermatoscopy, and others	aiding the diagnosis of tumors, assessing and predicting response to a given treatment, prognostication of the clinical course
Bioinformatics systems	Genetic sequencing data, and other -omics technologies	defining signatures of response to oncological treatments and targetable tumor subtypes
Natural language processing (NLP) systems	Spoken language or written notes (free text, unstructured data)	to automate the documentation or provide medical knowledge to doctors and patients, to extract data for further analysis from text, dialog systems (chatbots), analyze medical notes
Systems for real-world data (RWD) analysis	Electronic health records (EHR) and patient-reported outcomes (PRO)	Identification of adverse events, recommendation of treatment strategies
Decision support systems	Clinical and molecular data, data from multi-disciplinary tumor boards, clinical data, time series	Recommendation of treatment strategies for a given patient
Medical-device related systems	Measurements from physical sensors, such as wearables and smart watches	(out)patient monitoring and prediction of treatment related complications

relevance were summarized (Fig. 1B). These methods are applicable to a broad range of clinical problems and can process a range of different data formats (Table 1). However, it should also be noted that there are AI approaches that can fall into several categories, e.g., when image and -omics data are jointly analyzed. In the following sections, each of these AI systems will be explained and examples for practical use cases will be given.

### Image analysis systems

The analysis of digital images is one of the most common applications of AI in oncology (Kleppe et al. 2021; Shmatko et al. 2022; Farina et al. 2022; Shreve et al. 2022; Luchini et al. 2022; Echle et al. 2021). In fact, most clinically approved algorithms using AI in medicine are related to image data (Benjamens et al. 2020; Muehlematter et al. 2021; Alexander et al. 2020). Images are often prone to subjective human interpretation, and especially reading medical imaging data requires years of training. Given sufficient training data, computer models have been shown to be able to perform on narrow tasks at the level of human experts (Shmatko et al. 2022; Shen et al. 2019; Nagendran et al. 2020; Tschandl, et al. 2020). For example, AI has been used successfully to diagnose diseases such as diabetic retinopathy (Natarajan et al. 2019; Sosale 2020; Quelled et al. 2021), melanoma (Balasubramaniam 2021; Brinker et al. 2019), or lung cancer (Jacobs et al. 2021; Ibrahim et al. 2021) from image data. In addition, AI may be able to detect features that are not immediately apparent to the naked eye. For example, the prognosis of lung cancer patients can be predicted from routine computer tomography image data.

Traditionally, in the 2010s, “Radiomics” machine-learning methods have used sets of expert-defined visual “features”, coupled with simple ML models (Aerts et al. 2014). More recently, end-to-end Deep Learning has been increasingly applied to such tasks (Ghaffari Laleh et al. 2022). For example, AI has been used to prognosticate the course of colorectal cancer from digitized histopathology image data (Skrede et al. 2020), or to predict the response to immunotherapy from radiological imaging data (Trebeschi et al. 2019; Wu et al. 2019; Ligerio et al. 2021). Also, AI has been used to predict the presence of genetic alterations from image data (Shmatko et al. 2022; Kockwelp, et al. 2022; Kather et al. 2020), and is being discussed as a potential way to pre-screen patients for targeted molecular testing (Shmatko et al. 2022). Thus, AI-based image analysis systems can serve two purposes in oncology: they can speed up diagnostic processes, make them more consistent and readily available even in low-resource settings. On the other hand, AI-based image analysis systems can in some circumstances extract prognostic or predictive information from images, and thus serve as a biomarker for precision oncology. In both types of applications—automation and biomarkers—rigorous clinical evidence is required before these systems are used broadly in clinical routine (Geis et al. 2019).

### Bioinformatics systems

Omics technologies (e.g., genomics, proteomics, metabolomics) generate large amounts of data that can be difficult for humans to interpret (Eraslan et al. 2019; Elmarakeby et al. 2021; Lipkova et al. 2022). In the last decades, computer-based methods to analyze these data have co-evolved

with the laboratory assays (Shmatko et al. 2022). For example, the development of genome sequencing assays has been accompanied by the development of algorithms for sequence alignment and variant calling. Many bioinformatic machine-learning methods were developed for selecting molecular features and constructing molecular classifiers for disease entities or prognosis (Horn et al. 2018; Staiger et al. 2020). The development of all these methods has predated the era of Deep Learning. In fact, Deep Learning does not have a role in standard genetic diagnostics of cancer. However, additional useful information may be hiding in genome sequences and other -omics data. Several studies in the last few years also proposed Deep Learning methods to identify subtle patterns that were not identifiable by classical statistical approaches (Eraslan et al. 2019; Zeng et al. 2021; Tran et al. 2021). For example, AI has been used to predict outcomes and treatment response from sequencing data in cancer (Huang et al. 2020). Furthermore, -omics data might be combined with other data types (e. g., histopathology) to predict the clinical outcome of cancer patients more accurately (Chen et al. 2022). Researchers hope that by understanding genetic variants and their interplay in cellular networks in tumors they can find new therapeutic approaches, such as identifying potentially targetable neoantigens. Another field of application is the analysis of single-cell sequencing data often relying on neural network approaches such as variational auto-encoders. However, there is still a long way to go from academic research studies to embedding modern AI methods in clinical routine practice of genomically guided precision oncology.

### Natural language processing (NLP) systems

NLP is a branch of AI that deals with the interpretation and manipulation of human language (Yim et al. 2016; Sorin et al. 2020; Kung, et al. 2022; Yang et al. 2022; Singhal, et al. 2022). For example, NLP is used in chatbots and digital assistants such as Siri or Alexa, which are capable of understanding natural language commands and providing relevant information in response. In medicine, language is often used as an unstructured way to store and transmit information. In this context, NLP can be used to extract information from clinical reports or electronic health records (Yang et al. 2022; Thomas et al. 2014). It should be noted that current systems do not simply search and extract text but integrate the context (e.g., it is essential whether a diagnosis is present or excluded or how the temporal dependencies between reported events are). This information can then be used to support decision-making or generate predictions about disease progression or response to therapy. Although applications of NLP in oncology have been proposed more than five years ago (Yim et al. 2016), limitations of NLP methods have precluded widespread use. However, the

research field of NLP is evolving rapidly and applications which were unimaginable as little as one or two years ago are now reality. Most recently, at the end of 2022, new large language models (LLMs) GPT-3 and its variant chatGPT by the company OpenAI have raised broad interest. Modern LLMs can converse like humans, can respond to questions in medical examinations (Kung, et al. 2022) and generally can be used as a search tool, in particular to answer medical questions. In the next few years, we expect an exponential increase in the application of NLP in oncology. Human-level NLP systems are just emerging and hence, the process to translate this technology to clinical value in oncology is also just beginning.

### Real-world data (RWD) analysis systems

Electronic health records (EHR) are at the core of documenting any patient contact in oncology, and also integrate multimodal data related to the diagnosis of cancer and biomarkers for precision oncology (Parikh et al. 2022; Morin et al. 2021; Araki et al. 2022). Much EHR data are unstructured or just loosely structured, making it difficult to mine historically. Moreover, such data is often distributed across different primary IT systems (e.g., laboratory information system or a hospital information system), which in turn may not be designed to support interoperability (the ability of a system to function effectively with other systems). AI methods have been applied to EHR data and promise to make the data available in a structured way and extract hidden value from the EHR data (Morin et al. 2021; Araki et al. 2022). Poor design of EHR is an unpleasant experience for many doctors and contributes to physician burnout (Muhiyaddin et al. 2022; Tajirian et al. 2020; Kroth et al. 2019). Thus, AI-based support systems to parse EHR data could be useful for users. EHRs often contain time series data which are challenging to analyze, but have been analyzed with neural networks or dynamical models (Kheifetz and Scholz 2019; Tomašev et al. 2019). While EHR comprises mostly data generated by healthcare staff, the patient perspective can be underrepresented. This gap is filled by patient-reported outcome- and experience measures (PROMs, PREMs) including a data source of the patient's perspective which is increasingly being acknowledged in oncology as clinically relevant outcome measures in clinical trials and in certification processes evaluating health care in cancer centers (Parikh et al. 2022). EHR and PROM/PREM data are part of the loosely defined category of "real-world data" (RWD). We expect the AI-based analysis of RWD to be of much higher importance in the coming years, based on technological advances in multimodal AI models, NLP, and structured efforts to extract value from these data (Hegselmann, et al. 2022). The technology is now ready to be applied to many

use cases, but progress in the field will be limited by asking the right medical questions, identifying useful ways to apply this technology in the clinic and the data quality of the original documentation.

### Decision support systems

Several AI systems have been developed to automate part of the complex decision-making process in oncology (Kheifetz and Scholz 2019; Schmidt 2017; Rodríguez Ruiz et al. 2022). Some of them use multidisciplinary tumor boards as a blueprint. The defining feature of such boards is that specialists from different disciplines (e.g., surgery, medical oncology, radiation oncology, radiology, pathology) come together to discuss individual patient cases and make treatment recommendations. The clinical history of a given patient, all available results from diagnostic tests, patient preferences and current medical evidence are integrated in this process (Frank 2022). The level of complexity in clinical decision-making in multidisciplinary tumor boards is increasing with the expanding significance of genomic and molecular data for personalized treatment recommendations in cancer care (Büttner et al. 2019; Horak et al. 2021). As early as 2012, large-scale and well-funded programs have aimed to automate such recommendations, but so far, they have not reached clinical routine due to the complexity of extracting standardized recommendations from inconsistent data (Schmidt 2017). Despite these experiences in the past, the use of AI to automate decision-making in a way analogous to multidisciplinary tumor boards is still being commonly mentioned as a promising application. (Rodríguez Ruiz et al. 2022)

### Medical hardware-related systems

For users, the boundaries between medical hardware devices and consumer devices are blurring more and more and wearable sensors are becoming more common. For example, smart watches are widely used nowadays and can collect a wide variety of data, including information on heart rate, oxygenation, and movement. AI algorithms may be able to make sense of these high-dimensional data and provide insights into a patient's health (Sabry et al. 2022). For example, wearable sensors have been used successfully to detect early signs of disease such as sepsis (Ghiasi 2022). In hematology, this could be used to identify patients at risk for severe side effects and impending organ failure, including patients after myelosuppressive chemotherapy or stem cell transplantation (Nessle et al. 2022). Despite the obvious potential in this area, clinical evidence is still scarce and only a few dozen published studies have investigated an application of AI-based analysis of wearable sensor data in oncology (Sabry et al. 2022; Ghiasi 2022). Often, such systems

do not meet the criteria for regulatory approval as medical devices—especially if led by academic researchers. Again, this area will depend on hematologists and oncologists identifying the clinical need for new studies, running proof-of-concept studies and ultimately creating clinically relevant evidence how AI can be used for a patient benefit using properly designed clinical trials.

## Discussion and limitations

### Where are we headed?

In this article, we defined six main application areas in hematology and oncology where AI is already contributing or can contribute in the future. Some of these areas are already quite mature, for example, in AI-based image analysis, there are already dozens of approved medical devices that can be used in everyday clinical practice. Basic research or proof-of-concept work is still required in other areas, such as the support or partial automation of tumor board recommendations, modeling of individual time series data or the evaluation of sensor data with AI. Because of recent technological advances in AI, this technology is permeating our society more and more, and it is very likely that clinical management of patients in hematology and oncology will further benefit from AI approaches. As a result, it is critical that the transition to AI-assisted hematology and oncology is closely accompanied by a respective medical informatics infrastructure, new clinical trial concepts, and improved acceptance by doctors and patients.

### Data privacy and security

Whenever personal digital data are collected and linked together, the question of data privacy for the subject arises. It goes without saying that data in medicine must be securely stored, transferred, and protected from unauthorized access. This raises new issues in the age of artificial intelligence. Under certain conditions, it is possible to extract raw data from an AI network that has been trained on medical data and then published or sold for further use. In recent years, technological advancements such as differential privacy and secure multi-party computation have attempted to address this problem by introducing noise into the raw data of the training set or by privacy preserving computational approaches. However, in any case, it is critical that if possible, anonymized raw data are included from the start of the training process, and that, as is standard in medicine, an ethical approval and informed consent of the patient is available prior to the use and evaluation of patient-related data. (Seastedt et al. 2022)

## Biases, robustness and generalizability

Biases are another weakness of artificial intelligence and are particularly apparent in data-driven supervised machine-learning approaches (Andaur Navarro, et al. 2023). Machine-learning networks recapitulate and learn subtle patterns from training data, but these patterns are distorted in medicine and almost every other area of society due to prejudices and structural disadvantages for specific groups of people. If, for example, in the population represented in the training data, the implementation of complex molecular diagnostics is affected by place of residence, socio-economic factors, education, age, gender, or other factors, AI networks will learn these patterns on multiple levels and will eventually be able to apply them. Thus, use of an AI algorithm in the clinic may not generalize well enough to represent the diversity of future patients and therefore have an adverse effect on certain patient groups. The same is true if AI algorithms are trained and tested even on very large data sets from single institutions. Finally, high-parametric models are prone to overfitting, which reduces the performance in samples not used for model calibration. There is no perfect technical solution against such biases, but it is critical that both scientists and developers who set up the machine learning / AI system, as well as the end users, are made aware of them. Moreover, it should also be remembered that an appropriate study design is key to answer certain research questions (i.e., to evaluate the efficacy of an AI system, new concepts for randomized controlled trials will be needed). Further basic research in oncology is required to quantify the existence of such biases and to identify suitable areas of applications of these approaches and the potential harm to patients. It is precisely here that we see an important role for hematologists and oncologists, who, together with patients and patient advocacy groups, must ensure that such new technologies ultimately serve the well-being of patients and do not disadvantage specific patient groups based on their ethnicity, age, or other characteristics.

## Explainability and integration into clinical routines

Despite the possibilities of AI methods for hematology and oncology, these technologies often lack explainability since the underlying quantifications are based on highly complex calculations or learned network parameters which are not always directly relatable to biological mechanisms or structures. Despite the success of so-called explainable AI strategies (Minh et al. 2022), there are still many challenges, especially once these algorithms are used in critical situations such as clinical diagnoses and predictions (Ghassemi et al. 2021). Therefore and especially in the healthcare context, additional strategies have to be developed to enable informed clinical decisions. A possible approach could be to

inform hypotheses-free neural network models by biological knowledge or by relating learned network structures or classifiers to biological quantities or risk factors. Integration of AI into clinical reasoning has to be done carefully, meaning that the computed results need to be treated as additional evidence helping clinicians to gain a more complete picture. We see this as an important area of fundamental research for the next few years.

## Digital literacy and AI literacy

We expect that physicians will be increasingly confronted with artificial intelligence applications in the future. (Mosch et al. 2022) It is, therefore, imperative that the ability to assess the outputs of such artificial intelligence systems is part of medical education and training. Even today, simple computer skills pose a challenge to some doctors, such as typing quickly on keyboards or the intuitive use of graphical user interfaces. In the future, the complexity of our world will continue to increase massively due to artificial intelligence, and that makes it necessary for doctors to continuously learn the required skills during their studies, and later, in their careers. We see a particularly important role here for medical professional societies, which will set up and implement the appropriate further training curriculum for doctors.

## Outlook

We expect that AI will be broadly used to aid clinical decision-making and improve the quality of care in the near future. While these technologies are maturing fast, the sensible clinical use of AI in hematology and oncology is determined by defining appropriate areas of application and by establishing the required IT infrastructures. Moreover, as for every new treatment concepts, AI-based approaches need to demonstrate their superiority in well-designed clinical trials. AI methods could result in disruptive changes in the clinical practice. For example, access to data may change, and treatment decisions may become more and more influenced by IT systems rather than practitioners. Thus, the way medicine is performed requires monitoring and guidance. In Germany, large-scale medical informatics initiatives have been proposed as platforms to facilitate data-intensive research with clinical routine data. However, on top of this, new clinical trial strategies are also required to prove the advantage of such AI-guided individual therapy concepts. In all of these efforts, patients, physicians and a network of experts in methodology should guide and lead the AI transformation of hematology and oncology and that professional medical societies such as the German Society for Hematology and Oncology (DGHO), the German Association for Medical

Informatics, Biometry and Epidemiology (GMDS), and the German Informatics Society (GI), Special Interest Group Digital Health, together with other established initiatives and professional societies, will accompany and supervise this process.

**Author contributions** All authors jointly conceived and wrote the manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. The authors have not disclosed any funding.

## Declarations

**Conflict of interest** JNK reports consulting services for Owkin, France, Panakeia, UK, and DoMore Diagnostics, Norway and has received honoraria for lectures by MSD, Eisai, and Fresenius. WR reports travel support from Janssen, AstraZeneca and Amgen, honoraria for lectures from AstraZeneca and received a grant from Novartis. NvB received honoraria from Takeda and travel support from Janssen-Cilag. MS receives funding from Pfizer Inc. for a project not related to the topic of this paper. BR has no conflicts of interest. JMM reports consulting services for Janssen, Roche, Gilead, Abbvie, Jazz, Pfizer, Astellas, Novartis and funding of scientific projects from Janssen, Jazz, Novartis and Astellas. The other authors do not report any conflicts of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aerts HJWL et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 5:4006
- Alexander A, Jiang A, Ferreira C, Zurkiya D (2020) An intelligent future for medical imaging: a market outlook on artificial intelligence for medical imaging. *J Am Coll Radiol* 17:165–170
- Andaur NC et al (2023) Systematic review identifies the design and methodological conduct of studies on machine learning-based prediction models. *J Clin Epidemiol* 154:8–22
- Araki K et al (2022) Developing artificial intelligence models for extracting oncologic outcomes from Japanese electronic health records. *Adv Ther*. <https://doi.org/10.1007/s12325-022-02397-7>
- Balasubramaniam V, 2021 Artificial intelligence algorithm with SVM classification using dermoscopic images for melanoma diagnosis. *March*. 3: 34–42
- Benjamins S, Dhunoo P, Meskó B (2020) The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med* 3:118
- Brinker TJ et al (2019) Comparing artificial intelligence algorithms to 157 German dermatologists: the melanoma classification benchmark. *Eur J Cancer* 111:30–37
- Büttner R, Wolf J, Kron A (2019) Nationales Netzwerk genomische Medizin the national network genomic medicine (nNGM): Model for innovative diagnostics and therapy of lung cancer within a public healthcare system. *Pathologie* 40:276–280
- Bzdok D, Altman N, Krzywinski M (2018) Statistics versus machine learning. *Nat Methods* 15:233–234
- Chen RJ, Lu MY, Chen TY, Williamson DFK, Mahmood F (2021) Synthetic data in machine learning for medicine and healthcare. *Nat Biomed Eng* 5:493–497
- Chen RJ et al (2022) Pan-cancer integrative histology-genomic analysis via multimodal deep learning. *Cancer Cell* 40:865–878.e6
- Cifci D, Foersch S, Kather JN (2022) Artificial intelligence to identify genetic alterations in conventional histopathology. *J Pathol*. <https://doi.org/10.1002/path.5898>
- Dolezal JM et al (2022) Deep learning generates synthetic cancer histology for explainability and education. *Arxiv [eessIV]*. 22:432
- Echle A et al (2021) Deep learning in cancer pathology: a new generation of clinical biomarkers. *Br J Cancer* 124:686–696
- Elmarakeby HA et al (2021) Biologically informed deep neural network for prostate cancer discovery. *Nature* 598:348–352
- Eraslan G, Avsec Ž, Gagneur J, Theis FJ (2019) Deep learning: new computational modelling techniques for genomics. *Nat Rev Genet* 20:389–403
- Farina E, Nabhen JJ, Dacoregio MI, Batalini F, Moraes FY (2022) An overview of artificial intelligence in oncology. *Future Sci OA*. 8:787
- Frank B et al (2022) Multidisciplinary tumor board analysis: validation study of a central tool in tumor centers. *Ann Hematol*. <https://doi.org/10.1007/s00277-022-05051-y>
- Geis JR et al (2019) Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. *Radiology* 293:436–440
- Ghaffari Laleh N, Ligerio M, Perez-Lopez R, Kather JN (2022) Facts and hopes on the use of artificial intelligence for predictive immunotherapy biomarkers in cancer. *Clin Cancer Res* 29:1–8
- Ghassemi M, Oakden-Rayner L, Beam AL (2021) The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health* 3:e745–e750
- Ghiasi S et al (2022) Sepsis mortality prediction using wearable monitoring in low-middle income countries. *Sensors* 22:3866
- Hegselmann S et al (2022) TabLLM: few-shot classification of tabular data with large language models. *Arxiv [csCL]* 57:116
- Horak P et al (2021) Comprehensive genomic and transcriptomic analysis for guiding therapeutic decisions in patients with rare cancers. *Cancer Discov* 11:2780–2795
- Horn H et al (2018) Gene expression profiling reveals a close relationship between follicular lymphoma grade 3A and 3B, but distinct profiles of follicular lymphoma grade 1 and 2. *Haematologica* 103:1182–1190
- Huang Z et al (2020) Deep learning-based cancer survival prognosis from RNA-seq data: approaches and evaluations. *BMC Med Genomics* 13:41
- Ibrahim DM, Elshennawy NM, Sarhan AM (2021) Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. *Comput Biol Med* 132:104348
- Jacobs C et al (2021) Deep Learning for Lung Cancer Detection on Screening CT Scans: Results of a Large-Scale Public Competition and an Observer Study with 11 Radiologists. *Radiol Artif Intell* 3:e210027
- Kather JN et al (2020) Pan-cancer image-based detection of clinically actionable genetic alterations. *Nature Cancer* 1:789–799



- Kheifetz Y, Scholz M (2019) Modeling individual time courses of thrombopoiesis during multi-cyclic chemotherapy. *PLoS Comput Biol* 15:e1006775
- Kleppe A et al (2021) Designing deep learning studies in cancer diagnostics. *Nat Rev Cancer* 21:199–211
- Kockwelp J et al (2022) Cell selection-based data reduction pipeline for whole slide image analysis of acute myeloid leukemia. in: *Comp vis Pattern Recog Work*. 25:1825–1834
- Krause J et al (2021) Deep learning detects genetic alterations in cancer histology generated by adversarial networks. *J Pathol* 254:70–79
- Kroth PJ et al (2019) Association of electronic health record design and use factors with clinician stress and burnout. *JAMA Netw Open* 2:e199609
- Kung TH et al (2022) Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *Biorxiv*. <https://doi.org/10.1101/2022.12.19.22283643>
- Ligerio M et al (2021) A CT-based radiomics signature is associated with response to immune checkpoint inhibitors in advanced solid tumors. *Radiology* 299:109–119
- Lipkova J et al (2022) Artificial intelligence for multimodal data integration in oncology. *Cancer Cell* 40:1095–1110
- Luchini C, Pea A, Scarpa A (2022) Artificial intelligence in oncology: current applications and future perspectives. *Br J Cancer* 126:4–9
- Minh D, Wang HX, Li YF, Nguyen TN (2022) Explainable artificial intelligence: a comprehensive review. *Artif Intell Rev* 55:3503–3568
- Morin O et al (2021) An artificial intelligence framework integrating longitudinal electronic health records with real-world data enables continuous pan-cancer prognostication. *Nat Cancer* 2:709–722
- Mosch L et al (2022) The medical profession transformed by artificial intelligence: Qualitative study. *Digit Health* 8:20552076221143904
- Muehlematter UJ, Daniore P, Vokinger KN (2021) Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *Lancet Digital Health* 3:e195–e203
- Muhyaddin R et al (2022) Electronic health records and physician burnout: a scoping review. *Stud Health Technol Inform* 289:481–484
- Nagendran M et al (2020) Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 368:m689
- Natarajan S, Jain A, Krishnan R, Rogye A, Sivaprasad S (2019) Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone. *JAMA Ophthalmol* 137:1182–1188
- Nessle CN, Flora C, Sandford E, Choi SW, Tewari M (2022) High-frequency temperature monitoring at home using a wearable device: A case series of early fever detection and antibiotic administration for febrile neutropenia with bacteremia. *Pediatr Blood Cancer* 69:e29835
- Parikh RB et al (2022) Development of machine learning algorithms incorporating electronic health record data, patient-reported outcomes, or both to predict mortality for outpatients with cancer. *JCO Clin Cancer Inform* 6:e2200073
- Quelleg G et al (2021) ExplAIn: explanatory artificial intelligence for diabetic retinopathy diagnosis. *Med Image Anal* 72:102118
- Rajpurkar P, Chen E, Banerjee O, Topol EJ (2022) AI in health and medicine. *Nat Med* 28:31–38
- Rodríguez Ruiz N et al (2022) Data-driven support to decision-making in molecular tumour boards for lymphoma: A design science approach. *Front Oncol* 12:984021
- Sabry F, Eltaras T, Labda W, Alzoubi K, Malluhi Q (2022) Machine Learning for Healthcare Wearable Devices: The Big Picture. *J Healthc Eng* 2022:4653923
- Schmidhuber J (2022) Annotated history of modern AI and Deep learning. *Arxiv [csNE]*. 33:554
- Schmidt CMD (2017) Anderson breaks with ibm watson, raising questions about artificial intelligence in oncology. *J Natl Cancer Inst* 109:113
- Seastedt KP et al (2022) Global healthcare fairness: We should be sharing more, not less, data. *PLOS Digit Health* 1:e0000102
- Shen J et al (2019) Artificial intelligence versus clinicians in disease diagnosis: systematic review. *JMIR Med Inform* 7:e10010
- Shmatko A, Ghaffari LN, Gerstung M, Kather JN (2022) Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. *Nat Cancer*. 3:1026–1038
- Shreve JT, Khanani SA, Haddad TC (2022) Artificial Intelligence in Oncology: Current Capabilities, Future Opportunities, and Ethical Considerations. *Am Soc Clin Oncol Educ Book* 42:1–10
- Singhal K et al (2022) Large language models encode clinical knowledge. *Arxiv*. 5:103
- Skrede O-J et al (2020) Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* 395:350–360
- Sorin V, Barash Y, Konen E, Klang E (2020) Deep-learning natural language processing for oncological applications. *Lancet Oncol* 21:1553–1556
- Sosale B et al (2020) Simple, mobile-based artificial intelligence algorithm in the detection of diabetic retinopathy (SMART) study. *BMJ Open Diabetes Res Care* 8:892
- Staiger AM et al (2020) A novel lymphoma-associated macrophage interaction signature (LAMIS) provides robust risk prognostication in diffuse large B-cell lymphoma clinical trial cohorts of the DSHNHL. *Leukemia* 34:543–552
- Tajirian T et al (2020) The influence of electronic health record use on physician burnout: cross-sectional survey. *J Med Internet Res* 22:e19274
- Thomas AA et al (2014) Extracting data from electronic medical records: validation of a natural language processing program to assess prostate biopsy results. *World J Urol* 32:99–103
- Tomašev N et al (2019) A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 572:116–119
- Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 25:44–56
- Topol EJ (2020) Welcoming new guidelines for AI clinical research. *Nat Med* 26:1318–1320
- Tran KA et al (2021) Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med* 13:152
- Trebesch S et al (2019) Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann Oncol* 30:998–1004
- Tschandl P et al (2020) Human–computer collaboration for skin cancer recognition. *Nat Med* 26:1229–1234
- Vinyals O et al (2019) Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* 575:350–354
- Wu M et al (2019) Imaging-based biomarkers for predicting and evaluating cancer immunotherapy response. *Radiol Imaging Cancer* 1:e190031
- Yala A et al (2022) Optimizing risk-based breast cancer screening policies with reinforcement learning. *Nat Med* 28:136–143
- Yang X et al (2022) A large language model for electronic health records. *NPJ Digit Med* 5:194
- Yim W-W, Yetisgen M, Harris WP, Kwan SW (2016) Natural language processing in oncology: a review. *JAMA Oncol* 2:797–804
- Zeng Z et al (2021) Deep learning for cancer type classification and driver gene identification. *BMC Bioinformatics* 22:491

## Authors and Affiliations

Wiebke Rösler<sup>1</sup> · Michael Altenbuchinger<sup>2</sup> · Bettina Baeßler<sup>3</sup> · Tim Beissbarth<sup>2</sup> · Gernot Beutel<sup>4</sup> · Robert Bock<sup>5</sup> · Nikolas von Bubnoff<sup>6</sup> · Jan-Niklas Eckardt<sup>7,8</sup> · Sebastian Foersch<sup>9</sup> · Chiara M. L. Loeffler<sup>7,8</sup> · Jan Moritz Middeke<sup>7,8</sup> · Martha-Lena Mueller<sup>10</sup> · Thomas Oellerich<sup>11</sup> · Benjamin Risse<sup>12</sup> · André Scherag<sup>13</sup> · Christoph Schliemann<sup>14</sup> · Markus Scholz<sup>15</sup> · Rainer Spang<sup>16</sup> · Christian Thielscher<sup>17</sup> · Ioannis Tsoukakis<sup>18</sup> · Jakob Nikolas Kather<sup>7,8,19</sup>

<sup>1</sup> Department for Medical Oncology and Hematology, University Hospital Zurich, Zurich, Switzerland

<sup>2</sup> Department of Medical Bioinformatics, University Medical Center Göttingen, Göttingen, Germany

<sup>3</sup> Department of Diagnostic and Interventional Radiology, University Hospital Würzburg, Würzburg, Germany

<sup>4</sup> Department for Hematology, Hemostasis, Oncology and Stem Cell Transplantation, Hannover Medical School, Hannover, Germany

<sup>5</sup> IMMS Institute for Microelectronics and Mechatronics Systems GmbH (NPO), Ilmenau, Germany

<sup>6</sup> Department of Hematology and Oncology, Medical Center, University of Schleswig Holstein, Campus Lübeck, Lübeck, Germany

<sup>7</sup> Department of Medicine 1, University Hospital Carl Gustav Carus, Technical University Dresden, Dresden, Germany

<sup>8</sup> Else Kroener Fresenius Center for Digital Health (EFFZ), Technical University Dresden, Dresden, Germany

<sup>9</sup> Institute of Pathology, University Medical Center Mainz, Mainz, Germany

<sup>10</sup> MLL Munich Leukemia Laboratory, Munich, Germany

<sup>11</sup> Medizinische Klinik 2-Haematology/Oncology, University Hospital, Frankfurt am Main, Germany

<sup>12</sup> Computer Vision and Machine Learning Systems Group, Institute for Geoinformatics, University of Münster, Münster, Germany

<sup>13</sup> Institute of Medical Statistics, Computer and Data Sciences, Jena University Hospital - Friedrich Schiller University, Jena, Germany

<sup>14</sup> Department of Medicine A, University Hospital Münster, Münster, Germany

<sup>15</sup> Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany

<sup>16</sup> Department of Statistical Bioinformatics, University of Regensburg, Regensburg, Germany

<sup>17</sup> Competence Center for Medical Economics, FOM University, Essen, Germany

<sup>18</sup> Department of Hematology and Oncology, Sana Klinikum Offenbach, Offenbach, Germany

<sup>19</sup> Medical Oncology, National Center for Tumor Diseases (NCT), University Hospital Heidelberg, Heidelberg, Germany