

Statement

of the Central Ethics Committee at the German Medical Association

“Decision Support by Artificial Intelligence in Medical Practice”

In the event of any discrepancies between the English and German versions,
the German version shall prevail.

Preface

In the field of medicine, there are also high expectations associated with the application of artificial intelligence (AI) systems. Due to the versatile application possibilities and the constantly expanding technological developments, AI systems in medicine represent a very significant and at the same time extraordinarily dynamic field with the potential for further improving healthcare. However, the use of AI in medicine also arouses fears, for example when AI interferes with the relationship between doctors and patients, which is central to the medical profession.

In this context, the Central Ethics Committee at the German Medical Association (ZEKO) has decided to outline the current stage of development in a statement and to shed light on the questions associated with the use of AI for medical practice from a medical, ethical and legal perspective. For the purpose of narrowing down the topics, the ZEKO has placed the particularly relevant topic of AI-based systems for decision-making support in medical practice – so-called “Clinical Decision Support Systems” (CDSS) – in the focus of its considerations.

The possibilities and limits of AI in medicine are currently also being discussed internationally. Only a few days after ZEKO unanimously adopted the present statement on 23.06.2021, the World Health Organization (WHO) published its recommendation for action “Ethics & Governance of Artificial Intelligence for Health”. It describes the enormous potential that AI holds for medical care, but at the same time emphasises the challenges and risks from an ethical perspective, for

example with regard to competence and responsibility in AI-based decision support. In this sense, the recommendation for action also includes a proposal for six ethical principles for the use of AI in healthcare. The WHO’s recommendation for action underlines the topicality and importance of ZEKO’s statement presented here. With this statement, ZEKO would like to offer orientation in the dynamic field of AI-based decision support for physicians, but also for patients and the interested public – also in the sense of the information required by the WHO, especially for those working in the health sector. ZEKO hopes that the statement will contribute to raising awareness of the ethical challenges in the development and use of AI-based CDSS and thus to counteract undesirable developments at an early stage.

We would like to take this opportunity to thank all those involved for their constructive contributions and discussions as well as for their commitment in the preparation of this statement.

Berlin, July 2021

Prof. Dr. jur. Jochen Taupitz
Chairman of the Central Ethics Committee
at the German Medical Association

1. Introduction

In the wake of the digital transformation of the health care system in Germany, IT-based systems for decision support – so-called “Clinical Decision Support Systems” (CDSS) – have been developed in recent years. Many of these are still in the testing phase, but others are already in clinical use.

In the past, doctors were already able to draw on different interdisciplinary and interprofessional findings and observations (laboratory findings, assessments by colleagues, observations by staff, etc.), so that the integration of new technical components into the treatment process did not mean a fundamental change in the medical purpose, at least not in a structural sense. However, with the dynamic and broad field of CDSS, which includes different technologies, this support of medical activity is qualitatively expanded. Some of these CDSS work with artificial intelligence (AI) processes, i.e. software and hardware systems, “that, given a

complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data” [30] and using it to derive more or less highly automated recommendations for action or decision-making.¹

The purpose of CDSS is to support physicians and patients by providing a large amount of clinical-diagnostic information for the joint decision-making process, which is selected individually and case-oriented by integrated software systems. The use of CDSS can contribute to improving patient care through more precise diagnostics and personalised therapy selection [8]. The

¹ Due to the categorical difference between humans and AI systems, only personal actors can be considered as subjects of actions and decisions. Although AI systems can simulate certain individual intelligent tasks and support reasonable decision-making processes by doctors, they cannot make decisions or act themselves due to their lack of consciousness and language comprehension.

German Medical Association emphasises the need for CDSS so that “up-to-date medical knowledge is made available to the treating physician in a targeted manner at all times” [11]. Nevertheless, the framework conditions for the use of these new technologies are dynamic and their use is linked to ethical, legal and social challenges.

For this reason, ZEKO would like to offer orientation in the dynamic field of digital decision support for physicians with this statement. To accomplish this, it first informs about the current state of technical developments, describes the challenges in the use of digital decision support, evaluates these from an ethical point of view and describes the regulatory framework. With its recommendations, ZEKO would like to contribute to raising awareness of the special features of CDSS and to avoiding undesirable developments. The statement is directed both at physicians who already work with systems of (partially) automated decision support or are considering their use, and at physicians who have not yet had any contact with these technologies. The focus of this statement is on CDSS that complement the work of physicians with digital decision support and address one of the four areas of diagnostics, therapy, prognosis and prediction.

The statement does not deal with support systems for the work of health care professionals (e.g. nursing, physiotherapy) or systems that are used independently by patients (e.g. so-called “symptom checkers” or health apps). The statement also does not deal with IT-based systems that do not provide decision support but are intended to optimise and simplify health care in other ways (e.g. hospital information systems, “nursing robots”, electronic communication systems).

The question of whether and to what extent complex AI-based systems and their machine-learning functions can be attributed to an independent actor status in the medium and long term, as well as mental properties and a form of intelligence analogous to the practical reasoning of humans [23], cannot be dealt with in this statement. In view of the current state of development of technical support systems, this question points to a still distant future, which interpretation is controversially discussed with regard to so-called strong AI systems. Instead, this statement focuses on CDSS that are already in use or about to be introduced into clinical use.

2. Current Status

Possible and in part already realised clinical application examples/fields for CDSS exist in diagnostic, therapeutic, prognostic and predictive contexts.

CDSS in diagnostics have been developed for radiological imaging, for example, where conspicuous areas within images (e.g. suspicious areas in mammograms) are detected and marked. Here, the “path” from the measurement data to the image (reconstruction) is partly improved with machine learning methods. In machine learning, artificial systems are able to largely independently recognise patterns and regularities from training data and are also able to apply these findings to previously unknown data. CDSS are also already being used for clinical diagnostics in dermatology to assess the malignancy of skin lesions [75]. Ophthalmology (examination of the ocular fundus) can also be supported by CDSS in the future. Here, machine learning-based methods exist that enable the (partially) automated assessment of retinal structural anomalies and thus potentially support the monitoring of diseases of the ocular fundus. In addition, AMELIE (Automatic Mendelian Literature Evaluation) is an instrument that ranks

gene variants for genetic diagnoses and analyses the most probable causes of a phenotypic disease pattern on the basis of a literature search (PubMed). With the help of “Natural Language Processing” (NLP), full texts are searched and the result indicates on which studies the ranking is based [6; 26].

With regard to the use of **CDSS in the therapeutic field**, applications are available, for example, which support preoperative therapy planning and the intraoperative procedure. Thus, innovative systems using 3D visualisations or the improvement of endoscopic navigation (e.g. depth estimation, odometry) should increase the precision of interventions and reduce their invasiveness in computer-assisted surgery. However, there are also setbacks in the implementation of CDSS, for example in the non-surgical field: thus, the IBM development “Watson for Oncology” is under criticism because of uncertain and incorrect treatment recommendations [63]. This system was developed to provide individualised treatment recommendations for patients with cancer based on the processing of large amounts of scientific literature and case-based training by clinicians.

The use of CDSS, which provide information on the **clinical prognosis** of patients, is a particularly controversial area from an ethical point of view. By using artificial neural networks and incorporating both individual and population-based data, the aim is to predict risks for adverse events (e.g. cardiovascular events) or the survival time of patients suffering from cancer or end-stage renal disease, for example. This information could then be taken into account when deciding whether to initiate or reduce therapy. Particularly controversial in this context is the data on which a prognosis is based – especially in view of the sometimes considerable variation in individual clinical courses – and possible roles of health economic parameters.

The use of CDSS goes one step further than prognosis in the **prediction of diseases**, which starts with healthy people and aims to reveal dispositions and susceptibilities to certain diseases [74]. The focus here is the assessment of individual risk factors such as blood pressure, body mass index, lifestyle, biomarkers from genomics or metabolomics or the prediction of individual ways of reacting to drugs (pharmacogenetics). In the meantime, there are also solutions for companies that combine health monitoring of employees with occupational health services. However, prediction is often based on statistical probabilities that are difficult to interpret and are based less on methods of deductive causal determination than on inductive causal assumptions from correlation analyses, which can lead to erroneous conclusions [44; 60]. There is also often a gap between diagnostic possibilities and preventive therapeutic options. Many such approaches are also characterised by the fact that the raw data collected for risk assessment are not generated from an existing doctor-patient relationship, but as “direct-to-consumer” offers in the domestic context of prospective patients [55; 4]. The processing of different types of data poses challenges, which may include disclosure, data protection and quality assurance.

CDSS can already achieve results comparable to or even better than those of physicians in certain subtasks through the use of modern data processing methods. For example, a comparative study between an automatic classification system (based on “Deep Convolutional Neural Networks”) and dermatologists in the assessment of photographically documented skin lesions showed that the system was superior, especially with regard to the specificity of the results [27]. Similarly encouraging results

in human-machine comparisons have been reported in the assessment of conventional x-rays [15], arrhythmia detection in ECG [28] and paediatric diagnoses [37]. In the surgical field, it has been established that the therapeutic recommendations of the systems are highly consistent with the therapy chosen at highly qualified surgical centres [14]. Smaller studies also suggest that a physician's choice of therapy can be improved by using CDSS as compared to not using them or using other sources of information [51; 43]. However, caution should be exercised when evaluating the accuracy of diagnoses made by humans and deep learning algorithms in comparative studies. For example, according to a meta-analysis published in "The Lancet Digital Health", only a very small proportion of over 20,000 image diagnostic studies comparing diagnostic accuracy between humans and deep learning algorithms had externally validated results, and there were often flaws in study presentation [39]. This shows that not only are larger study populations needed, but also better quality studies, especially prospective ones, to assess the clinical validity of CDSS in practice [cf. 36].

It must also be taken into account that the design of such comparative studies does not indicate whether the therapy recommendations of CDSS and the doctors are therapeutically correct (because both can be wrong), nor does it indicate whether it is incorrect treatment if doctors do not follow the recommendation of the CDSS [67]. To investigate this question, a comparison between the outcomes of different treatment protocols would be methodologically relevant, but this is not usually done. Quality and safety problems as well as errors of CDSS can arise, for example, if the application conditions and training data differ too much ("distributional shift") and the CDSS do not take into account the significance of false-positive and false-negative diagnoses. If certain groups of patients are underrepresented in the training data, the specificity and sensitivity of the CDSS can be considerably lower with regard to this group of patients ([13], see Ch. 3.6).

Comparative studies also indicate that the results of human judgement and assessment are by no means inferior to the "outcome" of digital devices. In principle, it must be taken into account that humans make different mistakes than machines. Machine-learning systems lack the contextual knowledge of doctors, which can lead to considerable errors [12; 56: p. 52; 25]. The comprehensive knowledge of human experience is an important resource in health care that cannot be replaced by machines (see Ch. 3.5). Conflicts and burdens of acting in the presence of uncertainty cannot be completely resolved even through the use of CDSS. Complexity and uncertainty will persist in medical practice even with highly developed CDSS. However, well-evaluated CDSS can provide additional guidance and thus support human judgement.

2.1. Differentiation according to application options

A classification of the different CDSS can be made according to their possible applications. They differ in terms of

- the target group that will use the system (including doctors, health professionals, patients, citizens),
- the field of application (including diagnostics, therapy, prognosis, prediction, prevention, rehabilitation, routine/emergency),
- the specialised medical disciplines being addressed (including dermatology, radiology, oncology) and
- the concrete form of decision support in terms of how it is integrated into the clinical workflow.

2.2. Types of AI-based CDSS

CDSS can also be differentiated more precisely in terms of their technical implementation and mode of operation. All automation processes, even the more complex ones, are based on the use of algorithms. For one thing, there are logical instructions for action (if-then statements) programmed in software that describe to a system exactly what to do in a certain constellation. Then there are "learning" systems that not only execute predetermined instructions for action, but also "learn" new courses of action from the training data entered and their processing. An increasing part of CDSS is based on systems of this AI. Techniques such as machine learning, neuronal networks or deep learning refer to procedures by which an AI system "learns" to solve problems that are not precisely specified or whose solution paths cannot be precisely described by symbolic rules of reasoning. An example of such procedures is the recognition of images, language and texts, which provides automated reasoning from unstructured data to behavioural or outcome predictions. In many of these AI systems, the "how" and "why" of a solution path found by AI and an automated decision recommendation can hardly or not at all be comprehended from the outside ("black box", opacity, see [7; 50]). Despite various attempts (still in their initial stages) to contain the problem of opacity through new models of "explainable AI" [3; 58], fundamental questions arise from the functionality of such systems that go beyond the requirements for the methodological expertise of physicians as individual users and primarily concern the concrete responsibilities of developers of such systems. The fact that the functionality and the achievement of a result is in principle no longer comprehensible even for the developers, is a new phenomenon in technical systems.² Even if the comprehensibility of the algorithmic calculations can be increased through corresponding technical investments, the question arises whether this makes sense or is even necessary for individual users, since doctors also use technical aids in many other respects, the exact functioning of which they themselves are only slightly familiar with. In this context, further developments should be carefully observed.

Finally, it should be emphasised that the use of machine learning techniques is always associated with a certain percentage of errors, whereby accuracy is essential for the quality of an AI system. The type and amount of training data processed is significant here, whereby additional relevant considerations are aspects such as the correctness (e.g. incoming parameters and disease diagnoses) and representativeness of the data (e.g. gender, ethnicity, social factors), the algorithms used and the targets. Machine learning relies on a very large amount of training data. If this contains concealed biases or overemphasises certain correlations, it can lead to distortions ("bias") in the results. The design of AI systems also raises the question of the extent to which certain values and interests are prioritised or weighted over others, and which objectives (e.g. definition of quality of life) exclude other objectives (e.g. different interpretations of fairness).

² A distinction must also be made between contingent and absolute opacity. The first form is not specific to CDSS, but characterises the use of many technical systems whose exact mode of operation is de facto not understood by users. The latter, on the other hand, is a new phenomenon that results, among other things, from the limited computing capacity of the human brain: Since the machine can take correlations between a larger number of parameters into account due to its increased computing power, the question arises as to whether this problem is actually only an investment problem or cannot be solved in principle, so that a residual opacity appears insurmountable in principle.

2.3. Innovation potential of AI-based CDSS

The previous sections have already made it clear that the use of AI-based CDSS offers many opportunities to improve patient care through more effective data processing. AI systems can be used not only to analyse and evaluate large data sets, but also to discover statistical correlations that can indicate possible causal relationships, which, if validated, can be useful in diagnosis and treatment. AI can be used to enable the processing of very large quantities of relevant images and texts and to compare them with new findings in real time. As the technology is refined – for example, higher image resolution – the systems again require new training data. Some CDSS can be equipped with highly stratified risk profiles, so that future diagnostic processes and therapy recommendations could be more precise. If CDSS prove their reliability in use, it can be assumed that they will generate a boost in innovation for broader care across the board and possibly contribute to more equitable patient care through AI-based diagnostics in medicine. For example, in its current report “Digitalisation for Health”, the German Council of Experts on the Assessment of Developments in the Healthcare System (Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen) particularly emphasised the opportunities of CDSS for people with rare diseases and constellations of findings, since they can fill “existing gaps in diagnostics and therapy” [64].

3. Ethical evaluation of CDSS

It can be assumed that AI-based and specifically machine-learning CDSS will gain increasing importance in medicine in the future [1]. The use of AI-based CDSS offers a wide range of possibilities for improving patient care (see Ch. 2). However, it is also associated with various challenges, such as the clear allocation of responsibilities, the transparency of data generation, the protection of autonomy, trust and privacy, along with the quality of communication between doctors and patients [cf. 47; 49]. In this context, an early and appropriate weighing of opportunities and risks is an important ethical task with regard to future healthcare. This can only succeed if the moral challenges resulting from the new human-machine relationship are taken seriously.

The European Commission (EU Commission), for example, included this in its White Paper “On Artificial Intelligence – A European Approach to Excellence and Trust” by emphasising there that the use of AI in Europe must be governed by “our values and fundamental rights, such as human dignity and the protection of privacy” ([20]: 2). AI systems should be human-centred and trustworthy. Their impact should be assessed not only for the individual, but also for society as a whole, and risks should be avoided. In the European Union (EU), a consensus has been established by the EU Commission’s High Level Expert Group (HLEG) on seven core requirements: the ethical criteria of the principal of primacy for humans, technical robustness and safety, preservation of data sovereignty and transparency, consideration of diversity, non-discrimination and fairness, societal well-being and accountability are to be respected and implemented [31]. As the EU Commission emphasises, these core requirements must be met accordingly at the various levels of responsibility in order to ensure trust towards staff and systems.³

³ These core requirements are also included in the EU Commission’s draft regulation COM(2021) 206 on the management of artificial intelligence, published on 21.04.2021 [19].

3.1. Trust and trustworthiness

Doctors and patients must be able to trust that the (correct) use of CDSS systems will enable at least the same, ideally better, care. Patients must also be able to (still) trust that they are treated as individuals, that their well-being and self-determined will are at the centre of care, and that they do not get the impression that they are being reduced to their data. In the same respect, they are dependent on receiving all information relevant to them for consent to certain treatment proposals that is not only legally effective, but actually autonomous. Where this is not the case, there is a risk of erosion of trust in the system, which is indispensable for the provision of medical services. Preserving this systemic trust is also a task of the medical profession where doctors act as advocates for their patients in the event of undesirable developments through the use of CDSS.

From the perspective of doctors, they must be able to trust that the standards developed for the development and use of such systems meet high quality benchmarks. This requires trustworthiness on the part of the CDSS producers, the hospitals that use these systems, as well as the professional societies and other creators of guidelines, who should develop clear application criteria with corresponding qualitative standards. Conversely, by using CDSS responsibly, doctors contribute to maintaining society’s trust in the medical profession.

3.2. Responsibility

With regard to the use of CDSS in the health sector, it is necessary to distinguish between different groups of actors and levels of responsibility.

At the **micro level**, doctors must ensure that they have the appropriate digital skills for the use of CDSS. In particular, as individual users of these systems – as is also the case in other areas of the use of technical devices – they are personally responsible for acquiring the necessary and adequate knowledge of professional handling, e.g. including the limitations of the new technologies, and for using the systems with due diligence. Doctors must be aware of the following problems in particular:

CDSS increasingly provide recommendations through machine learning and especially through deep neural networks without being able to provide a tangible or verifiable explanation of how this result was achieved (black box, opacity, see Ch. 2.2.). Lack of transparency of AI-based CDSS is, in addition to the reasons already mentioned such as complexity, insufficient traceability, unpredictability and partially automated behaviour, also due to the fact that AI systems are protected by trade/commercial secrets or by intellectual property rights. It should also be noted that the medical profession is bound by the principles of medical ethics, while the development of AI-based CDSS often takes place in the private sector under high pressure and is driven by developers who are not bound by medical professional ethics [46].

Data-driven CDSS also form classes and categories of patients and disease types generated according to different ranking patterns, evaluation criteria and risk scores. They can identify new patterns through correlations. However, such correlations should not be misunderstood as causality, but require clinical validation [cf. 17]. It follows in particular, that physicians must in any case carry out plausibility checks of the automated CDSS decision recommendations. This applies all the more, the less comprehensible the genesis of a decision recommendation is. If CDSS are sufficiently validated and have proven themselves in practice, it

may also be required that their use is in accordance with the current medical standard and individual consideration. This is because medical responsibility is always committed to the individual patient, is subject to strict standards of care and the requirement to adapt treatment to the current state of knowledge of medical science and practice.

Doctors are always responsible for the complete diagnostic and therapeutic decision-making process. However, the demand for an “unrestricted preservation” of the “ultimate responsibility of the doctor” [8] in the use of CDSS must not lead to a situation where the overall burden of introducing and using these systems lies with the individual doctors. Instead, those responsible at the meso and macro levels play a central role in the approval of CDSS as medical devices and in the use of these systems, in order to ensure trust in the system, on which physicians and patients must be able to rely equally. This includes the establishment of technically appropriate feedback processes between the various parties involved.

At the **meso level**, the facility operator must carefully examine which of the CDSS are to be used in their own medical facility. The institution must also ensure that staff are adequately prepared for their use through appropriate training measures.⁴ The systems themselves must be adequately maintained on an ongoing basis and protected against misuse (e.g. hacking, etc.).

For the production of CDSS, the meso level must provide appropriate testing, certification and auditing measures, to ensure that the use of data is lawful, fair, secure, transparent and accountable. Technical, ethical and regulatory auditing procedures are needed to determine what kind of algorithms, models and (partially) automated decision-making systems should be developed and used in CDSS and on which data they shall be based and implemented, how their efficiency and benefits shall be verified as well as how their performance shall be validated and how to integrate them in healthcare. Further intensive research on the explainability and comprehensibility as well as non-discriminatory nature of AI-based forms (machine learning and deep learning) of CDSS is necessary to avoid undesirable developments [72].

At the **macro level**, the legislator is responsible for continuously reviewing the regulatory framework for so-called “intelligent medical devices” to ensure that it accommodates the dynamic technical developments in this area and that the appropriate framework conditions are created to ensure that the actors working in the health sector at the meso and micro levels can actually fulfil their respective responsibilities. At the same time, medical societies and other authors of guidelines are asked to address the use of CDSS with machine-learning functions that are already part of the medical standard (see Ch. 4.3.2.) in their areas and, for example, to take them into account in their guidelines. With regard to education, further education and training, the increased teaching of digital competences is a cross-sectoral topic of central importance in order to provide the actors working in the health sector with appropriate qualifications concerning the new tasks and challenges associated with digitalisation. In terms of CDSS, they should above all be enabled to evaluate the benefits and risks and to use the systems in a patient-oriented manner.

The scope of responsibility thus assigned extends to the following areas of focus.

3.3. Autonomy

The core value of autonomy for the perception of ourselves as human beings is based on the ability to shape our own actions in a freely responsible manner by means of independent insight into reasonable behavioural principles. Although autonomous human actions cannot be imagined according to the pattern of isolated unconditioned self-determination, it is crucial for such a conditioned autonomy that actors can critically reflect on the given conditions and can base their decisions accordingly. Challenges arise from the use of CDSS that affect both the autonomy of physicians and that of patients.

3.3.1. Medical autonomy

In the context to be discussed here, medical autonomy means above all the freedom to choose the examination or treatment method suggested to the patient, although this choice is under obligation to the patient (see Ch. 4.4.1.). This can be influenced in various ways by the new technical systems.

The concrete autonomy-related risks associated with human-machine interaction in the clinical use of CDSS include in particular

- the acceptance of diagnosis proposals by doctors without further examination of their own (“automation bias”, “automation-induced complacency” [57]);
- decreased or inadequate responses to warning signals from the system (“alert fatigue” [53]);
- self-fulfilling prophesy: if a system trained on outcome data, e.g. from cancer patients, predicts a poor prognosis, palliative rather than curative treatment is given, which reinforces the recommendation of the CDSS [13];
- the danger of over-diagnosis and over-treatment, e.g. when AI detects possible signs of a disease whose manifestation is statistically very unlikely and/or very implausible, but physicians initiate a corresponding invasive diagnosis or therapy as a precaution, e.g. also for their own protection [2].

Overall, crossing the line between decision-making assistance and adoption of decisions carries the risk of negligence or even loss of control [25; 71; 69]. Physicians must continue to be able to have authority across the overall process of diagnostics, therapy, prognostics and prediction, which increasingly involves suggestions from machine-learning systems. This task cannot be delegated.

3.3.2. Patient autonomy

The impact of the use of AI-based CDSS on patient autonomy is already the subject of intense debate. Central ethical questions concern, for example, the information and consent for the use of CDSS or also possible demands by patients to be treated without the use of AI.

This means that patients may have to be informed and give their consent as to whether and to what extent such systems should be integrated into the treatment process and what opportunities and risks are associated with their use ([cf. 59; 45], see also Ch. 4.4.). A high degree of communication and transparency is necessary to protect a relationship of trust between doctors and patients. However, informed consent must not be misunderstood as a delegation of responsibility to the patient [18].

It is therefore important that the use of CDSS is appropriately integrated into the shared decision-making process between doctors and patients. The CDSS assessments should only be seen

⁴ In 2019, the German Medical Association developed a curriculum for continuing medical education titled “Digital Health Applications in Practices and Clinics”, which teaches skills such as assessing the safety of digital health applications.

as recommendations that doctors introduce into the joint decision-making process after their own critical review. On this basis, if necessary also when patients themselves bring the results of CDSS into the decision-making process, a benefit can also arise for patient autonomy. However, if the automated recommendations are used in a way that bypass the informed decision of patients, there is a risk that the self-determination of patients in digital healthcare is no longer reliably guaranteed. There is a risk of “computer paternalism” that does not meet current standards of patient autonomy [41]. Against this background, there is much to be said for a “value-sensitive” design of AI in the clinical decision-making context: by including the interests and preferences of individual patients, opportunities arise for even more individualised treatment.

3.4. Communication and empathy

Even if it cannot be ruled out that existing communication deficits in the relationship between doctors and patients could be exacerbated by the use of AI and automated systems, since diagnostic and therapeutic processes will become more automated and functionalised, hopefully the increase in efficiency through the transfer of tasks to CDSS will create more space for communication between doctors and patients. Considering this, CDSS should contribute to doctors being able to devote themselves more strongly to pivotal tasks.

Making diagnoses, therapies, prognoses and predictions not only involves technical actions, but is also connected with emotions and values. These are the basis of the relationship of trust between doctors and patients. That is why, precisely within this relationship, communication is essentially determined by care and empathy as well as insight into the individual context of the patient [62; 24].

For this reason, in the field of application of AI-based assistance systems in medicine, it must be examined whether the interpersonal and emotional aspects of the relationship between doctors and patients are inappropriately pushed into the background, for example if human communication is supported or even replaced by technical voice assistance systems (such as “chatbots”). Areas of action that are characterised by empathic interaction between humans must not be devalued, and in addition to the factual information level, the important conversational messages between doctors and patients (listening, self-revelation, appeal function [cf. 24; 48]) must not be neglected.

3.5. Experiential medical knowledge

Medical healing is essentially based on the power of judgement of doctors who, on the basis of generalised knowledge about efficient treatment techniques, offer their patients an individually relevant and situationally appropriate diagnosis and therapy [73]. In contrast to a technically substitutable processing of individual health-relevant data, the most comprehensive perception possible of the concrete unique life circumstances of patients is of particular significance. In addition to the scientific competences that physicians acquire and continuously develop not only in medical school but also in continuing medical education and training [9], physicians acquire personal knowledge that does not represent a neutral interconnection of information, but in how the development and communication of this knowledge is incorporated into the physician’s experience.

Medical expertise is not superfluous, especially in diagnostic procedures, but demonstrates the special power of medical judgement. Even in services where the capabilities of AI exceed those of doctors, the accompanying examination and review by doctors is not dispensable and replaceable. This applies, for example, to diagnosis in cancer detection. In this field, it has been shown that a tandem of doctors and AI has the potential to improve the reliability of the diagnosis and thus constitute an optimal treatment strategy. In contrast to digital tools, it is possible for doctors to not only focus on looking for a tumour – as is the case with AI – but to simultaneously have the wider environment of the tissue in view [42; 40]. This makes it easier to see and assess irregularities that can only be identified from the overall picture or, if necessary, from the patient’s medical history. Therefore, AI still does not offer a substitute, but rather a support for doctors, for example in making a diagnosis.

One issue that is problematised as an unintended side effect of CDSS is the devaluation of physicians’ experiential knowledge (“tacit knowledge”, “embodied knowledge”), which is implicit and therefore cannot be included in the training data for machine learning. Other contexts of the individual case that cannot be easily mapped into data can also distort results. In addition, the routine use of CDSS can lead to doctors no longer acquiring this experiential knowledge at all, which means that real monitoring and, if necessary, correction of the machine is no longer possible, or that they are unable to offer remedial action in the event of system failure [12; 21].

At the same time, it should not be ignored that doctors also acquire a new type of experiential knowledge through the qualified and reflected use of CDSS, which relates to the handling of these systems to improve the treatment of individual patients.

Presumably, only productive human-machine collaboration will serve patients. Clearly identifying areas in medicine where AI research should focus and integrating AI algorithms with data in other areas will be critical if the benefits of this collaboration are to be realised. What impact this development will have on education, practice and the healthcare system needs to be critically monitored.

3.6. Risk of discrimination

AI-based CDSS can only be as good as the training data on which their calculations and pattern recognition are based. AI systems can consequently produce systematically biased or incorrect results that disadvantage certain groups of people without any factual justification. Such bias can originate from erroneous or incomplete data, misclassification and measurement errors, but it can also result from faulty AI [25]. However, bias is possible even with correct data and an error-free AI system, namely if the data basis used is itself biased. This happens, for example, when certain groups of people (e.g. women, children,⁵ ethnic minorities or people who have less access to the health system for socio-economic reasons, people with disabilities, people with rare diseases, trans and intersex people) are underrepresented or

⁵ Since children are still involved in the growth and developmental phases, additional data needs to be taken into account when making decisions compared to adults. In order for children to benefit from the development of CDSS to the same extent as adults, databases and algorithms would have to be expanded accordingly. However, this in turn is associated with ethical and legal challenges (vicarious consent, constant monitoring if necessary) as well as technical challenges (difficulties of standardisation).

inadequately represented in the data used [5].⁶ This also affects the profiling of certain segments of patients who are classified as less healthy, with higher health risks or less “compliant”. Then there is a danger that structural discrimination and stigmatisation in medical treatment⁷ will continue or even be reinforced by CDSS [56; 54; 52]. In particular, those responsible at the meso level (see Ch. 3.2.), but also doctors involved in the specific treatment context, should remain aware of this danger and keep a critical eye on the outcome of the systems they use from this perspective.

3.7. Data sovereignty and privacy

The medical use of CDSS with machine-learning functions also touches on issues of data protection and medical confidentiality, which in turn are closely linked to the necessary protection of privacy and intimacy (cf. [17]; see Ch. 4.6.). It concerns guaranteeing individual data sovereignty as responsible informational freedom. This concerns both the disclosure of sensitive health data for the provision of training data and the implications of the application of CDSS, especially for vulnerable groups. In order to use the opportunities of CDSS and to effectively counter the risks of new asymmetric power relations and the resulting losses of informational self-determination ([17]: 262), suitable and, if necessary, new protective mechanisms and design strategies must be developed that extend far beyond the health sector [65; 61].

4. Regulatory aspects

CDSS have the potential to improve and expand the examination and treatment options of physicians. However, CDSS also raise new types of legal concerns, some of which have not yet been satisfactorily clarified. What is lacking here is a specific regulatory framework for AI-based medical decision support systems. An overview of the legal issues associated with CDSS is provided in the following.

4.1. CDSS as a medical device

CDSS software can be an independent medical device or, as software embedded in medical devices or as an accessory to a medical device, be subject to the regulations for medical devices [50], namely the EU Medical Devices Regulation (MDR) and the Medical Devices Directive Implementation Act (MPDG). The decisive factor for classification as a medical device is that software is used for a specific medical purpose in the context of a concrete, individual treatment of a particular patient [16]. Whether the software acts directly in or on the human body is irrelevant (ECJ, judgment of 7 December 2017 – C-329/16 = European Business Law Journal/EuZW 2018, 166).

Systems that are subject to the regulations for medical devices must be clinically evaluated in a certification procedure before they can be placed on the market in Germany. If the clinical data

required for this do not exist, a clinical trial must be conducted, which in turn requires approval by the Federal Institute for Drugs and Medical Devices (BfArM) as well as a consenting assessment by the responsible ethics committee. Certification (CE marking) is intended in particular to exclude errors that endanger the health and safety of patients beyond an acceptable level. Details are regulated by the MDR and the MPDG.

If CDSS that are (certified) medical devices have defects, doctors may be required to report them. If, in the course of their professional activities, doctors become aware of malfunctions or deficiencies in the CDSS provided by the manufacturers that could significantly endanger the lives or health of patients, they must report this to the BfArM (§ 3 Medical Devices User Reporting and Information Ordinance [MPAMIV], § 6 Model Professional Ordinance for Physicians/MBO-Ä).

4.2. CDSS as method of examination or treatment

4.2.1. Type and scope of practice of physicians (see Ch. 4.3.) as well as the necessity of patient information and consent (see Ch. 4.4.) depend on whether CDSS represent an independent examination or treatment method or are to be regarded merely as an ancillary or accompanying service supporting the treatment. This can only be decided for the corresponding CDSS. If CDSS is an independent examination or treatment method, special duties of care apply to physicians, the type and scope of which depend on whether the method is already part of the medical standard or is to be regarded as new (see Ch. 4.3.2. and Ch. 4.4.). In contrast, for CDSS that are merely a supporting ancillary or accompanying service, only the general duties of care of physicians should apply (see Ch. 4.3.1.) and no patient information and consent should be required prior to their use, although this has not been clarified in a legally binding manner.

4.2.2. The extent to which CDSS are to be classified as an independent examination or treatment method or as a mere ancillary and accompanying service, and which criteria apply, has not yet been established. It is always necessary to consider the individual case, for which the following applies: the more the CDSS is used directly on patients (e.g. Da Vinci surgical robot; AI used to detect eye diseases in the human eye; interactive implant system for tinnitus) and the more it is used in the background to support medical decisions without direct patient contact (e.g. AI system for the analysis of tissue sections for the purpose of tumour detection; AI-based image recognition programme used to detect strokes on CT or MRI images), but effectively replaces medical decision-making, the more the principles for examination and treatment methods are likely to apply. In any case, if the medical decision is delegated to an AI-based automated system that acts independently on the basis of individual patient data and significantly influences the further course of treatment, it will need to be classified as an independent examination or treatment method.

4.2.3. A decision must also be made on the basis of the specific individual case as to whether CDSS, which are to be regarded as an independent examination or treatment method, belong to the medical standard or are new. The minimum requirement expected for classification as a medical standard for CDSS that are medical devices, is certification [16]. As an indication of the existence of a new examination or treatment method, the recommendations of the Federal Joint Committee (G-BA) according to § 135 in conjunction with § 92 (1) sent. 2 (5) of the German

⁶ It is well known that many economically weak countries (so-called low income countries) lack any infrastructure for data collection and processing, which considerably limits the potential use of such systems for these people. Therefore, the development and establishment of the CDSS should take into account the issues of access and distributional justice at the global level.

⁷ Examples of such structural discrimination in the health care system can be that certain patients or patient groups receive shorter or worse treatment than others or do not receive the treatment they need from doctors or in hospital, for example because of language problems or lack of barrier-free access to healthcare facilities. Another example would be that a health insurance company either does not cover high treatment costs for incomprehensible reasons or has classified private patients higher when concluding a contract.

Social Code/SGB, Book V can be applied both positively as well as negatively.

4.3. Duty of care of physicians

The application of CDSS by doctors in the context of medical treatment (§ 630a Civil Code/BGB) requires care, the nature and extent of which depend on the type and method of application, comprehensibility and safety of the CDSS in question and its stage of development.

4.3.1. Physicians must always sufficiently familiarise themselves with the functionality of a (partially) automated system before using it. This applies regardless of whether the CDSS in question is to be classified as an independent examination or treatment method or as a supporting ancillary or accompanying service (see Ch. 4.2.). In any case, physicians must ensure that they have the – sometimes considerable – technical knowledge and skills necessary for the use of the CDSS. If doctors lack the technical expertise required for the CDSS in question, they must consult technically competent experts. Clinicians must ensure that they and their staff receive the training and education necessary for the continued use, maintenance and care of the CDSS. Adequate precautions must also be taken against misuse of the systems (e.g. hacking) by unauthorised third parties.

With regard to the operation of the system, doctors need not understand all the technical details; in particular, they do not have to understand the algorithms on which the system is based. However, doctors must familiarise themselves with the functionality of the CDSS to the greatest extent possible and reasonable. This includes acquiring knowledge about the proper use, monitoring and maintenance of the CDSS. The system must be operated and maintained in accordance with the manufacturer's recommendations.

Physicians should be aware that CDSS can contain particular sources of error, such as faulty algorithms, electronic viruses or an inadequate training database (see Ch. 2, cf. also [29]). These processes are often non-transparent and not sufficiently traceable. This entails the risk of incorrect diagnostic and therapeutic recommendations and discrimination against patients. In view of their ultimate responsibility for diagnosis and treatment decisions, physicians should exercise particular care when dealing with CDSS. Specifically, they must at least carry out plausibility checks of the automated decision recommendations of CDSS [66]. The less transparent and comprehensible the CDSS is for physicians, the more carefully the automated decision recommendation should be monitored for plausibility and accuracy.

4.3.2. Special duties of care apply to those CDSS that are independent methods of examination or treatment. If the CDSS is already part of the medical standard, in individual cases doctors can choose between the CDSS and other treatment methods that are still part of the standard. However, if a (partially) automated system proves to be superior to the conventional treatment method on the basis of evidence after the trial and transition phase, the non-use of the (partially) automated system may constitute a breach of duty [66]. The non-use of a (partially) automated system is in breach of duty of care if it is scientifically essentially uncontroversial and is supported by the vast majority of experts (physicians, scientists), specifically due to the fact that it is less risky or promises better healing chances than conventional treatment methods [50]. Aspects of reasonableness for physicians (e.g. acquisition costs) must also be taken into account.

If, on the other hand, CDSS represent a new examination or treatment method, physicians are subject to increased duties of care when deciding on the use of the system. In particular, they must carry out a risk-benefit assessment compared to conventional treatment methods that correspond to the current medical standard and review the medical indication during use. According to the result of the assessment, the physicians must choose between the CDSS and standard treatment methods within the framework of their freedom of therapy in such a way that, according to their professionally justified conviction, the best possible treatment is provided in the individual case [cf. 38].

4.4. Patient education and informed consent

4.4.1. In the case of CDSS that are used as an independent examination or treatment method within the framework of medical treatment and are already part of the medical standard, the same principles apply with regard to the medical duty to inform and the patient's consent as for other established examination and treatment methods. Doctors must inform patients about the essential circumstances of the treatment so that they can understand the significance and implications of the treatment and, in exercising their right to self-determination, make an informed decision as to whether they consent to the treatment (§§ 630c, 630e German Civil Code/BGB).

According to the principle of therapeutic freedom, the choice of the right treatment method and thus also the decision on the use of CDSS is fundamentally up to the attending physicians ([32]; see also Ch. 3.3.1. above). If not asked by the patient, they do not have to explain existing treatment alternatives (e.g. a conventional treatment method) if they use CDSS that corresponds to the medical standard of care as a treatment method. This does not apply if alternatives to the use of CDSS exist in the specific case and the alternative methods are equally medically indicated and customary and can lead to significantly different burdens, risks or chances of recovery (§ 630e (1), sent. 3 German Civil Code/BGB). This means that information about treatment alternatives must be provided if the alternative method either has a lower risk burden for the patient with an equivalent prospect of cure or success or promises a greater prospect of cure or success with equivalent burdens and risks in terms of type and direction [33]. In this case, the patient's right to self-determination requires that they be informed about this and that the choice of method be left up to them [70]. Doctors must carefully examine whether there is a genuine alternative treatment on the basis of specialist medical knowledge.

4.4.2. CDSS, which represent an independent examination or treatment method and which are still being tested (new method), doctors only have to inform patients about, if they consider to use the CDSS – after due consideration (see chapter 4.5.) [cf. 34] – in treatment. Doctors must then inform patients that this is a new method that has not been practised for long, whose effectiveness has not yet been statistically proven and for which unknown risks, such as the lack of transparency and traceability of the system or the risk of distortions of the machine diagnosis or therapy decision, cannot be ruled out. Patients must also be thoroughly informed about the standard methods that can be considered. Patients must be in a position to carefully weigh up whether they would like to be treated according to the conventional method with known risks or according to the new method (CDSS), taking into account in particular the promised advantages and the risks

that are not yet known in every respect [35]. The decision for or against the use of the automated system is then up to the patient. In case of doubt, physicians must prove that they have provided proper information.

4.5. Liability of doctors for breach of duty of care when using CDSS

The use of CDSS by physicians can entail liability risks that to a large extent depend on the respective system and its area of application.

If physicians are accused of a breach of duty (non-observance of medical duties of care) in connection with the use of a (partially) automated system, they are liable (contractually acc. to § 280 in conjunction with § 630a ff. BGB (German Civil Code) and/or in tort acc. to § 630a ff. BGB (German Civil Code) depending on fault for attributable damage caused. Circumstances giving rise to liability may be, in addition to the use of a non-certified medical device, errors in the commissioning and use of CDSS. Such errors can occur, for example, if doctors do not sufficiently familiarise themselves with the functionality of CDSS prior to use, if they do not carry out a comprehensive risk-benefit assessment beforehand or if they entrust the use of CDSS to staff members without sufficiently instructing and monitoring them.

Cases in which the damage is due to a defect in the (partially) automated system, such as a manufacturing defect (e.g. defective production of individual CDSS contrary to the design and quality specifications of the manufacturer), a design defect (e.g. defective control algorithm due to defective design by the manufacturer) or an instruction defect (e.g. defective instructions for use by the manufacturer) must be considered separately. If systems are subject to the regulations for medical devices (see Ch. 4.1.) and, despite certification, have errors that endanger the health and safety of patients beyond an acceptable level, physicians are not liable in the event of damage if the error was not recognisable to them when exercising the necessary care (see Ch. 4.3.). As a rule, recognisability is likely to decrease as the degree of automation of the system increases. However, the manufacturer may be liable for defective CDSS under the Product Liability Act and under § 823 German Civil Code (BGB) in conjunction with the Medical Devices Act.

If the use of CDSS violates prohibition of discrimination under civil law within the meaning of § 19 of the General Equal Treatment Act (AGG), the disadvantaged patients have claims for removal, injunctive relief and damages. However, it has not been clarified whether and to what extent such a prohibition of discrimination can be violated by the use of CDSS. In any case, doctors are not liable for non-recognisable discrimination that emanates from a certified system and endangers the safety and health of patients (e.g. CDSS for skin cancer screening that detect skin changes in dark-skinned people less well without this being recognisable to the doctors).

The introduction of a special liability framework for AI system operators is currently being discussed at the EU level. Operators of high-risk AI systems should be held liable, regardless of fault, for all personal injury and damage to property caused by a physical or virtual activity or device driven by the AI system (strict liability). Operators should be under the obligation to take out appropriate liability insurance. Doctors who use CDSS may be operators of an AI system within the meaning of the draft regulation. Which systems pose a “high risk” is to be specified in the annex to the regulation. The annex to the current draft regula-

tion includes “autonomous robots”. For other AI systems that are not listed in the annex, fault-based liability of the operators should apply. Further developments remain to be seen.

4.6. Data protection and duty of confidentiality

4.6.1. The functional requirement of machine learning systems is to process personal health data. Data protection law (GDPR, BDSG) is applicable to the processing of personal data. In this respect, reference should be made to the “Notes and Recommendations on Medical Confidentiality, Data Protection and Data Processing in Medical Practice” of the German Medical Association [10]. The following only deals with the particularities of CDSS.

If CDSS falls under the scope of application of data protection law, health data processing is permissible if the data subjects expressly consent or a legal basis permits this. For the purpose of preventive health care, medical diagnostics and healthcare and treatment, data processing by doctors or other persons who are subject to a corresponding duty of confidentiality is also permissible without the explicit consent of the patient (Art. 9 (2) lit. h, (3) GDPR; § 22 (1) no. 1 lit. b BDSG). In this case, additional specific (protective) measures must be taken, such as documenting who was involved in processing the data, raising the awareness of those involved in the processing operations, appointing a data protection officer and pseudonymising and encrypting the data (see in detail § 22 (2) BDSG).

Before using CDSS, doctors should check whether they are required to carry out an assessment of the consequences of the intended processing operations for the protection of personal data (data protection impact assessment) in accordance with Art. 35 of the GDPR [cf. 22].⁸ In connection with the use of CDSS, users need support in the form of manuals from the CDSS manufacturers or other suitable service providers or professional associations.

Pursuant to Art. 22 (1) of the GDPR, the data subject has the right not to be subject to a decision based solely on automated processing which produces legal effects concerning them or similarly significantly affects them. This applies in particular to processing of health data (see Art. 22 (4) GDPR). AI systems that replace medical decisions may therefore only be used with the explicit consent of the patient after they have received the appropriate information. However, if a human decision intervenes, for example if doctors carry out a plausibility and accuracy check in individual cases, Art. 22 of the GDPR does not apply [68].

The right of data subjects to deletion (Art. 17 GDPR) can also take effect when self-learning systems are used. Revocation and deletion would deprive the systems of part of their training basis. However, revocation does not affect the lawfulness of processing what previously took place (Art. 7 (3) sent. 2 GDPR). Furthermore, the “black box” problem mentioned above could prove to be an advantage in this context, as the lack of traceability of the automated decision-making process can stand in the way of a subsequent personal allocation of the data entered; if the personal reference of data is eliminated, claims for revocation and deletion are ruled out. However, the details for regulatory clarification in this case have not yet been established.

⁸ Helpful here is the list of processing activities by the Data Protection Commission (DSK) for which a data protection impact assessment (DSFA) must be carried out (“must list”), version 1.1 dated 17.10.2018.

Overall, the use of AI-supported systems harbours uncertainties in terms of data protection law. In particular, it is unclear to what extent the processing of patient data for the purpose of individual diagnostics and therapy may at the same time serve the training and further development of CDSS inherent in the system and whether the patients must expressly consent to this. In general, it is uncertain how the conflict of the principle of data minimisation with the functional condition of AI systems, which are dependent on the processing of as much data as possible to facilitate the learning ability of the systems, can be legally resolved.

4.6.2. Medical confidentiality is central to the use of intelligent software solutions in medical practice, e.g. when third party providers are to be given access to confidential patient information. The disclosure of patient information to third parties, who are involved in the professional activity of the primary professional who holds the confidentiality, is permitted (e.g. service for the installation and maintenance of software or IT systems). What must be determined on a case-to-case basis is the issue of who shall be included among these contributors. The contractual arrangements between hospitals, doctors' practices and third-party providers should therefore always be screened to see whether they adequately take medical confidentiality into account when using intelligent software solutions.

5. Summary and recommendations

ZEKO welcomes it when the CDSS contribute to an improvement in the quality and effectiveness of diagnostics, therapy, prognosis and prediction. When AI-based CDSS are used in medicine, doctors are supported in tasks that were previously reserved for humans. In contrast to conventional, regularly programmed software and hardware systems, AI is partly characterised by machine-learning systems (neuronal networks, deep learning), which generate decision recommendations that are likely to become less and less traceable or explainable in terms of their origin in the future.

Consequently, ZEKO recommends that the following considerations in particular be taken into account when using such systems:

- Medical action is always dedicated to the individual patient. Doctors are responsible for ensuring that the use of CDSS is carried out with the aim of improving patient care.
- The responsibility and accountability for diagnosis, indication and therapy always lies with the doctors and must not be surrendered to a CDSS system. The threshold between decision support and automated decision-making must not be crossed.
- As a rule, optimal treatment results are only achieved through the interaction of CDSS and medical experience. This is because the manifestation of medical judgement, which cannot be replaced both in terms of medical expertise and on a communicative level, is necessary for optimal treatment results. Only doctors are able to understand the clinical picture in its entirety and take into account psychological and emotional factors, which are important for the diagnosis and can also be decisive for an appropriate therapy. A devaluation of a doctor's experience and a weakening of trust in the relationship between doctor and patient must therefore be prevented while emphasising the value of cooperation.
- Doctors should regularly inform themselves about new CDSS in their respective fields and keep track of their development.
- Doctors should be aware that CDSS may contain errors and

biases and that the processes of the systems are often not sufficiently comprehensible. This entails the risk of erroneous diagnostic and therapeutic recommendations. Physicians should therefore verify the plausibility of (partially) automated decision recommendations from CDSS.

- The use of CDSS by doctors in the context of medical treatment requires diligence and involves liability risks. Before using a (partially) automated system, physicians must therefore familiarise themselves with how it functions and the regulatory framework for its use.
- Before using CDSS, doctors must check whether the relevant system represents a new or standard method of medical examination or treatment. In the case of new methods, the medical decision on the use of the AI system is subject to increased duties of care.
- Doctors only have to inform patients about CDSS that are still being tested (new method) if – after due consideration – they want to use the system in treatment. Doctors must then inform patients that this is a new method that has not been in practice for long, whose effectiveness has not yet been statistically proven and for which unknown risks, such as the lack of transparency and traceability of the system or the risk of distortions of the machine diagnosis or therapy decision, cannot be ruled out. Patients must also be comprehensively informed about the standard methods that are available for consideration. The decision for or against the use of the automated system is then up to the patient.
- Since CDSS usually process personal data, doctors must observe the regulations of data protection law and the duty of confidentiality.
- Physicians share responsibility for quality assurance with regard to patient care. They should contribute to the further development of CDSS by providing feedback on their experiences and report undesirable developments at an early stage, especially in specific patient care processes (process quality). If, in the course of their professional activities, physicians become aware of malfunctions or inadequacies in the information provided by the manufacturers that could significantly endanger the life or limb of patients, they must report this to the Federal Institute for Drugs and Medical Devices (BfArM).
- The requirement for physician responsibility in the use of the systems must not lead to a situation where the overall burden of the introduction and use of CDSS lies with the individual physicians. Rather, when approving and using these systems, those responsible at the meso and macro levels have a central role to play in ensuring that there is a level of trust in the system that can be relied upon by doctors and patients alike.
- AI-based CDSS should be systematically evaluated with regard to their potential benefits and harms, including ethical, legal and social implications, before their widespread use in routine care. Particular importance is attached to the prospective validation of the systems within the framework of controlled clinical studies.
- Stigmatisation and discrimination of patients must be mitigated through rigorous and robust validation of models, training data and findings, as well as appropriate policy and regulatory measures to protect fundamental rights. Doctors must be confident that the training data used in the development of CDSS is free from bias. They will still have an observational role in the further development and subsequent monitoring of CDSS,

for which appropriate feedback procedures should also be established. They should undertake efforts to ensure that CDSS is supported by a database that is representative of all population groups.

- Doctors must ensure that they have the appropriate digital competences for the use of CDSS. To achieve this, hospital operators at the meso level must appropriately prepare their staff for the use of the CDSS, e.g. by offering suitable further training measures. At the macro level, increased instruction in digital competencies as an important interdisciplinary topic in medical education and training is of central importance, so that those working in the health sector can assess the benefits and risks of AI-based CDSS and use them in a patient-oriented manner.
- Medical societies and other creators of guidelines should address the use of CDSS in their fields at an early stage and take it into consideration, e.g. in their guidelines, in order to prevent different standards from being established in various sectors.
- There is a considerable need for research to develop measures for the explainability and comprehensibility of AI-based CDSS, but also for the containment of risks and the regulation of CDSS, which recognises the aforementioned technical, ethical, legal and social areas of tension as a design issue and offers a constructive solution-based approach.
- It is also necessary to have an accompanying social debate about these issues and the question of the involvement of doctors and patients in the development of these systems.

Literature

1. Acatech – Deutsche Akademie der Wissenschaften (Hg.): Machine Learning in der Medizintechnik. Analyse und Handlungsempfehlungen. München 2020, <https://www.acatech.de/publikation/machine-learning-in-der-medizintechnik/download-pdf?lang=de> (letzter Zugriff 04.03.2021).
2. Aschwanden C: Artificial Intelligence Makes Bad Medicine Even Worse. A new study out from Google seems to show the promise of AI-assisted health care. Actually, it shows the threat. WIRED 2020, <https://www.wired.com/story/artificial-intelligence-makes-bad-medicine-even-worse/> (letzter Zugriff 04.03.2021).
3. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al.: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 2020; 58: 82–115.
4. Bates M: Direct-To-Consumer Genetic Testing: Is the Public Ready for Simple, At-Home DNA Tests to Detect Disease Risk? IEEE Pulse 2018; 9(6): 11–4.
5. Baumgartner R: Künstliche Intelligenz in der Medizin: Diskriminierung oder Fairness? In: Bauer G, Kechaja M, Engelmann S, Haug L (Hg.): Diskriminierung und Antidiskriminierung: Beiträge aus Wissenschaft und Praxis. Bielefeld: transcript 2021; 149–164.
6. Birgmeier J, Haeussler M, Deisseroth CA, et al.: AMELIE speeds Mendelian diagnosis by matching patient phenotype and genotype to primary literature. Sci Transl Med 2020; 12(544).
7. Bjerring JC, Busch J: Artificial Intelligence and Patient-Centered Decision-Making. Philos. Technol. 2020.
8. Bundesärztekammer: Stellungnahme „Präzisionsmedizin: Bewertung unter medizinisch-wissenschaftlichen und ökonomischen Aspekten“. Dtsch Arztebl 2020; 117(22–23): A-1184 / B-1000. DOI: 10.3238/baek_sn_praezision_2020.
9. Bundesärztekammer: Stellungnahme „Wissenschaftlichkeit als konstitutionelles Element des Arztberufes“. Dtsch Arztebl 2020; 117(4): A-176 / B-160 / C-156. DOI: 10.3238/baek_wb_sn_wiss2020.
10. Bundesärztekammer, Kassenärztliche Bundesvereinigung: „Hinweise und Empfehlungen zur ärztlichen Schweigepflicht, Datenschutz und Datenverarbeitung in der Arztpraxis“. Dtsch Arztebl 2018; 115(10): A-453 / B-395 / C-395. DOI: 10.3238/arzt-ebl.2018.ds01.
11. Bundesärztekammer: Thesen zur Weiterentwicklung der ärztlichen Patientenversorgung durch Digitalisierung. 2021 (noch nicht veröffentlicht).
12. Cabitza F, Rasoini R, Gensini GF: Unintended Consequences of Machine Learning in Medicine. JAMA 2017; 318(6): 517–8.
13. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K: Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019; 28(3): 231–7.
14. Chen P-J, Li T, Sun T-T, et al.: Clinical Decision Support for High-Risk Stage II Colon Cancer: A Real-World Study of Treatment Concordance and Survival. Dis Colon Rectum 2020; 63(10): 1383–92.
15. Cicero M, Bilbily A, Colak E, et al.: Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs. Invest Radiol 2017; 52(5): 281–7.
16. Dettling H-U: Künstliche Intelligenz und digitale Unterstützung ärztlicher Entscheidungen in Diagnostik und Therapie. Pharm R 2019; 12:633.
17. Deutscher Ethikrat (Hg.): Stellungnahme „Big Data und Gesundheit: Datensouveränität als informationelle Freiheitsgestaltung“. Berlin 2017.
18. Di Nucci E: Should we be afraid of medical AI? J Med Ethics 2019; 45(8): 556–8.
19. Europäische Kommission: Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). COM(2021) 206 final vom 21.04.2021.
20. Europäische Kommission: Weissbuch „Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen“. COM(2020) 65 final vom 19.02.2020.
21. Froomkin AM, Kerr IR, Pineau J: When Als Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning. Arizona Law Rev 2019; 61:33.
22. Frost Y: Künstliche Intelligenz in Medizinprodukten und damit verbundene medizintechnik- und datenschutzrechtliche Herausforderungen. MPR 2019(4): 117.
23. Fuchs T: Verteidigung des Menschen: Grundfragen einer verkörperten Anthropologie 2020.
24. Geisler LS: Arzt und Patient – Begegnung im Gespräch: Wirklichkeit und Wege. 5th ed. Frankfurt am Main: pmi Verl. 2008.
25. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G: Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. JAMA Intern Med 2018; 178(11): 1544–7.
26. Gurovich Y, Hanani Y, Bar O, et al.: Identifying facial phenotypes of genetic disorders using deep learning. Nat Med 2019; 25(1): 60–4.
27. Haenssle HA, Fink C, Schneiderbauer R, et al.: Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 2018; 29(8): 1836–42.
28. Hannun AY, Rajpurkar P, Haghpanahi M, et al.: Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med 2019; 25(1): 65–9.
29. Helle K: Intelligente Medizinprodukte: Ist der geltende Rechtsrahmen noch aktuell? Medizinrecht 2020; 38(12): 993–1000.
30. High-Level Expert Group on Artificial Intelligence (HLEG): A definition of Artificial Intelligence: main capabilities and scientific disciplines. 2019, <https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines> (letzter Zugriff 04.03.2021).
31. High-Level Expert Group on Artificial Intelligence (HLEG): Ethics guidelines for trustworthy AI. 2019, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (letzter Zugriff 04.03.2021).
32. Katzenmeier, in: Hau/Poseck, BeckOK BGB, 57. Edition Stand: 01.02.2021, § 630e Rn. 25.
33. Katzenmeier, in: Hau/Poseck, BeckOK BGB, 57. Edition Stand: 01.02.2021, § 630e Rn. 26.
34. Katzenmeier, in: Hau/Poseck, BeckOK BGB, 57. Edition Stand: 01.02.2021, § 630e Rn. 28.
35. Katzenmeier, in: Hau/Poseck, BeckOK BGB, 57. Edition Stand: 01.02.2021, § 630e Rn. 29.
36. Kieser M, Caliebe A, Burger H-U, Knoerzer D: Big Data in der klinischen Forschung: Vieles ist noch Wunschdenken. Dtsch Arztebl 2019; 166(35–36): A 1534–9.
37. Liang H, Tsui BY, Ni H, et al.: Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nat Med 2019; 25(3): 433–8.
38. Lipp V: Heilversuch und medizinische Forschung, Kap. XIII Rn. 31. In: Laufs A, Katzenmeier C, Lipp V: Arztrecht. 8th ed. München: C.H. Beck 2021; 553.
39. Liu X, Faes L, Kale AU, et al.: A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The Lancet Digital Health 2019; 1(6): e271–e297.
40. Liu Y, Kohlberger T, Norouzi M, et al.: Artificial Intelligence-Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for Pathologists. Arch Pathol Lab Med 2019; 143(7): 859–68.
41. McDougall RJ: Computer knows best? The need for value-flexibility in medical AI. J Med Ethics 2019; 45(3): 156–60.
42. McKinney SM, Sieniek M, Godbole V, et al.: International evaluation of an AI system for breast cancer screening. Nature 2020; 577(7788): 89–94.
43. McNamara DM, Goldberg SL, Latts L, et al.: Differential impact of cognitive computing augmented by real world evidence on novice and expert oncologists. Cancer Med 2019; 8(15): 6578–84.
44. Mehra MR, Ruschitzka F, Patel AN: Retraction – Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. The Lancet 2020; 395(10240): 1820.

45. Miguel Beriaín I de: Should we have a right to refuse diagnostics and treatment planning by artificial intelligence? *Med Health Care Philos* 2020; 23(2): 247–52.
46. Mittelstadt B: Principles alone cannot guarantee ethical AI. *Nat Mach Intell* 2019; 1(11): 501–7.
47. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L: The ethics of algorithms: Mapping the debate. *Big Data & Society* 2016; 3(2): 205395171667967.
48. Morley J, Joshi I: Developing effective policy to support Artificial Intelligence in health and care. *Eurohealth* 2019; 25(2): 11–4.
49. Morley J, Machado CCV, Burr C, et al.: The ethics of AI in health care: A mapping review. *Soc Sci Med* 2020; 260: 113172.
50. Mühlböck L, Taupitz J: Haftung für Schäden durch KI in der Medizin. *AcP* 2021; 221(1/2): 179–218.
51. Neugebauer M, Ebert M, Vogelmann R: A clinical decision support system improves antibiotic therapy for upper urinary tract infection in a randomized single-blinded study. *BMC Health Serv Res* 2020; 20(1): 185.
52. Obermeyer Z, Powers B, Vogeli C, Mullainathan S: Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019; 366(6464): 447–53.
53. Olakotan O, Mohd Yusof M, Ezat Wan Puteh S: A Systematic Review on CDSS Alert Appropriateness. *Stud Health Technol Inform* 2020; 270: 906–10.
54. O’Neil C: *Angriff der Algorithmen: Wie sie Wahlen manipulieren, Berufschancen zerstören und unsere Gesundheit gefährden*. 2nd ed. München: Carl Hanser Verlag 2018.
55. Orth M, Luppá PB: Direct-To-Consumer-Testing: Fluch oder Segen für die Patienten? *Dtsch Arztebl* 2015; 112(5): A-174 / B-154 / C-154. <https://www.aerzteblatt.de/archiv/167563/Direct-To-Consumer-Testing-Fluch-oder-Segen-fuer-die-Patienten> (letzter Zugriff 18.05.2021).
56. Orwat C: Diskriminierungsrisiken durch Verwendung von Algorithmen: Eine Studie erstellt mit einer Zuwendung der Antidiskriminierungsstelle des Bundes. 2019, https://www.antidiskriminierungsstelle.de/SharedDocs/Downloads/DE/publikationen/Expertenstudie_diskriminierungsrisiken_durch_verwendung_von_algorithmen.html (letzter Zugriff 05.03.2021).
57. Parasuraman R, Manzey DH: Complacency and bias in human use of automation: an attentional integration. *Hum Factors* 2010; 52(3): 381–410.
58. Ploug T, Holm S: The four dimensions of contestable AI diagnostics – A patient-centric approach to explainable AI. *Artif Intell Med* 2020; 107: 101901.
59. Ploug T, Holm S: The right to refuse diagnostics and treatment planning by artificial intelligence. *Med Health Care Philos* 2020; 23(1): 107–14.
60. Price ND, Magis AT, Earls JC, et al.: A wellness study of 108 individuals using personal, dense, dynamic data clouds. *Nat Biotechnol* 2017; 35(8): 747–56.
61. Raab CD: Information privacy, impact assessment, and the place of ethics. *Computer Law & Security Review* 2020; 37: 105404.
62. Riedl D, Schüßler G: The Influence of Doctor-Patient Communication on Health Outcomes: A Systematic Review. *Z Psychosom Med Psychother* 2017; 63(2): 131–50.
63. Ross C, Swettlitz I: IBM’s Watson supercomputer recommended ‘unsafe and incorrect’ cancer treatments, internal documents show. 2018, <https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/> (letzter Zugriff 05.03.2021).
64. Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen (SVR): Digitalisierung für Gesundheit: Ziele und Rahmenbedingungen eines dynamisch lernenden Gesundheitssystems. Gutachten 2021. <https://www.svr-gesundheit.de/gutachten/gutachten-2021/> (letzter Zugriff 12.04.2021).
65. Strauß S: *Privacy and Identity in a Networked Society – Refining Privacy Impact Assessment*. London: Routledge 2019.
66. Taupitz J: Medizinische Informationstechnologie, leitliniengerechte Medizin und Haftung des Arztes. *AcP* 2011; 211(3): 352–94.
67. Tupasela A, Di Nucci E: Concordance as evidence in the Watson for Oncology decision-support system. *AI & Soc* 2020; 35(4): 811–8.
68. v. Lewinski, in: Wolff/Brink (Hg.), BeckOK-Datenschutzrecht, Stand: 01.02.2021, GDPR, Art. 22 Rn. 21.
69. Walsh SLF, Humphries SM, Wells AU, Brown KK: Imaging research in fibrotic lung disease; applying deep learning to unsolved problems. *Lancet Respir Med* 2020; 8(11): 1144–53.
70. 70. Walter, in: Benecke, BeckOGK, BGB, Stand: 15.03.2020, § 630e Rn. 12.
71. Wang F, Casalino LP, Khullar D: Deep Learning in Medicine – Promise, Progress, and Challenges. *JAMA Intern Med* 2019; 179(3): 293–4.
72. Watson DS, Krutzinna J, Bruce IN, et al.: Clinical applications of machine learning algorithms: beyond the black box. *BMJ* 2019; 364: i886.
73. Wiesing U, Bormann F-J: Urteilkraft und Gewissen: die ärztliche Entscheidung im Zeitalter der evidence-based medicine. In: Bormann F-J, Wetzstein U (Hg.): *Gewissen. Dimensionen eines Grundbegriffs Medizinischer Ethik*. De Gruyter, Berlin/Boston 2014; 39–49.
74. Wildner M: Diagnose, Prognose, Prädiktion. *Gesundheitswesen* 2014; 76(4): 185–6.
75. Winkler JK, Sies K, Fink C, Toberer F, Enk A, Haenssle HA: Digitalisierte Bildverarbeitung: künstliche Intelligenz im diagnostischen Einsatz. *Forum* 2020; 35(2): 109–16.

Members of the Central Ethics Commission

Prof. Dr. theol. Franz-Josef Bormann, Tübingen
 Prof. Dr. jur. Frauke Brosius-Gersdorf, LL.M., Hannover
 Prof. Dr. theol. Elisabeth Gräß-Schmidt, Tübingen
 Prof. Dr. med. Wolfram Henn, Homburg
 Prof. Dr. (TR) Dr. phil. et med. habil. İlhan İlkelic, M.A., Istanbul
 Dr. phil. Julia Inthorn, Hannover
 Prof. Dr. med. Dipl.-Soz. Tanja Krones, Zürich/CH
 Prof. Dr. phil. Dirk Lanzerath, Bonn (Coordinator)
 Prof. Dr. jur. Dr. h. c. Volker Lipp, Göttingen
 Prof. Dr. med. Georg Marckmann, MPH, München
 Dr. med. Stephan M. Probst, Bielefeld
 Prof. Dr. med. Dr. phil. Sabine Salloch, Hannover
 Prof. Dr. med. Jan Schildmann, Halle (Saale)
 Prof. Dr. phil. Ingrid Schneider, Hamburg
 Prof. Dr. jur. Jochen Taupitz, Mannheim (Chair)
 Prof. Dr. med. Dr. phil. Eva Winkler, Heidelberg

Correspondence address

Zentrale Ethikkommission bei der Bundesärztekammer
 Herbert-Lewin-Platz 1, 10623 Berlin
 E-Mail: zeko@baek.de
<https://www.zentrale-ethikkommission.de>