Responsibility and decision-making authority in using clinical decision support systems: an empirical-ethical exploration of German prospective professionals' preferences and concerns

Florian Funer [1,2] Wenke Liedtke [3] Sara Tinnemeyer,[1] Andrea Diana Klausen,[4] Diana Schneider [5] Helena U Zacharias,[6] Martin Langanke,[3] Sabine Salloch [1]

[1]Institute of Ethics, History and Philosophy of Medicine, Hannover Medical School, Hannover, Germany
[2]Institute of Ethics and History of Medicine, Eberhard Karls University Tübingen, Tübingen, Germany
[3]Department of Social Work, Protestant University of Applied Sciences RWL, Bochum, Germany
[4]Institute of Medical Informatics, RWTH Aachen University, Aachen, Germany
[5]Competence Center Emerging Technologies, Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe, Germany
[6]Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School, Hannover Medical School, Hannover, Germany

Correspondence to
Dr Florian Funer, Institute of Ethics and History of Medicine, Eberhard Karls University Tübingen, Tübingen, Baden-Württemberg, Germany; florian.funer@uni-tuebingen.de

ABSTRACT

Machine learning-driven clinical decision support systems (ML-CDSSs) seem impressively promising for future routine and emergency care. However, reflection on their clinical implementation reveals a wide array of ethical challenges. The preferences, concerns and expectations of professional stakeholders remain largely unexplored. Empirical research, however, may help to clarify the conceptual debate and its aspects in terms of their relevance for clinical practice. This study explores, from an ethical point of view, future healthcare professionals' attitudes to potential changes of responsibility and decision-making authority when using ML-CDSS. Twenty-seven semistructured interviews were conducted with German medical students and nursing trainees. The data were analysed based on qualitative content analysis according to Kuckartz. Interviewees' reflections are presented under three themes the interviewees describe as closely related: (self-)attribution of responsibility, decision-making authority and need of (professional) experience. The results illustrate the conceptual interconnectedness of professional responsibility and its structural and epistemic preconditions to be able to fulfil clinicians' responsibility in a meaningful manner. The study also sheds light on the four relata of responsibility understood as a relational concept. The article closes with concrete suggestions for the ethically sound clinical implementation of ML-CDSS.

BACKGROUND

Facing complex decision-making situations is an integral part of healthcare's daily practice. Clinical decision support systems (CDSSs) have been supposed to enhance and accelerate professional decision-making over the last 30 years.[1] Given the plethora of patient data provided by electronic health records, machine learning-driven CDSSs (ML-CDSSs) seem particularly promising for future routine and emergency care. So far, ML-CDSSs have been successful mainly in the field of (imaging) diagnostics, but their use in prediction, prognostics and therapy also offers many opportunities. In response to the recent progress, ethical debates on the requirements for the clinical implementation of ML-CDSS have been unleashed, mainly about topics such as patient safety, privacy, data ownership, opacity/transparency/explainability, biases, trustworthiness, validity and reliability.[2–5]

Many of these important ethical aspects culminate in theoretical and practical questions about responsibility and its allocation when such technologies are used.[6–13] Often, in line with the designation as a support system, the importance that such systems should assist professionals in their decision-making has been highlighted. The ML-CDSSs pursue the goal of supporting professionals in their decision-making by making medical recommendations[14–16] and covering their existing information needs as human experts,[12 17 18] 'for example, by providing evidence that would have otherwise not been available within a reasonable time frame'.[19] Even if ML-CDSSs seem particularly suitable for this goal in view of the increasing technical capacities and the professional's epistemic dependency, the question arises whether and to what extent it is realiter still the professional who is responsible for the decision-making.

Responsibility is considered to be a relational concept in ethics. There is an ongoing discussion on the exact number of relata constituting this relationship.[13 20–24] While older contributions to the discussion of ethics of responsibility predominantly refer to the subject and object of responsibility and suffer from a lack of precision regarding the role of the normative standard[25] newer concepts expand the structure of the relation of responsibility in order to mirroring a holistic view and to clarify the normative implications.[20 23] But, according to a broad consensus, at least four relata are essential: subject of responsibility(A), object of responsibility(B), addressees of responsibility(C) and normative standards(D).[7 26] Thus, the relationship of responsibility can be expressed as:

A is responsible to C for B because of D.

While the subject and the addressee of responsibility can be either a person or institution, the object is an action and/or its consequences. The normative basis for attribution of responsibility, the normative standard, is a set of explicit (eg, legal) or implicit rules. Based on this relational concept, responsibility can be adopted in at least two directions: first, retrospectively, after harm has been caused; who is responsible, who should be held accountable and who is liable for the harm caused? With significant changes in the doctor–patient relationship since the 1970s one can detect a general move from a paternalistic understanding of the subject of responsibility to the concept of shared decision-making

which implies a co-responsibility of doctors and patients; as well as a co-responsibility to developers and manufacturers.[27] This perspective is important not to simply assign blame, but to trace causes and prevent future reoccurrences more effectively.[7 28] Second, prospectively, insofar as moral requirements must be formulated and justified to prevent harm from the outset.[9] How, for example, should a healthcare worker act if (s)he does not agree with the ML-CDSS recommendation? Such specific decision-making situations (sometimes called 'peer disagreement'[10 18 29])[i], in which the professional's clinical judgement differs from the recommendation of the ML-CDSS and these judgements are incompatible due to their epistemic characteristics (ie, opacity), seem to be particularly problematic regarding the attribution and assumption of responsibility.[10 12 18 27 29] In these instances, it can be necessary to train professionals' suitable competencies 'as a safeguard to decrease the risk of harm in cases of cognitive misalignment between the physicians and the AI [=artificial intelligence] system—when an AI output cannot be confirmed (verified or falsified)'.[9]

Despite the extensive theoretical literature on ML-CDSS, responsibility and decision-making, to the best of our knowledge, little evidence currently exists on professionals' attitudes regarding these topics.[30] Empirical research mostly examines factors related to doctors' acceptance of ML-CDSSs and its promotion,[31 32] whereas aspects of moral responsibility or decision-making play only a minor role.

Therefore, responsibility and its attribution in the overall context of ML-CDSS merits more explicit research and analysis. Given the range of the existing theoretical debates, the inquiry of clinical stakeholders' views and their underlying reasoning may address important aspects of professional practice and adjust them regarding their theoretical relevance for clinical decision-making. This study explores the opinions, preferences and concerns of future healthcare professionals about potential changes of responsibility allocation and decision-making authority when using ML-CDSS from an ethical point of view. Based on our information, this study is the first to collect professionals' attitudes while confronting them with different case vignettes of ML-CDSS, thereby, enabling comparisons between different types of decision support. We consider different healthcare fields (surgery, nephrology, long-term care) and different degrees of decision support (alert, information, concrete recommendation for action). Furthermore, the study addresses future physicians and future nurses, which allows a closer look at the interprofessional similarities and differences. In this respect, our study represents an important enrichment of the theoretical discussion on responsibility and situations of disagreement that has taken place up to this point.

METHODS

We conducted a qualitative interview study to investigate how medical students and nursing trainees expect and assess responsibility and decision-making authority in their future clinical practice when using ML-CDSS. We used semistructured interviews with 15 medical students and 12 nursing trainees of a German maximum-care hospital. For more information on the personal characteristics of the interviewer as well as the coders (credentials, occupation, gender, experience), see online supplemental 1.

---

[i]'Peer disagreement' is used in literature to express that one expert opinion (the professional's) stands parallel to another expert opinion (that of the CDSS) and has equal epistemological weight.

Data collection

A convenience sampling was used for data collection. Interview partners were included if they belonged to the groups of interest (medical students: fourth/fifth year of study; nursing trainees: second/third year of training), were ≥18 years old and had sufficient knowledge of German. There was no relationship established between the participants and the interviewer prior to the commencement of the study. Participants were informed in advance about the topic 'Digital Decision Support Systems and Digitisation in Medicine' before the interviews were conducted. There was no relationship established between interviewer and participants; only information about the interviewer's current affiliation and educational background was provided to the participants. All interviews were conducted via video calls between June and October 2021. Participants were generally at home and alone during the interview. They received a common expense allowance for participation. None of the participants dropped out after study enrolment.

The interviewer used a semistructured interview guide including case vignettes (see online supplemental 2). In the case of medical students, the guide included two case vignettes with ML-CDSSs to support doctors (intra-abdominal surgical navigation and the prognosis and therapy planning of chronic kidney diseases), and in the case of nursing trainees, one case vignette of a CDSS to support the monitoring of home-ventilated patients. The ML-CDSSs were selected regarding their diversity in terms of the clinical application field (surgery, nephrology, long-term care) and their degree of support (guidance for incision lines; information, prognosis estimation and therapy planning; alarm and intervention recommendation). A broad concept of responsibility was chosen due to the exploratory objective. The case vignettes were uniformly accompanied by non-theory-based questions about prospective and retrospective dimensions of responsibility. Audiorecordings and field notes were made to document the interviews. Data collection was terminated when informational saturation was reached, that is, when additional interviews did not provide any additional information about the research question.

Data analysis

Interviews were anonymised and transcribed ad verbatim. Transcripts were not sent to the participants for review. Data analysis relied on qualitative content analysis according to Kuckartz,[33] which is a multistage procedure combining inductive category building along data and theoretically derived categories. The coding system was developed collaboratively, starting from specific passages in the data to identify recurring themes and concepts. Topics typical for the research question were drawn from the literature and their occurrence in the data was investigated. We clarified coding rules for the initial coding categories and identified exemplary passages (see online supplemental 1). The coding system was constantly revised and expanded. After the coding system remained the same and the redundancy of findings did not contribute anything substantially new, we assumed theoretical saturation. The software MAXQDA (2020) was used to support the data analysis. Any ambiguities and potential disagreements were discussed critically between the first and last author and decided by consensus.

The focus during data analysis was directed at specific topics, such as the research question of the present article. All codes related were selected from the coding system. Finally, in an iterative process, types and subtypes were identified, suitable example codes were selected, translated from German and included in the article. The presentation of methods and results was guided

Table 1	Sociodemographic data of interviewees	
	Medical students	**Nursing trainees**
No	15	12
Average age (range, median)	25.5 years (23–36, 29.5)	25 years (20–50, 35)
Gender (self-reported)	8 ♀/7 ♂	10 ♀/2 ♂

by the Consolidated Criteria for Reporting Qualitative Research (COREQ).[34]

RESULTS

The interviews lasted an average of 51:26 min (with a range from 29:44 to 75:37 min) and the interviewees had the following sociodemographic characteristics (table 1).

In line with the interview guide, the ML-CDSSs were introduced to the interviewees, and they were confronted with two clinical scenarios: first, harm caused due to an erroneous ML-CDSS recommendation, and second, an ML-CDSS recommendation differing from their own professional judgement. Thus, they were asked about responsibility—retrospectively and prospectively—in a situation of (potentially) harmful treatment. The respondents' answers regarding these scenarios can be grouped into three strongly interrelated categories: (self-)attribution of responsibility, decision-making authority and need of (professional) experience.

(Self-)attribution of responsibility

The causation of errors and the assignment of responsibility for those errors is described as significant. This is seen as particularly difficult when ML-CDSSs are used: 'this question of when a mistake happens, who's to blame, the nurse, the person who made the robot [=CDSS], or the hospital, so that's, I think, one of the biggest complications I could imagine right now' (TI-6)[ii]. Since the error's originator is often not clearly known, interviewees hold that responsibility would lie with several entities and cite this 'shared responsibility' (SI-15) or 'joint failure' (SI-6).

Interviewees generally mention the following subjects of responsibility: developers/providers, regulatory control instances, healthcare institutions/supervisors and clinical professionals. Consensus exists that ML-CDSS could not bear responsibility. Instead, interviewees are concerned that colleagues could invoke ML-CDSS as an excuse and 'shift responsibility to the system' (SI-15).

Developers and providers were seen to be responsible for reliable functioning and would, therefore, be accountable in situations where the cause of the damage is 'faulty programming' (TI-10) or 'faulty prognosis based on a faulty weighting of statistics' (SI-8). Regulatory instances and purchasing institutions, such as hospitals or nursing services, are seen as additional assurances of reliability. However, the institutional responsibility depends significantly on the concrete use directive: exemplarily, if employees are required to use ML-CDSS or the latter are used for dealing with staff shortages, respondents determine more that the institution bears a greater responsibility. Some interviewees even recognise the risk of 'coercion': 'If I am now forced to use this support system and I actually

don't feel safe with it [...], then perhaps the hospital management with its guidelines would somehow also be responsible in a broader sense' (SI-2).

Respondents emphasise a professional's 'final responsibility' for decisions. One student underlines the merely supporting character of ML-CDSS: '[It] is supposed to support you in making your decisions, but ultimately you are the person who bears the risk of what decision you make' (SI-2), and 'it's my free decision whether I make the cut or not. It's not like it's forcing me to do it' (SI-2). More closely, final responsibility is characterised as the ability to critically scrutinise recommendations before action: 'And this is now a support, a tool, and I have to check and evaluate or question this tool again and again' (SI-8). The consolidation and interpretation is seen as integral part of a doctors' task: 'I think it always needs the one person who can somehow connect everything together a little bit and who then also takes responsibility for interpreting something out of it' (SI-9). Nursing trainees describe the final responsibility comparably but with a stronger reference to the caring relationship and the well-being of their patients: 'Yes, it's always still the nurse, because a device like that is all well and good, but patient observation and such is still the main task, so it's still my responsibility whether the person survives or not or whatever happens to them' (TI-1).

Although the participants see themselves as professionally responsible for treatment decisions, some problematise—in view of the complexity of ML-CDSS—that they could be no longer in a position to fulfil it: 'but if I have so much data that I can no longer keep track of it myself as a doctor, then I can also no longer actually control this algorithm' (SI-8).

Decision-making authority and coping strategies

When interviewees were asked how they would deal with a situation in which the recommendation of the ML-DCSS differs (significantly) from their own judgement, different rationales emerged.

Some point out a need for open-mindedness among human decision-makers, that ML-CDSS could perform some tasks better, and that proof of 'scientific quality criteria,' such as failure rates, would be crucial for risk assessment: 'If it has really been shown that my [decision] is usually worse than [that of] the AI, and, thus, I end up accepting fewer errors in exchange for preventing many errors on my part, then it was still the right decision to follow' (SI-6).

Contrarily, some call for human control to assume responsibility for decisions: 'to work with it, I would still like to have my complete background knowledge. I would still like to be able to control what I do and what the device tells me. So, I wouldn't want to just blindly rely on it' (SI-10). Others formulate the same fact as professionals urgently being the final bearer of decisions: 'The primary role of physicians will be not to let themselves get screwed, but to keep an eye on the fact that the final decision is made by people' (SI-11).

In order to explain the importance of taking the final decision, professionals state their necessary ability to justify themselves: 'We always have to justify ourselves for what we do. [...] If I relied solely on the app without looking at the scientific basis for it, then it's my fault' (SI-13). No longer being able to decide about the correctness of an ML-CDSS recommendation is seen as a potential danger: 'That means that at some point, as a doctor who has an overview of this and can assess it again, I am in a certain way disengaged. And I just have to concentrate completely, just like the patient, on this app' (SI-8).

[ii] TI stands for interviews with nursing trainees; SI stands for interviews with medical students.

A dissenting ML-CDSS recommendation would compromise the professional's belief in her/his judgement: 'I think I would first check again all the data that I have entered. Then, of course, I would also question myself, that is, I would question myself on what basis do I come to the other conclusion. And there is, of course, again the question, how much experience do I have with the disease and with the course of the disease and on which data is the algorithm based' (SI-8). In order to resolve discrepant recommendations, joint deliberation could usually help, as it does with colleagues: 'Well, of course, I would prefer to ask […], so in the best case, the system could somehow explain to me how it came to this decision […], that I can just reassure myself' (SI-8).

However, if the ML-CDSS does not provide an explanation, different scoping strategies will be chosen. Some would prefer a consultation with (more experienced) colleagues or superiors: 'I would probably ask someone else again, because then it's basically opinion against opinion and then a third opinion is perhaps quite good to hear again' (SI-10). Given the undecidability about the 'correct' recommendation, other interviewees would communicate both versions to patients: 'Then I would, I think, openly explain the discrepancy to the patient. So, I would say, on the one hand, that's this algorithm, it comes to that result. But I personally, from my clinical experience, would see this rather positively' (SI-8).

Need for (professional) experience

Respondents initially underline ML-CDSS' potential to provide support to clinicians with less experience: 'I think the device is actually quite good if you haven't been qualified for a long time, so if you're sort of freshly qualified and you're coming to the ward' (TI-4). However, over-reliance on incorrect support is suspected especially for inexperienced professionals: 'So, if you're very inexperienced, you're more likely to stick to those kinds of systems than if you're more experienced' (SI-14). A differing ML-CDSS recommendation is compared with dealing with supervisors: 'but you don't necessarily contradict your boss, especially as a beginner. And so, then I could imagine it a bit similar with the device as well' (SI-15).

The bearing of responsibility for decision-making depends largely on sufficient (clinical) experience: 'The question is to what extent you can still decide for yourself if the robot [=CDSS] really, let's say, tells you how to cut. Would you have cut at this point, or would you have cut somewhere else? […] But that's probably where the experience that the surgeon has to bring plays a role again' (SI-7). Sufficient experience ensures that ML-CDSS recommendations could be critically scrutinised and evaluated: 'I think I would still need to have quite a lot of experience myself and know I'm about as good as this system. So, I think, if I start now as a physician, probably, I would think all the time like: "Yeah, who's deciding?"' (SI-9). Consequently, the use of ML-CDSS is permissible only if the professionals could largely perform the decisions even without it: 'So I think, I personally, in my idea of a good education it is, I think, eminently important to gain experience and just to be able to do theoretically what the system supports even without the system' (SI-13).

Respondents conclude that standardised use of the ML-CDSS presented jeopardises both the acquisition and the maintenance of a required level of competence. Instead, there is the danger of a potential 'lack of experience' (SI-13) if the system does not function properly: 'Well, if, I don't know, it can still be that something, that it, that the system fails or something similar and then you stand in the operating room and think to yourself: "Yeah, great. Now I don't have the support." It can be anything,

it's still technology that can fail and then if a surgeon is not trained to do it without that system, of course it's difficult' (SI-3). A nursing trainee said it similarly: 'If people only work with the device, they then get so used to it that they can no longer work stand-alone. That they suddenly stand there and no longer know what to do' (TI-4).

DISCUSSION

Respondents used 'responsibility' to describe theoretically distinctive objects, such as the positive assumption of individual moral accountability (including culpability) or legal liability. Nevertheless, when taking a relational perspective on responsibility,[13 23 24] the need of identifiable bearers of responsibility (subjects of responsibility) for clinical decision-making and its results (objects of responsibility) is consistently emphasised. Regarding clinical decision-making, respondents see it as their duty to justify—or at least to be able to justify—their clinical advice both to legal or institutional authorities and to their patients as moral authorities (addressees of responsibility). The normative standard was rarely made explicit. When this was done, interviewees mostly referred to presumptions about legal standards and to a moral obligation of justification towards patients and their autonomy.

The object of responsibility here includes even the delegation of decisions which are considered part of the physician's or nurse's role. The few studies already available indicate that physicians are willing to assign certain clinical activities to ML-CDSSs, while other tasks are considered 'as being central to who they are as physicians and as human beings'.[30] They emphasise that one of the high-valued core parts of their clinical role is a perceived 'final responsibility,' which means that they 'should always have a supervising role and, at least, every important decision should be made by (them)'.[30] To delegate responsibility within a cooperation—that is, 'share responsibility'[29]—then means handing over the supervising role for a definable task to a third party. Of course, concrete consideration should be given to what extent the decision-making is actually delegated.[19] However, there are preconditions for an assumption or assignment of responsibility.[23] For the results presented here, they can be divided into at least two requirements which are considered to be at risk when using ML-CDSSs.

First, the *structural freedom* or *institutional voluntariness*—as absence of institutional coercions—which would allow the use of ML-CDSS generally or the following of single ML-CDSS recommendations. In this sense, directives by institutional managements or superiors are seen as potential constraints to bearing responsibility. Additionally, more sublime types of coercion are seen as detrimental to responsibility, such as being pressured to use ML-CDSS in the face of human resource constraints or monetary profitability.

Second, the *epistemic freedom* or *level of information*, in which it is presupposed to have the necessary—mainly cognitive—competencies to make ML-CDSS advice seriously useful for the professional's decision-making. To make this possible, sufficient medical and technical background knowledge, practical experience in clinical decision-making without an ML-CDSS[14] and a comprehensible presentation of its outcomes are considered to be prerequisites. The alleged coping strategies if respondents do not (or no longer) see themselves in the epistemic position to reliably synthesise the ML-CDSS recommendation with their own judgement, that is, do not have the decision-making authority, are interesting. Professionals would seek advice from either colleagues or more experienced supervisors to ascertain

'shareable reasons' for weighing the ML-CDSS recommendation (as Kempt and Nagel suggested[29]). In this case, reasons for or against the final advice to the patient are deliberated between clinicians.[7 18 29] Alternatively, they would pass on the decision-making authority regarding the preferable advice to the patient (cf. the 'irresponsible outsourcing of responsibility' by Di Nucci[19]).

If professionals are structurally and epistemically able to act differently, they see themselves as responsible for their decision-making. Once one or both dimensions are restricted, the bearing of responsibility is assessed to be gradually reduced or even impossible.

The study results highlight the need of further balancing responsibility with other normative claims as well as the importance of preconditions. Insofar as the professionals continue to see themselves as integrated in a process of responsibility, this can lead to greater acceptance of normative regulations to meaningfully enable professionals to embrace responsibility as a bottom-up strategy with regard to shared responsibility. The study results are ambivalent insofar as the withdrawal or rejection of responsibility due to epistemic and/or structural limitations can also be observed. With regard to a normative solution, caution is required, not least in the interweaving with potential affected parties and their needs. Further investigations concerning the concrete epistemic and structural challenges are needed.

As a limitation of this study's results, it must be kept in mind that interviewees had little clinical experience. We also assume that especially those who were relatively more interested in questions of digitalisation of healthcare agreed to participate (selection bias) which might have had an influence on the answering behaviour.

CONCLUSIONS

Bearing responsibility for clinical decisions is linked to several requirements that were brought to the fore by our interview study. Particularly, structural opportunities for or against the use of ML-CDSS as well as a sufficient level of competency and clinical experience to meaningfully scrutinise ML-CDSS recommendations—that is, to have decision-making authority—were highlighted as necessary requirements. Even if the use of ML-CDSS may lead to shifts in roles and responsibilities,[6 35] legal and moral 'responsibility gaps'[12 36–38] should be prevented. As long as clinical professionals are assigned with the responsibility of being the final decision-maker, respectively, supervisor of ML-CDSS, they should also be given sufficient opportunities and qualifications to fulfil this responsibility. The ML-CDSS potentially offer numerous prospects to improve healthcare. However, the empirical findings illustrate that using ML-CDSS will require a consistent and transparent allocation of responsibility, not only for reasons of acceptance but also for the benefit of moral embedding.

for this study. She had no involvement in developing the interview guide, conducting the interviews or analyzing the interviews.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and this study was approved by the Research Ethics Committee of Hannover Medical School, Germany (Reg. No. 9805_BO_K_2021). All participants provided written informed consent to participate in this study. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. The data are not publicly available as they might contain information that could compromise research participant privacy and consent.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Florian Funer http://orcid.org/0000-0001-9242-0827
Wenke Liedtke http://orcid.org/0000-0002-1433-658X
Diana Schneider http://orcid.org/0000-0002-7131-5005
Sabine Salloch http://orcid.org/0000-0002-2987-2684

REFERENCES

1 Middleton B, Sittig DF, Wright A. Clinical decision support: a 25 year retrospective and a 25 year vision. *Yearb Med Inform* 2016;Suppl 1(Suppl 1):S103–16.
2 Nebeker C, Torous J, Bartlett Ellis RJ. Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Med* 2019;17:137.
3 Čartolovni A, Tomičić A, Lazić Mosler E. Ethical, legal, and social considerations of ai-based medical decision-support tools: a scoping review. *Int J Med Inform* 2022;161:104738.
4 Morley J, Machado CCV, Burr C, *et al.* The ethics of AI in health care: a mapping review. *Soc Sci Med* 2020;260:113172.
5 Murphy K, Di Ruggiero E, Upshur R, *et al.* Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Med Ethics* 2021;22:14.
6 Bleher H, Braun M. Diffused responsibility: attributions of responsibility in the use of AI-driven clinical decision support systems. *AI Ethics* 2022;2:747–61.
7 Coeckelbergh M. Artificial intelligence, responsibility Attribution, and a relational Justification of explainability. *Sci Eng Ethics* 2020;26:2051–68.
8 Verdicchio M, Perin A. When doctors and AI interact: on human responsibility for artificial risks. *Philos Technol* 2022;35:11.
9 Sand M, Durán JM, Jongsma KR. Responsibility beyond design: physicians' requirements for ethical medical AI. *Bioethics* 2022;36:162–9.
10 Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. *J Med Ethics* 2020;46:205–11.
11 Tigard DW. Artificial moral responsibility: how we can and can not hold machines responsible. *Camb Q Healthc Ethics* 2021;30:435–47.
12 Grote T, Nucci ED. Algorithmic decision-making and the problem of control. In: Beck B, Kühler M, eds. *Technology, Anthropology, and Dimensions of Responsibility*. Metzler: Stuttgart, 2020: 97–113.
13 Liedtke W. Verantwortung als Herausforderung – Entwicklungs- und Nutzungsbeteiligte bei Decision Support Systemen. In: Friedrich OS, Schleidgen S, eds. *Mensch-Maschine-Interaktion. Konzeptionelle, soziale und ethische Implikationen neuer Mensch-Technik-Verhältnisse*. Brill; Mentis: Paderborn, 2023: 332–60.
14 Gaube S, Suresh H, Raue M, *et al.* Do as AI say: susceptibility in deployment of clinical decision-aids. *NPJ Digit Med* 2021;4:31.
15 Bjerring JC, Busch J. Artificial intelligence and patient-centered decision-making. *Philos Technol* 2021;34:349–71.
16 Amann J, Blasimme A, Vayena E, *et al.* Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak* 2020;20:310.
17 Denekamp Y. Clinical decision support systems for addressing information needs of physicians. *Isr Med Assoc J* 2007;9:771–6.
18 Funer F. The deception of certainty: how non-interpretable machine learning outcomes challenge the epistemic authority of physicians. A deliberative-relational approach. *Med Health Care Philos* 2022;25:167–78.

19 Di Nucci E. Should we be afraid of medical AI? *J Med Ethics* 2019;45:556–8.
20 Lenk H. *Technikverantwortung: Güterabwägung, Risikobewertung, Verhaltenskodizes*. Frankfurt: Campus-Verlag, 1991: .
21 Ropohl G. Neue Wege, die Technik zu verantworten. In: Lenk H, Ropohl G, eds. *Technik und Ethik*. Stuttgart: Reclam, 1993: 149–76.
22 Grunwald A. Verantwortungsbegriff und Verantwortungsethik. In: Grunwald A, ed. *Rationale Technikfolgenbeurteilung. Konzeption und methodische Grundlagen*. Berlin; Heidelberg; New York: Springer, 1999: 75–194.
23 Langanke M, Liedtke W, Buyx A. Patients' responsibility for their health. In: Schramme T, Edwards S, eds. *Handbook of the Philosophy of Medicine*. Cham; Heidelberg; New York; Dordrecht; London: Springer, 2016: 619–40.
24 Loh J. Strukturen und Relata der Verantwortung. In: Heidbrink L, Langbehn C, Loh J, eds. *Handbuch Verantwortung*. Wiesbaden: Springer Fachmedien, 2017: 35–56.
25 Jonas H. *The Imperative of Responsibility: In Search of an Ethics for the Technological Age*. Chicago: University of Chicago Press, 1985.
26 Werner MH. Verantwortung. In: Grunwald A, Hillerbrand R, eds. *Handbuch Technikethik*. Stuttgart: J.B. Metzler, 2021: 44–8.
27 Braun M, Hummel P, Beck S, *et al*. Primer on an ethics of ai-based decision support systems in the clinic. *J Med Ethics* 2020;47.
28 Racine E, Boehlen W, Sample M. Healthcare uses of artificial intelligence: challenges and opportunities for growth. *Healthc Manage Forum* 2019;32:272–5.
29 Kempt H, Nagel SK. Responsibility, second opinions and peer-disagreement: ethical and epistemological challenges of using AI in clinical diagnostic contexts. *J Med Ethics* 2022;48:222–9.
30 Van Cauwenberge D, Van Biesen W, Decruyenaere J, *et al*. "Many roads lead to Rome and the artificial intelligence only shows me one road": an interview study on physician attitudes regarding the implementation of computerised clinical decision support systems. *BMC Med Ethics* 2022;23:50.
31 Oh S, Kim JH, Choi S-W, *et al*. Physician confidence in artificial intelligence: an online mobile survey. *J Med Internet Res* 2019;21.
32 Ford E, Edelman N, Somers L, *et al*. Barriers and facilitators to the adoption of electronic clinical decision support systems: a qualitative interview study with UK general practitioners. *BMC Med Inform Decis Mak* 2021;21:193.
33 Kuckartz U. *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. In: Grundlagentexte Methoden*. Beltz: Weinheim, 2016: .
34 Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19:349–57.
35 Jha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. *JAMA* 2016;316:2353–4.
36 Matthias A. The responsibility gap: ascribing responsibility for the actions of learning automata. *Ethics Inf Technol* 2004;6:175–83.
37 Santoni de Sio F, Mecacci G. Four responsibility gaps with artificial intelligence: why they matter and how to address them. *Philos Technol* 2021;34:1057–84.
38 Gunkel DJ. Mind the gap: responsible robotics and the problem of responsibility. *Ethics Inf Technol* 2020;22:307–20.