Meaningful Transparency for Clinicians: Operationalising HCXAI Research with Gynaecologists

Bianca G.S. Schor
Department of Computer Science and Technology,
University of Cambridge
United Kingdom

Emma Kallina
Institute for Manufacturing, University of Cambridge
United Kingdom

Jatinder Singh
University of Cambridge
United Kingdom
The Alan Turing Institute
United Kingdom

Alan F. Blackwell
Department of Computer Science and Technology,
University of Cambridge
United Kingdom

ABSTRACT

AI systems can bring great benefits to our healthcare systems, e.g. by improving patient outcomes. Yet implementing them into clinical practice remains challenging. To bridge the gap between academic research and design implementation, we argue clinicians need transparency about such systems that is *meaningful*—i.e. contextually appropriate—to them. Towards this, we explore recent HCXAI recommendations for building transparent AI systems for users in a specific domain: gynaecology. By better understanding clinicians' perspectives on meaningful transparency, our aim is to complement and help operationalise such recommendations. We conduct a co-design workshop and interviews with n=15 gynaecologists in the UK and the Netherlands. We show that HCXAI must better account for clinical teams with different types of gynaecologist users, and that the timeliness and relevance of the information provided about the AI-based tool throughout its design lifecycle—in particular before a tool is implemented into clinical practice—is crucial for transparency to become meaningful. Our contributions include: i) testing recommendations from the latest HCXAI literature with a prospective, real-life AI application in a relatively less-studied clinical domain; ii) describing and visualising gynaecologists' understanding of meaningful transparency for clinicians; iii) outlining four design recommendations towards realising meaningful transparency for clinicians and opportunities for research; and iv) expanding HCI and AI research in women's health by directly engaging with gynaecologists as users and co-designers. Exploring such issues is key to facilitate the implementation of AI systems that meet clinicians' information needs and that they can trust.

CCS CONCEPTS

• **Applied computing** → **Life and medical sciences**; • **Social and professional topics** → **User characteristics**; • **Human-centered computing** → **Empirical studies in HCI**; **HCI theory, concepts and models**.

KEYWORDS

clinicians, transparency, clinical Artificial Intelligence, explainable AI, gynaecologists, co-design

1 INTRODUCTION

Safe and ethical implementations of clinical AI can bring significant benefits to our healthcare systems , e.g. by improving patient outcomes or accelerating drug discovery [12, 79, 88]. AI research in healthcare has boomed in the last decade [7], yet real-life implementations are still facing obstacles [49, 53, 72]. *Algorithmic system transparency* can assist AI applications in such a high-risk contexts, e.g. by facilitating accountability and calibrating clinician's trust – i.e. helping them decide how much to trust the system [21, 26, 93, 103]. We define it here as the transfer of information from a developer or a system to clinicians about the system's design, training data, behaviour, and potential impact [47, 95].

To bridge the gap between academic research and clinical AI implementation, we argue that transparency must be *meaningful* to clinicians. Transparency is not a panacea, nor an end in itself [2, 6, 57, 71, 83]. To become meaningful, it must be *contextually appropriate*, i.e. the information transferred to clinicians must be relevant, accurate, proportionate, and comprehensible to them [26]. Whilst more research studies clinicians' perception of AI systems [81, 88, 100], what meaningful transparency means to them remains unanswered [15, 86, 104]. Exploring their understanding of meaningful transparency is important to facilitate the adoption of clinical AI systems and leverage their benefits in healthcare.

To answer this question, we build on recent literature in human-centered explainable AI (HCXAI), grounding it into the specific clinical context of gynaecological imagery. Using systematic reviews of the HCXAI literature, we extract design recommendations for transparent and explainable AI systems, and compare these to the clinicians' own concerns. By better understanding clinicians' perspective on what constitutes *meaningful* transparent AI systems for them, we aim to complement and help operationalise current HCXAI recommendations.

We focus on gynaecologists as users to test and ground these HCXAI design recommendations into a specific and relatively less-studied field of practice. As highlighted by Ehsan et al. [31, 33]: "understanding who interacts with the black-box of AI is just as important as 'opening' it, if not more." Clinicians in a variety of specialities have already been studied as users both of clinical AI systems, and of particular transparency interventions such as image saliency maps [20, 104]. In contrast, while gynaecologists frequently use medical imagery such as ultrasound scans (e.g. for diagnosis) [51], to date AI has had limited impact in the field [28]. Advances are likely in the near future, given current active research in AI for gynaecological imagery [18, 38, 46, 62, 63], and efforts are being made to integrate them into clinical practice [36]. It is therefore an appropriate point for the explainable AI (XAI), FAccT, and HCI communities to consider the transparency needs of gynaecologists who would interact with AI systems in their clinical practice (§2.2).

Specifically, we involve gynaecologists as co-designers to better describe meaningfully transparent AI systems for use by clinicians. As a real-life case study for co-design, we use a recent medRxiv preprint describing a deep learning model to assist in ovarian cancer diagnosis [18]. We carried out a co-design workshop (§2.3) with n=10 gynaecologists and interviews with n=6 gynaecologists in the UK and the Netherlands (§3). We show that HCXAI must better account for clinical teams with different types of gynaecologist users, and that the timeliness and relevance of the information provided about the AI-based tool is key for transparency to become meaningful for clinicians throughout the design lifecycle. Gynaecologist participants particularly ask for transparent information about the tool's conceptualisation, role, and purpose prior to its implementation in clincal practice. Our main contributions include: i) testing recommendations from the latest HCXAI literature, with a prospective, real-life AI application in a relatively less-studied clinical domain; ii) describing and visualising how gynaecologists understand meaningful transparency for clinicians; iii) outlining research and design recommendations towards meaningful transparency for clinicians; and iv) expanding HCI and AI research in women's health by directly engaging with gynaecologists as users and co-designers. Exploring such issues is key to supporting the implementation more transparent AI systems that meet clinicians' information needs.

2 BACKGROUND

We first describe some key publications on meaningful algorithmic system transparency (§2.1), before outlining the use of AI in gynaecological imagery (§2.2) and co-design studies with clinicians in the HCXAI literature (§2.3). The HCXAI design recommendations are described in detail in the next section (§3.1).

2.1 Meaningful algorithmic system transparency for clinicians: definitions, advantages and limitations

Algorithmic system transparency is crucial in high-risk domains like healthcare. It can support AI implementation, e.g. by helping clinicians effectively trust AI tools [21], but also help address risks for patients by supporting accountability, i.e. allowing scrutiny and challenge to individual decisions of a system [68, 94]. Indeed, Ehsan et al. [30] suggest social transparency—i.e. an explanation of AI-mediated decision-making that incorporates the socio-technical context—can help calibrate users' trust in algorithmic systems and improve decision-making. Nonetheless, defining transparency for clinicians remains an open challenge. For example, we observe the terms 'transparency' and 'explainability' are often used interchangeably in the academic literature [26]. Indeed, the research on transparency is growing and currently spans, among other domains, technical [16, 96], ethics [50, 56], legal [3, 29], and interdisciplinary work [6, 60]. Even in computer science, algorithmic system transparency is a broad concept with different interpretations [9, 60, 68].

Moreover, the need for transparency to be tailored to a specific stakeholder group is largely acknowledged [33, 68, 92]. However, because algorithmic system transparency is not a panacea [6, 71], transparency needs to become *meaningful* to them [70]. Cobbe et al. [26] and Norval et al. [69] describe meaningful transparency as information that is *contextually appropriate*, i.e. it must be a) relevant to the kinds of accountability needed, b) correct, complete, and representative, c) proportionate to the level of information each stakeholder's needs, and d) comprehensible to a given stakeholder [68]. Similarly, and in line with this, we argue that here transparency is *meaningful* to clinicians when the system or its designers provide, according to the clinician's needs, information on data, goals, outcomes, compliance, influence, usage and the algorithms employed [30, 56, 60, 95], and this information is contextually appropriate for clinicians. Note, however, that transparency can be detrimental in some circumstances [3, 57, 92]: among other risks, it can sometimes clash with values of ethical algorithmic systems, such as privacy, which is crucial to protect medical data [37, 61, 92]. Berendt [13] highlights the risk of mistaking transparency for a remedy to issues often associated with algorithmic systems, e.g. biases and discrimination. Moreover, Kizilecec [55] shows providing too much information to individuals (in our paper clinicians) can erode trust in algorithmic systems. Stohl et al. [83] call this the 'transparency paradox:' where high availability of information can produce opacity. It is thus important to understand what constitutes meaningful transparency for clinicians, so as to make sure we build AI systems that meet their information needs and that they can trust. We now ground this investigation in a specific clinical domain: gynaecological imagery.

2.2 AI in gynaecological imagery: an under-researched application domain in the XAI literature

Among all clinical domains, we argue it is urgent for the FAccT, XAI, and HCI communities to consider the transparency needs of gynaecologists. It is beyond the scope of this paper to review all

current and potential applications of AI in gynaecology, so we focus here on the use of AI in gynaecological imagery, e.g. ultrasound (US) and Computed tomography (CT) scans [28] (see also §5.2). Combining X-rays and tomographic reconstruction algorithms, CT scans epitomise algorithmic systems introduced into healthcare in the 1970s, mostly for diagnosis [27], and while US scans are common in gynaecology [51], yet a 2020 literature review concludes: "AI has had a little impact on this field so far" [28].

Despite the lack of substantial impact to date, AI research for gynaecological imagery has been growing in the last five years, e.g. on magnetic resonance imagery (MRI) [46, 66], CT [18], and US scans [38, 62, 63]. A 2023 literature review found 41 papers using AI on gynaecological US scans, with some new sub-fields emerging from 2019 [51]. There has also been some research towards integrating such systems into clinical practice [7, 36]. Despite this recent interest, and the rise of XAI research in healthcare more generally [8, 67], few XAI publications focus on gynaecological imagery as an application domain: a 2023 systematic literature review mentions only one paper in the field, which uses convolutional neural networks to achieve an explainable diagnosis method of cervical cancer using pathological images [67, 87]. Gynaecologists' transparency needs therefore deserve further attention. We aim to contribute towards better understanding these with a case study and co-design approach. The next section considers recent co-design and HCXAI research involving clinicians from other fields to guide our co-design experiment with gynaecologists.

2.3 HCXAI literature and co-design with clinicians

Human-centered AI (HCAI) has increasingly been studied by HCI and AI researchers in recent years [1, 25, 75]. While Yang et al. highlight Human-AI challenges [101], Ehsan et al. [32, 34] call for more emphasis on operationalising the human-centered perspectives in XAI at the conceptual, methodological, and technical levels. They also emphasise that the question of who is transparent and explainable AI for is key [31]. We build on this approach in this paper. The question of who AI is made for has been the focus of a growing number of HCI research that explore clinicians' perception of AI systems [24, 81, 98, 100]. For example, Bussone et al. [21] explore the role of explanations on misplaced trust and over-reliance in Clinical Decision Support Systems (CDSS). More recently, Verma et al. [88] interviewed medical-imaging experts to "scrutinize physicians' engagement with AI (...) and disentangle its future alignment across the clinical and research workflows, diverging from the existing 'one-size-fits-all' paradigm within Human-Centered AI discourses." Whilst such publications inform our understanding of how to build clinical AI systems, defining transparency meaningful to clinicians remains an open question [15, 86, 104].

To further account for AI users' perspectives, co-design studies have been shown to be a viable approach [7, 65]. Including users as co-designers is considered as helpful for safe and ethical AI design and implementation [73]. Co-design can be defined as "the creativity of designers and people not trained in design working together in the design development process" [77], and can be used for all stages of the design process. Moreover, co-design methods are widely used in health [82]. Among other, co-design have been used to build

various (explainable) clinical AI systems [59, 90, 97], e.g. a machine learning–based predictive CDSSs to analyze health records [78], and a trustworthy deep learning-based skin lesion classifier [105]. These efforts include human-centered approaches: for instance, Panigutti et al. recently co-designed an explainable AI technique and user interface for CDSSs with medical professionals [73]. While others have focussed on defining and addressing the challenges hindering AI implementation in healthcare [7, 53], we build on these studies and the design recommendations offered in the latest HCXAI literature as a starting point to explore what constitutes *meaningful* transparency for clinicians, towards bridging the gap between research and design implementation. We now outline these recommendations and our co-design methodology.

3 METHODS: CO-DESIGN WORKSHOP AND INTERVIEWS WITH GYNAECOLOGISTS

To bridge the gap between research and the practicalities of implementating clinical AI systems, we explore clinicians' understanding of meaningful transparency by testing and expanding recent HCXAI recommendations. We ground our investigation by using the real-life, prospective case study of a Deep Learning model to help diagnose ovarian cancer (§3.2) with n=15 gynaecologists. We first detail the HCXAI literature reviews we based it on, before describing our case study and co-design activities, i.e. an in-person workshop (§3.3), and individual interviews (§3.4). This study took place in the UK and the Netherlands from July until August 2023 and received ethical approval from our institution.

3.1 The HCXAI literature reviews tested and extended in our co-design activities.

We start by analysing recent HCXAI literature and using it as a broad database of design recommendations currently available for building transparent clinical AI for users. By doing so, we aim to help operationalise such recommendations and pave the way for more responsible AI in healthcare. Because HCXAI research is wide and spans various disciplines, including social sciences [64], we base our co-design activities on recent systematic literature reviews relevant to our case study (§3.2) and research question. We do not claim to perform a meta-analysis of such literature, nor do we have the space to discuss each HCXAI literature survey in detail. Instead, we compiled a total of 28 design recommendations for users from the HCXAI literature in three steps. First, we identified 10 relevant literature reviews published in the last four years [4, 8, 40, 58, 67, 76, 80, 90, 91, 99]. Based on i) their recent publication date (2019-2022), and ii) them explicitly offering design recommendations to build transparent AI systems for users, we then selected among these three reviews [58, 90, 91]. In doing so, we excluded relevant reviews where design recommendations were not clearly provided, such as [8]. Lastly, we extracted all the recommendations from these three reviews, as reported in Tables 1 and 2, and tested them in our card sorting activity. Note we report rather than endorse what is presented in the literature.

Laato et al.'s "recommendations for end-users" [58] do not specifically focus on a clinical context, while Wang et al. (2019) [90] and Wang et al. (2023) [91] do. Morevoer, Laato et al. [58] cite both Wang et al. (2019) [90] and Xie et al. [98], and yet we identify respectively

Table 1: List of the 16 design recommendations for users provided in Laato et al. [58] (nb. 1-16)

Recommendation	Reasoning
1. Context is everything – There is no one-size-fits-all type of solution	What to explain is dependent on several factors including what kind of AI system or decision we are explaining, who are the target audience and do we want to optimise for trust, for understandability or do we wish to simply comply by legislation
2. Provide explanations on demand, not all the time	For certain decisions and in certain moments users' may be interested in seeing more information on AI system decisions. However, constant display of full XAI documentation can hurt the user experience
3. Personalise explanations	There are various kinds of people with different levels of understanding of AI systems and XAI needs. This could be taken into account when explaining the system
4. Consider visualising explanations	Users tend to anthropomorphise AI and may benefit from human-like explanations. Visualising explanations may help some users to accept the AI system and its decisions better
5. Acknowledge the existence of trade-offs	For example, optimising explanations for understandability can lead to less details, which can hurt end users' confidence in the explanation
6. Consider potential misconceptions	Users may end up forming or having formed misconceptions regarding the AI system. These may shape behaviour and interpretation of explanations in a certain way. Explanations that are able to reshape misconceptions in a constructive way of conceptual change are valuable
7. Link explanations to users' mental models	This makes the AI system easier to understand for end users, increasing transparency
8. Strengthen users' curiosity towards the system	To increase user satisfaction, provide interesting and even surprising elements to keep the users' curiosity at a high level
9. Ensure the visibility and discoverability of explanations	Make sure AI system end users find and become aware of explanations
10. Use metaphors to demystify demystify how AI systems work	Metaphors can be more useful in increasingend users' understanding of AI systems than precise but difficult technical language
11. Support users' own thinking	The AI system should provide counterfactuals and explanations so users can reflect on and test their own thinking and hypotheses
12. Provide access to source data	Users may want to request access to raw data to build their trust in the AI system
13. Provide users with generalised explanations rather than case-based explanations	Users may consider it quirky if the decision is explained to them with a particular event from the past. To increase user acceptance, refer to generalised past events instead
14. Consider what part of the AI system to explain	Depending on the situation, users may wish to know more about, for example: (1) inputs; (2) outputs; (3) application; (4) situation; (5) model; (6) certainty; and (7) control
15. Explain unfavourable decisions	Users are likely to demand explanations when they disagree with the system
16. Communicate the uncertainties involved in the system's decision making	If there is a mismatch between users' expectations of the AI system and its actual capabilities, it hinders users' acceptance and trust building in it. Users should understand the risks of the AI system's making errors

six and two design recommendations in these publications that have not been included in Laato's list. We thus include Wang et al. (2019)'s recommendations in Table 2. Although we identified two design recommendations from Xie et al. [98], we excluded them from our study for it is not a systematic literature review, and our aim is not to make an exhaustive list of design recommendations but rather provide an informative one. Moreover, please note the minor edits we used in our list of recommendations to test it with gynaecologists: i) we use British spelling throughout for consistency and removed explicit mentions of specific application domains, not to bias participants; ii) Wang et al. (2019)'s recommendations (nb. 17-22) have been shortened for clarity and usability, e.g. we removed

Table 2: List of the 12 design recommendations for users provided in Wang et al. [90] (nb. 17-22) and Wang et al. [91] (nb. 23-28)

Recommendation	Reasoning
17. Support hypothesis generation	The system should provide explanations to allow users to generate and test hypotheses to further narrow down potential causes
18. Support forward (data-driven) reasoning	Avoid confirmation bias and backward reasoning, where the user does not consider other hypotheses independently
19. Support coherent factors	Users may expect some features to be correlated or have some other relationship and would be confused if these features contradict their typical relationship. Such feature attributions should be aggregated together or have their interaction relationship visualised
20. Supporting access to source and situational data	When adopting a new AI, users may want to manually perform decision making with a few instances to build up their trust. Showing raw data or supplementary data about the aggregated together or have their interaction situation, even if not used directly by the AI, can help with this verification goal
21. Support Bayesian reasoning	It is important to show probabilities, namely,(1) prior probability to indicate the prevalence of classes in general, and (2) intermediate posterior probabilities, where after filtering on a set of salient features or factors to indicate the conditional prevalence of an outcome
22. Integrating multiple explanations	Users employ a diverse range of XAI facilities, to reason variedly. Therefore, more work is needed to integrating multiple explanations into single explanations
23. Provide additional information about AI	For example, the importance of providing a holistic, global view of the AI to users during the onboarding process, such as the system's capabilities, functionality and design objective
24. Align AI design with current workflow	Design AI in a way that can seamlessly fit into the local context and workflow. For example, making the AI "unremarkable"; that is, embedding the system into the point of decision-making and the infrastructure of the organisation in an unobtrusive way
25. Consider social, organizational, and environmental factors	For instance, it is suggested to build accountable relationships with the leadership, engage stakeholders early and often, rigorously define the problem in context, make the system an enterprise-level solution by involving other departments and units (e.g., security and legal departments), and create an efficient communication mechanism with end users
26. Respect professional autonomy	Allow users to freely operate their professional judgment and decision-making without any interference
27. Adopt a human-centered design approach	Create ongoing feedback loops with users and stakeholders. By doing so, researchers and system designers can continuously and iteratively collect user feedback to inform the design of AI
28. Other considerations	Providing in-depth training to users to flatten the the learning curve, designing AI as a multi-user system to better engage users in decision-making, expanding the application of AI tools to a variety of scenarios, and examining the data quality for potential bias and fairness in AI

examples and references; and iii) Wang et al. (2023)'s recommendations (nb. 23-28) were extracted from paragraphs of text by keeping the title (as reported in the "Recommendation" column) and one or two key sentences (the "Reasoning" column), while removing academic references and medical connotations, so as not to bias the participants. Consequently, not all these recommendations emerged from user studies in a clinical field, but we kept the full list to enable participants to consider how they would categorise them.

3.2 Case study: Deep Learning segmentation of ovarian cancer in CT scans

To ground our co-design workshop and interviews in gynaecologists' clinical practice, we use an open-source Deep learning model built on CT scans that segments multi-site lesions in ovarian cancer [18]. Although this model is not currently implemented in clinical pathways, we use it to illustrate the kind of models that are proposed and could be potentially introduced into gynaecological imagery going forward. We used slides with a CT scan example to introduce the model and co-design activities. Moreover, most gynaecologists are familiar with ovarian cancer and CT scans, making it an easily understandable example for our participants. We acknowledge cancer diagnosis is performed differently according to the country and healthcare system: in the UK for example, a team of specialised clinicians including gynaecologist oncologists, oncologists, and radiologists, is involved in discussing the CT scans.

According to the country, we thus focus on the perspective of gynae-cologists' and gynaecologist oncologists experienced in operating ovarian cancer patients. We now describe our co-design activities based on this case study in more detail.

3.3 In-person co-design workshop with gynaecologists in Amsterdam UMC, the Netherlands

We conducted a co-design workshop with gynaecologists to bet-ter understand gynaecologists' perspective about meaningful AI transparency for clinicians towards operationalising the design recommendations listed above. We recruited n=10 gynaecologists experienced with ovarian cancer patients in Amsterdam UMC, the Netherlands via email invitation; all knew each other, and the work-shop was attended by a PhD student in XAI and facilitated by two PhD students in HCI. The workshop lasted ~1.5hrs and after being introduced to the concept of meaningful transparency and the case study orally with slides, the participants were teamed into groups of two to three, and completed a card sorting activity in about 20 minutes. The aim of the card sorting activity was to evaluate to what extent the design recommendations currently offered in the HCXAI literature encapsulates the concerns of clinicians, and ex-plore the sorting and category labels participants identify based on these recommendations. It included 28 cards directly copying the design recommendations from the latest HCXAI literature reviews described in §3.1. Each card was printed on a A4 paper sheet with a title and a short description on the reverse. Each team sorted the cards into categories of their choice, which they named without prompts. These were discussed, challenged, and sometimes merged in the second activity.

The second activity consisted of co-designing a 'transparency map:' we copied all the categories identified by participants in the card sorting activity onto large sticky notes and co-designed a con-ceptual diagram with the whole group, which we called a map for simplicity. We placed the sticky notes on a wall so that physical proximity between the notes represented the conceptual closeness of the categories, e.g. superimposing identical categories. Partic-ipants collectively decided where to place the categories on the map. When consensus did not emerge quickly, we duplicated the category to reconcile disagreements on their location. Once all the categories where placed on the wall, the participants were given the opportunity to add, remove categories, and edit the map. The aim of the transparency map is to i) challenge, edit, and validate the main categories identified by participants in the card sorting activity (based on HCXAI design recommendations for building transparent AI systems for users); ii) highlight the connections they saw between these categories as indicative of their understanding of meaningful transparency; and iii) visually represent gynaecologists' understanding of meaningful transparency to extract insights that might lead towards practical design implementation. This activity was audio recorded and transcribed by one author. Finally, the work-shop participants filled a short evaluation form focusing on their lived experiences of the co-design activities. We also asked about their experience with clinical AI, attitude towards AI in general, as well as demographic information.

3.4 In-person interviews with gynaecologists in the UK and the Netherlands

To test and confirm the data collected during the workshop, we conducted n=6 in-person, individual interviews with gynaecolo-gists. We recruited participants via email invitation in hospitals in the Netherlands and the UK. As a result, only one participant took part in the workshop just before their interview in Amsterdam UMC with two PhD students in HCI. All remaining interviews were conducted in the UK by one PhD student in HCI with gynaecol-ogists who had not attended the workshop. We recruited mostly (n=4) gynaecologist oncologists to reflect the sub-specialisation of clinical practice in the UK and the specific transparency needs of this user group, though all n=6 participants had experience with ovarian cancer patients. These took place in the Nuffield Health Cambridge Hospital and Addenbrooke's Hospital in Cambridge, Queen Elizabeth Hospital in London, and Norfolk and Norwich University Hospital in Norwich, UK.

The interviews lasted 45 minutes and consisted of one short training on decision-making and AI with slides and two co-design activities. After being introduced to the definition of meaningful transparency and the ovarian cancer case study (for the five in-terviews in the UK only), we discussed the theory of *system 1*, or intuitive, fast thinking process, and *system 2*, or analytical, slow thinking process [35, 52] (for all n=6 interviews). 'System thinking' theory is largely considered relevant in user research with clini-cians [84, 90]. It was thus chosen to set context as an initial prompt to help participants create 1-2 gynaecologist *personas:* characters who represent different user types that they had encountered in their professional environments. Participants were asked to briefly describe the persona's usual thinking style during the diagnosis process, whilst thinking about our case study, and could describe them referring to system 1 and 2 if helpful. We printed forms titled 'Person description' and the interviewer(s) created 1-2 personas as well, after reproducing the workshop's transparency map by sticking on a wall large sticky notes. This persona activity helped investigate decision-making processes used by gynaecologists, and generate different clinical AI users based on participants' real-world experience in healthcare settings.

After discussing their personas together, the participant and interviewer(s) critiqued and co-edited the map in relation to the personas identified; the 'Person description' forms were placed on the map above the category (i.e. the group of HCXAI design recom-mendations) that seemed most relevant for this persona according to its creator. This activity allowed to confirm previous results and elicited further discussion. Finally, the interviews concluded with a debrief and the same evaluation form as the workshop's. The interviews were audio recorded and transcribed by one author.

3.5 Gynaecologists' demographics

Table 3 outlines our total n=15 participants' demographic infor-mation. Three participants (P13, P14 and P15) took part in the workshop (and were included in our analysis) but did not provide demographic information. The participants comprised six consul-tant gynaecologists, four consultant gynaecologist oncologists, two resident gynaecologists, and three other gynaecologists (P13-15), all 15 having surgical experience. Our sample takes into account

Table 3: Participants' self-described roles, gender, years of experience, co-design participation, and country. Participants are referred as P[number] in the following sections.

ID	Self-Described Role	Gender	Years of experience	Participation	Country
P1	consultant gynaecologist	Woman	11	Workshop	Netherlands
P2	consultant gynaecologist	Woman	Not given	Workshop	Netherlands
P3	consultant gynaecologist	Woman	20	Workshop	Netherlands
P4	resident gynaecologist	Man	5	Workshop	Netherlands
P5	consultant gynaecologist	Woman	Not given	Workshop	Netherlands
P6	resident gynaecologist	Woman	Not given	Workshop	Netherlands
P7	consultant gynaecologist	Man	>15	Both	Netherlands
P8	consultant gynaecologist oncologist	Man	13	Interview	UK
P9	consultant gynaecologist	Man	17	Interview	UK
P10	consultant gynaecologist oncologist	Man	40	Interview	UK
P11	consultant gynaecologist oncologist	Man	15	Interview	UK
P12	consultant gynaecologist oncologist	Man	8	Interview	UK
P13	gynaecologist	Not given	Not given	Workshop	Netherlands
P14	gynaecologist	Not given	Not given	Workshop	Netherlands
P15	gynaecologist	Not given	Not given	Workshop	Netherlands

the sub-specialisation and targets in place in the two countries, e.g. NHS patients with suspicion of ovarian cancer should be referred to an gynaecologist oncologist within two weeks. Most participants declared having a strong interest but limited experience and understanding of AI systems. Most state they use no AI system in their work – though one specific clinical AI system called IOTA is mentioned in Amsterdam. Finally, most declare ignoring how AI systems work, or describe them as: "magic," and "no human errors."

3.6 Data analysis

Our co-design activities produced qualitative and quantitative data. The card sorting activity was analysed by counting the 'transparency recommendation' cards, categories, and comparing the labels and groupings made by the four teams. We compared the outputs to the three papers [58, 90, 91] the cards were copied from. The transparency maps produced during the workshop and interviewees were compared to one another and in relation to the card sorting results and HCXAI literature reviews. We used Clarke and Braun's thematic analysis method [17] to extract key themes from the audio recordings and transcripts. We conducted all six stages (i.e. familiarizing yourself with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, producing the report) over two iterations. The main author searched for semantic themes in the transcripts and audio recordings using an inductive, data-driven approach. Together with another co-author, we discussed, and validated the themes, before repeating the process. We kept participants' wording, as reported in Table 4. Finally, the demographic data collected, the information about participants' experience with AI, and their experience of the co-design activities helped contextualise our findings.

4 RESULTS

We ran a co-design workshop and interviews with a total of n=15 gynaecologists towards operationalising current HCXAI recommendations in healthcare. We now describe the themes they raise and

key insights into their understanding of meaningful transparency for clinicians in the context of our ovarian cancer case study.

4.1 Main themes raised by participants

Table 4 summarises the main themes raised by participants during the workshop and interviews. We group these into seven categories, named with participants' wording: 'users,' 'relevance,' 'timeline,' 'cyclic process,' 'user interface,' 'challenges,' and 'questions on transparency.' There was a consensus among workshop participants and interviewees regarding the importance of the first three axes, which therefore became the focus of our study. These three transparency categories are visualised in Fig. 1, showing (i) the distinction between the information needs of different user groups among gynaecologists (Users axis), (ii) the relevance of various types of information (Relevance axis), and (iii) the chronological order of the information users receive (Timeline axis).

4.2 Axis one (Users): differentiating the information needs of different gynaecologist user groups

Participants stressed the importance of differentiating the information needs of various types of gynaecologist users among clinicians likely to use an AI-based tool to help diagnose ovarian cancer. Differences in medical expertise, clinical experience, decision-making processes, and habit of using a given AI tool were highlighted. We argue that a level of granularity in analysing and considering users is needed to add nuance to the design recommendations described in Tables 1 and 2 and to help operationalise them. For example, interviewees raised these distinctions when creating gynaecologist personas: out of the 18 identified, 8 were described as using mainly system 1, i.e. a fast and intuitive thinking process, 7 as primarily relying on system 2, i.e. a slow, analytical thinking process, and 3 as using a combination of both [35, 52]. An interviewee notes: "I think a chronological thinking process is very apparent in surgeons (...) because [surgery] is a procedure (...) I think if you do any intervention, you will develop a [thinking] structure in steps"

Table 4: Summary of main themes raised by participants, the three main categories for the transparency map are highlighted.

Overarching Category	Themes
Users (Axis 1)	Experienced & beginner [tool] users, specialist & generalist gynaecologists, trainees & senior doctors, fast- & slow-thinkers, personality traits & cultures, expectations & medical norms, decision-making & communication styles, approaches to surgery, multidisciplinary team (MDT)
Relevance (Axis 2)	More/less crucial, "the why [using AI]" more important than "the how [AI works]"
Timeline (Axis 3)	Before/during/after implementation, integration into clinical practice, during use, pre-op planning & during surgery
Cyclic process	Evaluation (e.g risks for patients, patient outcomes, trials), feedback loop (e.g. updating the data), collaboration with clinicians, cancer-specific diagnostic process
User interface	Communication with users, clinician-friendly & intuitive, personalisation & levels of information, risks (i.e. biased data, cumbersome information), a key to look into the system, interactive tool
Challenges	Unusable systems, trust in system/model, trustworthy colleagues, tech skepticism, national & regional IT disparities
Questions on transparency	difference between transparency & explanation, relationship between transparency & usability

Figure 1: Diagram representing the transparency map co-designed with gynaecologists in the workshop in Amsterdam.
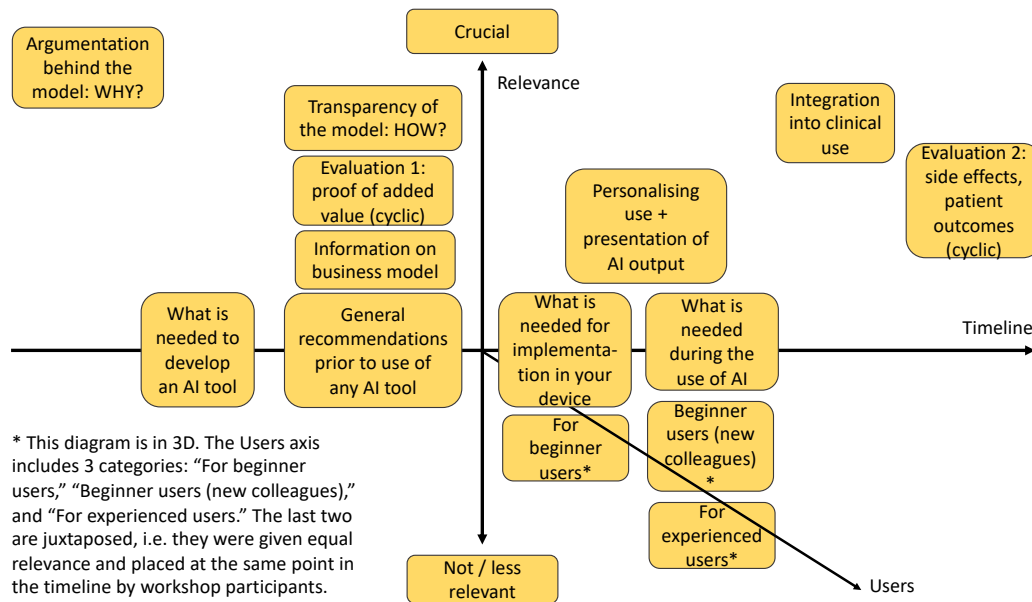
P11. These were often paired with personality traits, e.g. quiet and thorough, or social and dogmatic. By placing their personas on the transparency map, participants revealed how different gynaecologists can prioritise different information needs, in their experience (see §4.4). Another interviewee commented: "we [gynaecologist oncologists] don't need the same level of information as generalist [gynaecologists]. To me, [information about] the risk [for a patient] is more crucial" P8. Similarly, workshop participants drew a clear distinction between "beginner users" (including "new colleagues") and "experienced users" in terms of interaction with an AI tool (Fig. 1), and all interviewees confirmed it. These different users often work closely in teams, as one interviewee summarises it: "each individual [in an Multidisciplinary team (MDT)] has a specific set of expertise and, jointly, it's like the pieces of a puzzle coming together" P12. The HCXAI recommendations from prior work (Tables 1, 2) do not distinguish these various types of users, nor reflect their collaborative interactions with technology. Connecting them to

specific types of clinicians can help operationalising them [88]. By providing leads to define such users, we thus contribute towards facilitating the implementation of more clinical AI systems.

4.3 Axis two (Relevance): providing information in a prioritised manner in terms of perceived relevance for clinicians

As visualised in the second axis (Fig. 1), participants expressed the need to receive information about a clinical AI tool in a prioritised manner in terms of perceived relevance for them. Importantly, their priorities seem to differ from some of the recommendations from recent HCXAI literature. For example, interviewees placed virtually all the personas on the categories higher up on the map, which confirms their perceived importance for gynaecologist users. Indeed, the three main axes of the map reveal that the timeliness and relevance of the information provided about an AI-based tool throughout the AI tool's design lifecycle is key in participants'

understanding of meaningful transparency for clinicians. This prioritised and user-centric logic has not been described nor visualised for this context in the HCXAI literature yet. This insight is key to make sure we design AI systems that are meaningfully transparent for clinicians going forward. Moreover, when workshop participants grouped the 28 design recommendations extracted from the HCXAI reviews (Tables 1 and 2) into the category of their choice (unprompted), some recommendations were grouped as "not that relevant." Two teams (50% of workshop participants) grouped 11 recommendations (i.e. over a third of the recommendations) as "not/less relevant" and four of these recommendations were mentioned by both teams as being of lower relevancy: "Personalise explanations" (nb. 3), "Use metaphors to demystify how AI systems work" (nb.10), "Provide users with generalised explanations rather than case-based explanations" (nb. 13), and "Integrating multiple explanations" (nb. 22). While the recommendations nb. 3, 10, and 13 are not specifically listed for clinical users in the HCXAI literature reviews (Tables 1, 2), "Integrating multiple explanations" is [90]. Participants therefore challenge these recommendations for clinicians. Finally, the transparency map reveals another gap between the HCXAI literature and gynaecologists' perspective on meaningful transparency. Workshop participants seem relatively less concerned by the "what to explain" and "how to explain" categories mentioned in Laato et al. [58] than the "why," and "integration into clinical practice" categories they identify during the card sorting activity, as well as the additional "evaluation" categories they introduce during the map co-design activity (§4.4). Interviewees equally emphasise these categories, but none of the three selected literature reviews highlight them for clinicians [58, 90, 91]. This shows that, whilst recent HCXAI literature provides a starting point to discuss meaningful transparency for clinicians with gynaecologists, there are differences between some HCXAI recommendations and what gynaecologists consider as most relevant to them, and thus contextually appropriate transparency for clinicians.

4.4 Axis three (Timeline): prioritising information about the AI-based tool's conceptualisation, role, and purpose over (post-hoc) explanations

The third axis, called 'Timeline' in the transparency map, represents the stages in the AI design lifecycle when gynaecologist participants want to receive transparent information about the AI-based diagnostic tool (Fig. 1). Combined with the second axis (Relevance), it reveals that participants find information provided to them *before* the AI tool gets implemented more important for gynaecologists than (post-hoc) explanations, such as how and why it reaches a certain output with a given input. The former type of information includes among others insights about why the tool can be helpful to diagnose a given patient, what help they can expect from it, and why they should use it. Indeed, the category "Argumentation behind the model: WHY?" was placed by workshop participants before and higher than every other category on the transparency map, including above "Transparency of the model: HOW?" (Fig. 1). One workshop participant explained: "you need to have an idea why you're going to start this entire project" P3. Moreover, whilst all interviewees independently selected the "evaluation," and "how?"

categories as most relevant for personas described with system 2 (slow-thinking), and "integration into clinical practice" and "personalising use" for personas described with system 1 (fast-thinking), "the Why?" was once again considered by interviewees the most important category of transparent information when considering all types of personas (described with system 1, system 2 or both). Gynaecologists' preference for such type of transparent information might be partly explained by one interviewee's comment on the information needs of beginner and experienced users, who mainly use system 1 thinking: "I can imagine an interface that in the beginning gives me more information and in the end it ends up in a small question mark at the bottom. And if I use the model ten times, it's going to decrease the amount of background information that it's giving me, and it's more and more just showing me the results (...) one of my colleagues, they always want the background information. I don't, if I trust the system, I'm fine" P7. This confirms what Burgess et al. recently call "the front-loading" trust, whereby clinicians want "to determine their trust of an AI insight system when first introduced to the tool" [19]. However, the emphasis on holistic transparent information provided prior to the AI tool's implementation into clinical practice as a mechanism to support such trust was not in the design recommendations from the recent HCXAI literature reviews (Tables 1, 2). It warrants further research to be confirmed, and seems to indicate we need to rethink how and *when* to provide transparent information to such clinicians.

5 DISCUSSION

After testing the recommendations currently offered in the HCXAI literature (Tables 1, 2) and providing insights into gynaecologists' understanding of meaningful transparency (Table 4) in the context of our ovarian cancer case study (§3.2), we suggest four design recommendations for meaningful transparency for clinicians (§5.1). We then discuss the importance of direct engagement with gynaecologist users in relation to women's health and HCI and AI research (§5.2), before outlining our study limitations (§5.3) and research avenues to bridge the gap between HCXAI literature and implementation of transparent AI systems for clinicians (§5.4).

5.1 Design recommendations for meaningfully transparent AI systems for clinicians

Based on our findings, Table 5 outlines our four design recommendations for transparency meaningful to clinicians.

5.1.1 Prioritising information categories about the why, evaluation, personalisation, and integration into clinical practice. Given what participants expressed as their main information needs (Fig. 1), our first design recommendation moves beyond the list of 28 recommendations we have extracted from the HCXAI literature by prioritising four categories of transparent information: "the why," "evaluation," "personalisation," and "integration into clinical practice." Participants created the 'evaluation' category during the map co-design activity, so it is not directly connected to some pre-existing HCXAI recommendation listed in Tables 1 and 2. It includes, among others, information about the risks for patients, the patient outcomes, and the cost effectiveness of the tool (Table 4). This category is key to better align HCXAI design recommendations with real-life clinical practice, and is likely applicable to other clinical domains as well.

Table 5: Four design recommendations for AI systems meaningfully transparent for clinicians (in no particular order).

Design Recommendation	Reasoning related to our case study with gynaecologists
1. Prioritising information categories about the why, evaluation, personalisation, and integration into clinical practice	This focus can help cover the information needs of most gynaecologist personas we have identified with gynaecologist participants (§4.2, §4.3, §4.4)
2. Providing prioritised information in terms of timeliness and relevance for clinicians throughout the AI design lifecycle	This must align with gynaecologists' thinking process (Fig. 1, §4.3) and clinical experience as surgeons
3. Accounting for multidisciplinary teams (MDT) using a clinical AI system by taking into account the perspective, role, and human interactions of various types of gynaecologists	For example, in our case study, by better defining the role of gynaecologist oncologists in discussing with radiologists the diagnostic CT scans towards reaching a pre-operative plan (Table 4, §4.2)
4. Providing transparent information on the AI-based tool's conceptualisation, role, and purpose prior to its implementation	Information should be provided in particular before the tool gets implemented into clinical practice, e.g. during the initial clinical trials (Table 4, §4.4)

However, our gynaecologist participants did not mention the need to embody ethics and liability nor to develop AI educational opportunities, unlike Verma et al.'s oncologist interviewees [89]. This might reveal another specificity about gynaecologist users, who may be less exposed to clinical AI than those in other specialities.

5.1.2 Providing prioritised information in terms of timeliness and relevance for clinicians throughout the AI design lifecycle. Our second recommendation aims to align transparency with clinicians' and, in particular, surgeons' structured thinking process and clinical workflow (see §4.2). This confirms previous research on surgeons' procedural approach to tasks [39] but also AI onboarding [24]. Note our participants are surgeons, but not all gynaecologists are. We argue transparency can become more meaningful to such clinicians by following a structured order, key dimensions to which are visualised in Fig.1, though further research could assess whether this also applies to other types of gynaecologists (Table 4).

5.1.3 Accounting for multidisciplinary teams (MDT) using a clinical AI system by taking into account the perspective, role, and human interactions of various types of gynaecologists. Our third design recommendation specifies the HCXAI emphasis on accounting for the socio-technical context when designing and evaluating an AI-based system for users [26, 30, 33] in relation to a less-studied population and clinical domain—gynaecology (§2.2)—and uses a real-life case study to ground our findings into real-word settings [7]. Indeed, in the UK, gynaecologist oncologists approach ovarian cancer diagnosis and treatment in MDTs, and thus an AI-based diagnostic tool needs to support their interactions within such teams (see §4.2). By describing various gynaecologist users of a potential, real-life AI-based CDSS (§4.2), we follow Verma et al. in departing from the "one-size-fits-all" paradigm within HCAI research [88], and Berg in arguing that clinical IT development requires a user-centered approach, due to the complex network of people and practices in healthcare [14]. Further research is needed on characterising the transparency needs of gynaecologists as individual clinicians and within MDTs. Others have also stressed the need to better support patient-provider collaboration [19, 48].

5.1.4 Providing transparent information on the AI-based tool's conceptualisation, role, and purpose prior to its implementation. Our final recommendation stresses not only the need for designers to

communicate and engage with clinical users iteratively throughout the design process [14], but also the importance of providing them with transparent information about the AI-based tool's conceptualisation, role, and purpose *before* its implementation (§4.4). This could perhaps help to keep AI-based CDSS less obstructive and increase their contextual fit [102]. It also suggests that, in order to improve transparency for such clinicians, HCXAI research should shift from focusing predominantly on XAI to also exploring more inclusive approaches, and provide clinicians with information before an AI tool gets implemented (and, indeed, throughout the entire AI lifecycle, so as to enable understanding and scrutiny [26]).

5.2 Women's health in the HCI and AI literature: engaging directly with gynaecologist users

One key contribution of our study is to expand the AI and HCI literature on women's health by directly engaging with gynaecologist (professional) users. In doing so, we build upon the growing body of literature exploring women's bodily experiences and interactions with health and wellbeing technologies [5, 54], such as Homewood et al.'s innovative and phenomenological period tracking designs – e.g. Ovum [43, 44] and Ambient Cycle [45]– and Bardzell's feminist HCI research [10, 11, 41]. As we explicitly focus on gynaecologists' information needs in the context of a real-life case study of ovarian cancer, our study can thus complement such feminist HCI approach by contributing to implement more transparent (and thus responsible) clinical AI to advance women's health. Indeed, we aim to promote further HCXAI research relevant for improving women's health realities, where both the medical practitioners and patients' perspectives are critical for realising the potential of clinical AI. Moreover, and in line with the limitations of AI transparency (§2.1), we also acknowledge that clinical AI can bring some risks for women, e.g. discrimination. Thus more HCI and AI research is needed into "inaction as a design decision" [42] in relation to women's health, i.e. where designers decide not to implement AI because of its risks for users.

5.3 Study limitations

Due to the logistical challenges of organising in-person, co-design activities with gynaecologists, one limitation of this study is the

number of participants we were able to recruit. There is also a self-selecting bias in our sample of participants, based on their interest in AI systems. However, co-design can yield insights when conducted with a small group of participants, and we reached consensus on most of our activity findings. We also acknowledge that despite our best efforts to reproduce each experiment in the same conditions, this was not always possible. For example, the map co-designed during the workshop was reproduced in each interview on a different wall with different features. However, each interviewee had equal opportunities to discuss, challenge, and critique this map. Similarly, there was insufficient time with clinicians to discuss system 1 and 2 thinking in the workshop, but this was fully covered in the interviews. Moreover, we based the co-design activities, e.g. card sorting, on three of the latest HCXAI systematic reviews that explicitly included design recommendations for users [58, 90, 91]. Thus, we do not claim the list of 28 recommendations based on these surveys is exhaustive, e.g. we identified two additional ones in Xie et al. that were not tested here [98]. We argue this list is nonetheless indicative of the types of recommendations currently offered in the HCXAI literature to build transparent AI systems for users. Indeed, explainability and transparency are different concepts, but given the various use of both terms in the academic literature, we used the HCXAI literature reviews as broadly indicative of the current research done in algorithmic system transparency. In all, our findings towards operationalising such recommendations can be confirmed with more studies in the field, as described below.

5.4 Further research towards implementing meaningfully transparent AI systems for clinicians

Lastly, we encourage more co-design research into implementing meaningfully transparent AI systems for clinicians. Indeed, Thieme et al. show it can effectively lead to the integration of a production interface [85]. Patients must also be engaged as critical stakeholders of clinical AI. This was outside the scope of our study, however we encourage more studies directly engaging with patients of clinical AI, in particular women and non-binary individuals, to help close the gender data gap [22, 23, 74]. We suggest for example further studies to validate our design recommendations for meaningfully transparent clinical AI systems with a different gynaecological case study (§5.1). As our participants have unanimously highlighted the importance of the timeline when designing such clinical AI system, we also encourage future work to involve gynaecologists as co-designers as early as possible, and at every stage of the design process. Indeed, a 2021 systematic literature reviews shows that "clinical experts are less prevalent in developmental stages to verify clinical correctness, select model features, preprocess data, or serve as a gold standard"[78]. Building upon our findings, engaging clinicians as co-designers might be particularly relevant to design for the "why," "personalisation," "integration into clinical practice," and "evaluation" categories described above (§5.1). Moreover, we have described and visualised (Fig. 1) how some gynaecologists understand what constitutes meaningful transparency in clinical AI, in the context of ovarian cancer diagnosis with CT scans (§4.3). However, as our study has shown (§4.2), there is a need for future research to account for more detailed descriptions of clinical AI

primary users, for example by focusing on the perspective and information needs of clinicians who might not be as interested in AI systems as our participants. Finally, although we have explored how transparency could become meaningful to clinicians, investigating contexts when algorithmic system transparency might not be enough or could fail, for example to facilitate accountability or calibrate trust, was beyond the scope of this study. This question is equally important for safe and ethical AI application in healthcare, and requires further research with clinicians, including various types of gynaecologists.

6 CONCLUSION

To this day, implementing AI research into clinical practice remains challenging, and rather limited in gynaecology. To bridge the gap between the academic literature and design implementation of clinical AI systems, we investigate gynaecologists' understanding of *meaningful* transparency for clinicians. While transparency is not a panacea, it is recognised as helpful to calibrate clinicians' trust and facilitate accountability. Towards this, we have conducted a co-design workshop and interviews with n=15 gynaecologists in the UK and the Netherlands. Using the case study of a Deep Learning model for ovarian cancer, we have tested and extended recent HCXAI recommendations for building transparent AI systems for users, by grounding them into a specific and less-studied clinical domain: gynaecological imagery. In doing so, our aim is to complement and help operationalise such recommendations in clinical practice. Our study reveal that HCXAI must better account for clinical teams with different types of gynaecologist users. We also show that the timeliness and relevance of the information given users about the AI-based tool is key for transparency to become meaningful for clinicians throughout the design lifecycle, and in particular information about the tool's conceptualisation, role, and purpose provided prior to its implementation in clinical practice. Our main contributions include: i) testing recommendations from the latest HCXAI literature in a specific, relatively less-studied clinical domain with a prospective, real-life AI application; ii) describing and visualising gynaecologists' understanding of meaningful AI transparency for clinicians; iii) outlining research and four design recommendations towards meaningful transparency for clinicians; and iv) expanding HCI and AI research on women's health by directly engaging with gynaecologist as users and co-designers. Exploring such issues is key to implement more transparent AI systems that more effectively meet clinicians' information needs.

REFERENCES

[1] Ashraf Abdul, Jo Vermeulen, Danding Wang, Brian Y Lim, and Mohan Kankanhalli. 2018. Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–18.
[2] Alessandro Acquisti, Idris Adjerid, and Laura Brandimarte. 2013. Gone in 15 Seconds: The Limits of Privacy Transparency and Control. *IEEE Security Privacy* 11, 4 (2013), 72–74. https://doi.org/10.1109/MSP.2013.86
[3] Rachel Adams. 2020. *Transparency: New trajectories in law*. Routledge.
[4] Johannes Allgaier, Lena Mulansky, Rachel Lea Draelos, and Rüdiger Pryss. 2023. How does the model make predictions? A systematic literature review on the explainability power of machine learning in healthcare. *Artificial Intelligence in Medicine* 143 (2023), 102616.
[5] Teresa Almeida, Madeline Balaam, Shaowen Bardzell, and Lone Koefoed Hansen. 2020. Introduction to the Special Issue on HCI and the Body: Reimagining Women's Health. *ACM Trans. Comput.-Hum. Interact.* 27, 4, Article 20 (aug 2020), 32 pages. https://doi.org/10.1145/3406091

[6] Mike Ananny and Kate Crawford. 2018. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society* 20, 3 (2018), 973–989.

[7] Tariq Osman Andersen, Francisco Nunes, Lauren Wilcox, Enrico Coiera, and Yvonne Rogers. 2023. Introduction to the Special Issue on Human-Centred AI in Healthcare: Challenges Appearing in the Wild. , 11 pages.

[8] Anna Markella Antoniadi, Yuhan Du, Yasmine Guendouz, Lan Wei, Claudia Mazo, Brett A Becker, and Catherine Mooney. 2021. Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review. *Applied Sciences* 11, 11 (2021), 5088.

[9] Carolyn Ball. 2009. What is transparency? *Public Integrity* 11, 4 (2009), 293–308.

[10] Shaowen Bardzell. 2010. Feminist HCI: taking stock and outlining an agenda for design. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1301–1310.

[11] Shaowen Bardzell and Jeffrey Bardzell. 2011. Towards a feminist HCI methodology: social science, feminism, and HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 675–684.

[12] Ivana Bartoletti. 2019. AI in healthcare: Ethical and privacy challenges. In *Artificial Intelligence in Medicine: 17th Conference on Artificial Intelligence in Medicine, AIME 2019, Poznan, Poland, June 26–29, 2019, Proceedings 17*. Springer, 7–10.

[13] Bettina Berendt. 2022. The AI Act Proposal: Towards the next transparency fallacy? Why AI regulation should be based on principles based on how algorithmic discrimination works.

[14] Marc Berg. 1999. Patient care information systems and health care work: a sociotechnical approach. *International journal of medical informatics* 55, 2 (1999), 87–101.

[15] Astrid Bertrand, Tiphaine Viard, Rafik Belloum, James R Eagan, and Winston Maxwell. 2023. On Selective, Mutable and Dialogic XAI: a Review of What Users Say about Different Types of Interactive Explanations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–21.

[16] Umang Bhatt, Alice Xiang, Shubham Sharma, Adrian Weller, Ankur Taly, Yunhan Jia, Joydeep Ghosh, Ruchir Puri, José M. F. Moura, and Peter Eckersley. 2020. Explainable Machine Learning in Deployment. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) *(FAT* ' *20)*. Association for Computing Machinery, New York, NY, USA, 648?657. https://doi.org/10.1145/3351095.3375624

[17] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[18] Thomas Buddenkotte, Leonardo Rundo, Ramona Woitek, Lorena Escudero Sanchez, Lucian Beer, Mireia Crispin-Ortuzar, Christian Etmann, Subhadip Mukherjee, Vlad Bura, Cathal McCague, et al. 2023. Deep learning-based Segmentation of Multi-site Disease in Ovarian Cancer. *medRxiv* (2023), 2023–01.

[19] Eleanor R Burgess, Ivana Jankovic, Melissa Austin, Nancy Cai, Adela Kapuścińska, Suzanne Currie, J Marc Overhage, Erika S Poole, and Jofish Kaye. 2023. Healthcare AI Treatment Decision Support: Design Principles to Enhance Clinician Adoption and Trust. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[20] Alyxander Burns, Christiana Lee, Ria Chawla, Evan Peck, and Narges Mahyar. 2023. Who Do We Mean When We Talk About Visualization Novices?. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.

[21] Adrian Bussone, Simone Stumpf, and Dympna O'Sullivan. 2015. The Role of Explanations on Trust and Reliance in Clinical Decision Support Systems. In *2015 International Conference on Healthcare Informatics*. 160–169. https://doi.org/10.1109/ICHI.2015.26

[22] Mayra Buvinic, Rebecca Furst-Nichols, and Gayatri Koolwal. 2014. Mapping gender data gaps. *Data2X report, United Nations Foundation, Washington, DC. http://data2x. org/wp-content/up8 loads/2014/11/Data2X_MappingGenderDataGaps_FullReport. pdf* (2014).

[23] Mayra Buvinic and Ruth Levine. 2016. Closing the gender data gap. *Significance* 13, 2 (2016), 34–37.

[24] Carrie J Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2019. " Hello AI": uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making. *Proceedings of the ACM on Human-computer Interaction* 3, CSCW (2019), 1–24.

[25] Tara Capel and Margot Brereton. 2023. What is Human-Centered about Human-Centered AI? A Map of the Research Landscape. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 359, 23 pages. https://doi.org/10.1145/3544548.3580959

[26] Jennifer Cobbe, Michelle Seng Ah Lee, and Jatinder Singh. 2021. Reviewable Automated Decision-Making: A Framework for Accountable Algorithmic Systems. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) *(FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 598–609. https://doi.org/10.1145/3442188.3445921

[27] collective. [n. d.]. Wikipedia: CT scans. https://en.wikipedia.org/wiki/CT_scan

[28] L Drukker, JA Noble, and AT Papageorghiou. 2020. Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology. *Ultrasound in Obstetrics & Gynecology* 56, 4 (2020), 498–505.

[29] Lilian Edwards and Michael Veale. 2018. Enslaving the algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"? *IEEE Security & Privacy* 16, 3 (2018), 46–54.

[30] Upol Ehsan, Q. Vera Liao, Michael Muller, Mark O. Riedl, and Justin D. Weisz. 2021. *Expanding Explainability: Towards Social Transparency in AI Systems*. Association for Computing Machinery. https://doi.org/10.1145/3411764.3445188

[31] Upol Ehsan, Samir Passi, Q Vera Liao, Larry Chan, I Lee, Michael Muller, Mark O Riedl, et al. 2021. The who in explainable ai: How ai background shapes perceptions of ai explanations. *arXiv preprint arXiv:2107.13509* (2021).

[32] Upol Ehsan, Philipp Wintersberger, Q. Vera Liao, Martina Mara, Marc Streit, Sandra Wachter, Andreas Riener, and Mark O. Riedl. 2021. Operationalizing Human-Centered Perspectives in Explainable AI. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 94, 6 pages. https://doi.org/10.1145/3411763.3441342

[33] Upol Ehsan, Philipp Wintersberger, Q Vera Liao, Elizabeth Anne Watkins, Carina Manger, Hal Daumé III, Andreas Riener, and Mark O Riedl. 2022. Human-Centered Explainable AI (HCXAI): beyond opening the black-box of AI. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–7.

[34] Upol Ehsan, Philipp Wintersberger, Elizabeth A Watkins, Carina Manger, Gonzalo Ramos, Justin D. Weisz, Hal Daumé Iii, Andreas Riener, and Mark O Riedl. 2023. Human-Centered Explainable AI (HCXAI): Coming of Age. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 353, 7 pages. https://doi.org/10.1145/3544549.3573832

[35] Seymour Epstein. 1994. Integration of the cognitive and the psychodynamic unconscious. *American psychologist* 49, 8 (1994), 709.

[36] Lorena Escudero Sanchez, Thomas Buddenkotte, Mohammad Al Sa'd, Cathal McCague, James Darcy, Leonardo Rundo, Alex Samoshkin, Martin J Graves, Victoria Hollamby, Paul Browne, et al. 2023. Integrating Artificial Intelligence Tools in the Clinical Research Setting: The Ovarian Cancer Use Case. *Diagnostics* 13, 17 (2023), 2813.

[37] Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar. 2020. Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. *Berkman Klein Center Research Publication* 2020-1 (2020).

[38] Stefano Guerriero, MariaAngela Pascual, Silvia Ajossa, Manuela Neri, Eleonora Musa, Betlem Graupera, Ignacio Rodriguez, and Juan Luis Alcazar. 2021. Artificial intelligence (AI) in the detection of rectosigmoid deep endometriosis. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 261 (2021), 29–33.

[39] Atul Guwande. 2010. The checklist manifesto. *New York: Picadur* (2010).

[40] AKM Bahalul Haque, AKM Najmul Islam, and Patrick Mikalef. 2023. Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting and Social Change* 186 (2023), 122120.

[41] Sarah Homewood. 2018. Designing for the changing body: A feminist exploration of self-tracking technologies. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–4.

[42] Sarah Homewood. 2019. Inaction as a design decision: Reflections on not designing self-tracking tools for menopause. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[43] Sarah Homewood, Harvey Bewley, and Laurens Boer. 2019. Ovum: Designing for fertility tracking as a shared and domestic experience. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 553–565.

[44] Sarah Homewood, Laurens Boer, and Anna Vallgårda. 2020. Designers in white coats: deploying ovum, a fertility tracking device. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[45] Sarah Homewood and Anna Vallgårda. 2020. Putting phenomenological theories to work in the design of self-tracking technologies. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 1833–1846.

[46] Dingdu Hu, Junming Jian, Yongai Li, and Xin Gao. 2023. Deep learning-based segmentation of epithelial ovarian cancer on T2-weighted magnetic resonance images. *Quantitative Imaging in Medicine and Surgery* 13, 3 (2023), 1464.

[47] IEEE. 2022. 7001-2021. IEEE Standard for Transparency of Autonomous Systems. https://ieeexplore.ieee.org/document/9726144

[48] Maia Jacobs, Jeffrey He, Melanie F. Pradier, Barbara Lam, Andrew C Ahn, Thomas H McCoy, Roy H Perlis, Finale Doshi-Velez, and Krzysztof Z Gajos. 2021. Designing AI for trust and collaboration in time-constrained medical decisions: a sociotechnical lens. In *Proceedings of the 2021 chi conference on human factors in computing systems*. 1–14.

[49] Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, and Yongjun Wang. 2017. Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology* 2, 4 (2017).

[50] Anna Jobin, Marcello Ienca, and Effy Vayena. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1, 9 (2019), 389–399.

[51] Elena Jost, Philipp Kosian, Jorge Jimenez Cruz, Shadi Albarqouni, Ulrich Gembruch, Brigitte Strizek, and Florian Recker. 2023. Evolving the Era of 5D Ultrasound? A Systematic Literature Review on the Applications for Artificial Intelligence Ultrasound Imaging in Obstetrics and Gynecology. *Journal of Clinical Medicine* 12, 21 (2023), 6833.

[52] Daniel Kahneman. 2011. *Thinking, fast and slow.* macmillan.

[53] Christopher J Kelly, Alan Karthikesalingam, Mustafa Suleyman, Greg Corrado, and Dominic King. 2019. Key challenges for delivering clinical impact with artificial intelligence. *BMC medicine* 17 (2019), 1–9.

[54] Os Keyes, Burren Peil, Rua M. Williams, and Katta Spiel. 2020. Reimagining (Women's) Health: HCI, Gender and Essentialised Embodiment. *ACM Trans. Comput.-Hum. Interact.* 27, 4, Article 25 (aug 2020), 42 pages. https://doi.org/10.1145/3404218

[55] René F Kizilcec. 2016. How much information? Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI conference on human factors in computing systems.* 2390–2395.

[56] Ansgar Koene, Chris Clifton, Yohko Hatada, Helena Webb, and Rashida Richardson. 2019. A governance framework for algorithmic accountability and transparency. (2019).

[57] Joshua Alexander Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, and Harlan Yu. 2017. Accountable algorithms. *University of Pennsylvania Law Review* 165 (2017), 633–705. https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3/

[58] Samuli Laato, Miika Tiainen, AKM Najmul Islam, and Matti Mäntymäki. 2022. How to explain AI systems to end users: a systematic literature review and research agenda. *Internet Research* 32, 7 (2022), 1–31.

[59] Jean-Baptiste Lamy, Boomadevi Sekar, Gilles Guezennec, Jacques Bouaud, and Brigitte Séroussi. 2019. Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach. *Artificial intelligence in medicine* 94 (2019), 42–53.

[60] Stefan Larsson and Fredrik Heintz. 2020. Transparency in artificial intelligence. *Internet Policy Review* 9, 2 (2020).

[61] David Lehr and Paul Ohm. 2017. Playing with the Data: What Legal Scholars Should Learn About Machine Learning. *UC Davis Law Review* 51 (2017).

[62] Gabriel Maicas, Mathew Leonardi, Jodie Avery, Catrina Panuccio, Gustavo Carneiro, M Louise Hull, and George Condous. 2021. Deep learning to diagnose pouch of Douglas obliteration with ultrasound sliding sign. *Reproduction and Fertility* 2, 4 (2021), 236–243.

[63] Praful Mathur, Karan Kakwani, Srinivas Kudavelly, GA Ramaraju, et al. 2020. Deep learning based quantification of ovary and follicles using 3d transvaginal ultrasound in assisted reproduction. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC).* IEEE, 2109–2112.

[64] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence* 267 (2019), 1–38.

[65] Henrik Mucha, Dennis Mevißen, Sebastian Robert, Ricarda Jacobi, Kirsten Meyer, Winfried Heusler, and Daniel Arztmann. 2020. Co-Design Futures for AI and Space: A Workbook Sprint. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20).* Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3375203

[66] Najme Zehra Naqvi, Kritleen Kaur, Shubhi Khanna, and Shivangi Singh. 2023. An Overview of Machine Learning Techniques Focusing on the Diagnosis of Endometriosis. *Machine Vision and Augmented Intelligence: Select Proceedings of MAI 2022* (2023), 61–84.

[67] Sajid Nazir, Diane M Dickson, and Muhammad Usman Akram. 2023. Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks. *Computers in Biology and Medicine* (2023), 106668.

[68] Chris Norval, Jennifer Cobbe, and Jatinder Singh. 2021. Towards an accountable Internet of Things: A call for 'reviewability'. In *Privacy by Design for the Internet of Things: Building Accountability and Security.* The Institution of Engineering and Technology.

[69] Chris Norval, Kristin Cornelius, Jennifer Cobbe, and Jatinder Singh. 2022. Disclosure by Design: Designing information disclosures to support meaningful transparency and accountability. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency.* 679–690.

[70] Chris Norval and Jatinder Singh. 2023. A Room With an Overview: Towards Meaningful Transparency for the Consumer Internet of Things. *IEEE Internet of Things Journal* (2023).

[71] Jonathan A. Obar. 2020. Sunlight alone is not a disinfectant: Consent and the futility of opening Big Data black boxes (without assistance). *Big Data & Society* 7, 1 (2020), 1–5.

[72] Nazmun Nisat Ontika, Hussain Abid Syed, Sheree May Saßmannshausen, Richard HR Harper, Yunan Chen, Sun Young Park, Miria Grisot, Astrid Chow, Nils Blaumer, Aparecido Fabiano Pinatti de Carvalho, et al. 2022. Exploring human-centered AI in healthcare: diagnosis, explainability, and trust. (2022).

[73] Cecilia Panigutti, Andrea Beretta, Daniele Fadda, Fosca Giannotti, Dino Pedreschi, Alan Perotti, and Salvatore Rinzivillo. 2023. Co-design of human-centered, explainable AI for clinical decision support. *ACM Transactions on Interactive Intelligent Systems* (2023).

[74] Caroline Criado Perez. 2019. *Invisible women: Data bias in a world designed for men.* Abrams.

[75] Gonzalo Ramos, Jina Suh, Soroush Ghorashi, Christopher Meek, Richard Banks, Saleema Amershi, Rebecca Fiebrink, Alison Smith-Renner, and Gagan Bansal. 2019. Emerging Perspectives in Human-Centered Machine Learning. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19).* Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3290607.3299014

[76] Waddah Saeed and Christian Omlin. 2023. Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. *Knowledge-Based Systems* 263 (2023), 110273.

[77] Elizabeth B-N Sanders and Pieter Jan Stappers. 2008. Co-creation and the new landscapes of design. *Co-design* 4, 1 (2008), 5–18.

[78] Jessica M Schwartz, Amanda J Moy, Sarah C Rossetti, Noémie Elhadad, and Kenrick D Cato. 2021. Clinician involvement in research on machine learning–based predictive clinical decision support for the hospital setting: A scoping review. *Journal of the American Medical Informatics Association* 28, 3 (2021), 653–663.

[79] Rushabh Shah and Alina Chircu. 2018. IoT and AI in healthcare: A systematic literature review. *Issues in Information Systems* 19, 3 (2018).

[80] Ruey-Kai Sheu and Mayuresh Sunil Pardeshi. 2022. A Survey on Medical Explainable AI (XAI): Recent Progress, Explainability Approach, Human Interaction and Scoring System. *Sensors* 22, 20 (2022), 8068.

[81] Venkatesh Sivaraman, Leigh A Bukowski, Joel Levin, Jeremy M Kahn, and Adam Perer. 2023. Ignore, trust, or negotiate: understanding clinician acceptance of AI-based treatment recommendations in health care. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems.* 1–18.

[82] Peter Slattery, Alexander K Saeri, and Peter Bragge. 2020. Research co-design in health: a rapid overview of reviews. *Health research policy and systems* 18, 1 (2020), 1–13.

[83] Cynthia Stohl, Michael Stohl, and Paul M Leonardi. 2016. Managing Opacity: Information Visibility and the Paradox of Transparency in the Digital Age. *International Journal of Communication* 10 (2016).

[84] Shu Wen Tay, Paul Ryan, and C Anthony Ryan. 2016. Systems 1 and 2 thinking processes and cognitive reflection testing in medical students. *Canadian medical education journal* 7, 2 (2016), e97.

[85] Anja Thieme, Maryann Hanratty, Maria Lyons, Jorge Palacios, Rita Faia Marques, Cecily Morrison, and Gavin Doherty. 2023. Designing human-centered AI for mental health: Developing clinically relevant applications for online CBT treatment. *ACM Transactions on Computer-Human Interaction* 30, 2 (2023), 1–50.

[86] Sana Tonekaboni, Shalmali Joshi, Melissa D McCradden, and Anna Goldenberg. 2019. What clinicians want: contextualizing explainable machine learning for clinical end use. In *Machine learning for healthcare conference.* PMLR, 359–380.

[87] Kazuki Uehara, Masahiro Murakawa, Hirokazu Nosato, and Hidenori Sakanashi. 2020. Multi-scale explainable feature learning for pathological image analysis using convolutional neural networks. In *2020 IEEE International Conference on Image Processing (ICIP).* IEEE, 1931–1935.

[88] Himanshu Verma, Jakub Mlynar, Roger Schaer, Julien Reichenbach, Mario Jreige, John Prior, Florian Evéquoz, and Adrien Depeursinge. 2023. Rethinking the Role of AI with Physicians in Oncology: Revealing Perspectives from Clinical and Research Workflows. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23).* Association for Computing Machinery, New York, NY, USA, Article 17, 19 pages. https://doi.org/10.1145/3544548.3581506

[89] Himanshu Verma, Roger Schaer, Julien Reichenbach, Mario Jreige, John O Prior, Florian Evéquoz, and Adrien Depeursinge. 2021. On improving physicians' trust in AI: Qualitative inquiry with imaging experts in the oncological domain. *BMC Medical Imaging, in review* (2021).

[90] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y Lim. 2019. Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019 CHI conference on human factors in computing systems.* 1–15.

[91] Liuping Wang, Zhan Zhang, Dakuo Wang, Weidan Cao, Xiaomu Zhou, Ping Zhang, Jianxing Liu, Xiangmin Fan, and Feng Tian. 2023. Human-centered design and evaluation of AI-empowered clinical decision support systems: a systematic review. *Frontiers in Computer Science* 5 (2023), 1187299.

[92] Adrian Weller. 2019. Transparency: motivations and challenges. In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning.* Springer, 23–40.

[93] Maranke Wieringa. 2020. What to Account for When Accounting for Algorithms: A Systematic Literature Review on Algorithmic Accountability. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) *(FAT* '20).* Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3351095.3372833

[94] Maranke Wieringa. 2020. What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability. In *Proceedings of*

the 2020 Conference on Fairness, Accountability, and Transparency. 1–18.

[95] Alan FT Winfield, Serena Booth, Louise A Dennis, Takashi Egawa, Helen Hastie, Naomi Jacobs, Roderick I Muttram, Joanna I Olszewska, Fahimeh Rajabiyazdi, Andreas Theodorou, et al. 2021. IEEE P7001: a proposed standard on transparency. *Frontiers in Robotics and AI* (2021), 225.

[96] Ning Xie, Gabrielle Ras, Marcel van Gerven, and Derek Doran. 2020. Explainable deep learning: A field guide for the uninitiated. *arXiv preprint arXiv:2004.14545* (2020).

[97] Yao Xie, Melody Chen, David Kao, Ge Gao, and Xiang'Anthony' Chen. 2020. CheXplain: enabling physicians to explore and understand data-driven, AI-enabled medical imaging analysis. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[98] Yao Xie, Ge Gao, and Xiang 'Anthony' Chen. 2019. Outlining the Design Space of Explainable Intelligent Systems for Medical Diagnosis. *CoRR* abs/1902.06019 (2019). arXiv:1902.06019 http://arxiv.org/abs/1902.06019

[99] Qian Xu, Wenzhao Xie, Bolin Liao, Chao Hu, Lu Qin, Zhengzijin Yang, Huan Xiong, Yi Lyu, Yue Zhou, Aijing Luo, et al. 2023. Interpretability of Clinical Decision Support Systems Based on Artificial Intelligence from Technological and Medical Perspective: A Systematic Review. *Journal of Healthcare Engineering* 2023 (2023).

[100] Qian Yang, Yuexing Hao, Kexin Quan, Stephen Yang, Yiran Zhao, Volodymyr Kuleshov, and Fei Wang. 2023. Harnessing Biomedical Literature to Calibrate Clinicians' Trust in AI Decision Support Systems. In *Proceedings of the 2023*

CHI Conference on Human Factors in Computing Systems (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 14, 14 pages. https://doi.org/10.1145/3544548.3581393

[101] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–13.

[102] Qian Yang, Aaron Steinfeld, and John Zimmerman. 2019. Unremarkable AI: Fitting intelligent decision support into critical, clinical decision-making processes. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–11.

[103] Yunfeng Zhang, Q Vera Liao, and Rachel KE Bellamy. 2020. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 295–305.

[104] Yayan Zhao, Mingwei Li, and Matthew Berger. 2023. Graphical Perception of Saliency-based Model Explanations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.

[105] Roberto V Zicari, Sheraz Ahmed, Julia Amann, Stephan Alexander Braun, John Brodersen, Frédérick Bruneault, James Brusseau, Erik Campano, Megan Coffee, Andreas Dengel, et al. 2021. Co-design of a trustworthy AI system in healthcare: deep learning based skin lesion classifier. *Frontiers in Human Dynamics* 3 (2021), 688152.