Acceptability of artificial intelligence in breast screening: focus groups with the screening-eligible population in England

Lauren Gatting ,[1,2,3] Syeda Ahmed,[4] Priscilla Meccheri,[4] Rumana Newlands,[5] Angie A Kehagia ,[2] Jo Waller [1,3]

ABSTRACT

Introduction Preliminary studies of artificial intelligence (AI) tools developed to support breast screening demonstrate the potential to reduce radiologist burden and improve cancer detection which could lead to improved breast cancer outcomes. This study explores the public acceptability of the use of AI in breast screening from the perspective of screening-eligible women in England.

Methods 64 women in England, aged 50–70 years (eligible for breast screening) and 45–49 years (approaching eligibility), participated in 12 focus groups—8 online and 4 in person. Specific scenarios in which AI may be used in the mammogram reading process were presented. Data were analysed using a reflexive thematic analysis.

Results Four themes described public perceptions of AI in breast screening found in this study: (1) *Things going wrong and being missed* summarises a predominant and pervasive concern about an AI tool being used in breast screening; (2) *Speed of change and loss of control* captures a positive association of AI with technological advances held by the women but also feelings of things being out of their control, and that they were being left behind and in the dark; (3) *The importance of humans* reports concern around the possibility that AI excludes humans and renders them redundant and (4) *Desire for thorough research, staggered implementation and double-checking* of scans included insistence that any AI be thoroughly trialled, tested and not solely relied on when initially implemented.

Conclusions It will be essential that future decision-making and communication about AI implementation in breast screening (and, likely, in healthcare more widely) address concerns surrounding (1) the fallibility of AI, (2) lack of inclusion, control and transparency in relation to healthcare and technology decisions and (3) humans being left redundant and unneeded, while building on women's hopes for the technology.

INTRODUCTION
Background
In the UK, the National Health Service (NHS) offers breast screening every 3 years

WHAT IS ALREADY KNOWN ON THIS TOPIC
⇒ Few studies have investigated public and patient acceptability of the use of artificial intelligence (AI) in breast screening, and none have carried out an in-depth qualitative exploration of the eligible population's perceptions including a diverse sample across England.

WHAT THIS STUDY ADDS
⇒ This study found women eligible for breast screening in England have an awareness and understanding of AI while also having low confidence in its use in breast screening. Acceptability for this population depended on the evidence of, or belief in, the following: thorough research, staggered implementation, double-checking of X-rays and results, data protection and safety and continued human involvement.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY
⇒ The specific concerns, attitudes and information needs identified in this study hold major implications for communication and implementation of AI in breast screening. Not attending to the issues highlighted here risks damaging trust, alienating the public and patients, and undermining the benefits of breast screening.

to women aged 50–70 years. The programme aims to reduce breast cancer mortality through early detection and is estimated to prevent 1300 breast cancer deaths a year.[1] At present, all breast screening mammograms in England are reviewed by two trained healthcare professionals ('readers') and if there is disagreement, a third reader arbitrates.[2] However, the programme is currently facing critical staff shortages,[3] with increasing pressure on radiologists, exacerbated by backlogs from the COVID-19 pandemic,[4] and, like all screening programmes, carries harm in terms of overdiagnosis[1] and false positive results.[5]

Artificial intelligence (AI), defined by the Office for Artificial Intelligence[6] as 'the use of digital technology to create systems capable of performing tasks commonly thought to require intelligence', may help address some of these issues in breast screening.[7 8]

Freeman *et al*[6] outline four common ways in which AI has been applied to reading mammograms whereby it may (1) replace the second reader, (2) assist the reader, (3) be used as a triage tool and (4) replace all human readers. Currently, a range of AI algorithms exist, with different performance characteristics and outputs ranging from binary outcomes to more detailed reports on the location of a potential abnormality. Some of the algorithms developed for the reading of mammograms have been found to yield positive outcomes in the breast screening context.[7 9 10] AI more generally has the potential to improve screening sensitivity and specificity (distinguishing between non-cancerous lesions, lesions that ought to be monitored and cancers that require immediate treatment), address staff shortages, reduce practitioner workload and burn-out and may also be able to identify cases where women are at higher risk of having fast-growing interval cancers.[6]

In addition to the need for robust prospective studies ahead of clinical implementation, it is essential that acceptability to the public and workforce is considered. The UK National Screening Committee highlights clinical, social and ethical acceptability as key criteria to be assessed prior to implementation in any screening programme.[11] Healthcare acceptability is an overarching concept related to the approval and adoption of a healthcare tool or innovation. It comprises cognitive, affective and behavioural levels, bringing together constructs across which judgements of approval can be made.[12] There is consensus that greater acceptability leads to higher healthcare engagement including screening uptake.[12] In the case of breast cancer screening, supporting engagement is vital as lower acceptability could reduce participation rates, adversely affecting patient outcomes.[13] Patient and public involvement (PPI) in decisions regarding AI use, including those regarding its deployment and regulation, is fundamental for transparency, trust and confidence.[14 15] This is particularly so, given the opaque nature of AI algorithms, their as yet unconfirmed benefits and harms, associated regulatory challenges surrounding AI and the undetermined broader long-term impacts of AI use on individuals and societies[16 17]

Perceptions of AI in healthcare and breast screening

Previous work on public perceptions of AI in breast screening, including a meta-synthesis of 12 qualitative studies[18] and a systematic review with thematic analysis (TA),[19] distills public understanding and perceptions surrounding healthcare AI into a list of perceived benefits, concerns and wishes. However, none has examined views on AI being used in the UK's NHS Breast Screening Programme. Additionally, a deeper qualitative exploration to understand the meaning-making underpinning public perceptions of AI in breast screening could greatly benefit future communication efforts, as public perceptions around AI evolve in response to increased contact and familiarisation, technological developments, media and advertising.

Little is known about participant experience and acceptance of AI in breast screening.[20] Studies from the Netherlands,[21 22] Sweden,[23] Italy,[24] England[25] and Australia[26] are inconsistent in terms of whether, and to what degree, women want to be informed about the use of AI. None has addressed what may drive differences in information preference. A study carried out in the East Midlands region of England, and primarily among NHS employees, had limited scope.[25] Qualitative research with a larger, more diverse sample is warranted.

Research within the healthcare setting[27] as well as the wider literature on AI[28] suggests that understanding is important for establishing trust in AI. A rapid review by the Accelerated Access Collaborative identified lack of understanding and misunderstanding by the UK public surrounding AI in healthcare.[14] A stakeholder analysis conducted in Scotland concluded that a key strategy to build support for AI in breast screening is to improve knowledge for all stakeholders.[29] To improve communications surrounding AI to the public, further research on information needs is necessary. Additionally, disparities in breast screening uptake and outcomes between socioeconomic and ethnic groups are well documented,[30–34] and it is paramount that these are not exacerbated by the implementation of AI tools, or indeed by suboptimal communication around these. It is, therefore, necessary and ethically warranted that the views and concerns of women from diverse backgrounds are captured.

Aims

This study explored perceptions of AI in the context of the NHS Breast Screening Programme from the perspective of women eligible or approaching eligibility for breast screening in England.

MATERIALS AND METHODS

Design

We carried out focus groups with women from across England aged 45–70 years. 20 best practice recommendations for effectively conducting and reporting TA[35] were followed (online supplemental additional file 1).

The study protocol was preregistered on Open Science Framework (https://osf.io/gqtcr).

This study was carried out as part of an independent evaluation of Kheiron's Mia AI tool but took a vendor-agnostic approach, and neither the specific tool nor the developer was presented to participants.

Recruitment

Participants were recruited via a research participant recruitment agency (Saros Research UK; strategy 1; n=52) and through local community groups (strategy 2; n=12) between October and November 2022. The agency

recruits to their panel through social media, online advertising, print press, community outreach and word of mouth and screen every new member with a telephone call, achieving substantial demographic diversity. Strategy 1: Women from the online panel, aged between 45 and 70 and living in England were sent an invitation, participant information sheet and screener questionnaire (capturing ethnicity, AI attitudes and breast screening history) by email. Eligible women were selected and allocated to focus groups, with groups dedicated to women with often under-represented backgrounds (lower occupational social grade—C2, D or E; Indian, Pakistani, Bangladeshi backgrounds; Black, Black British, Caribbean or African backgrounds), experience (no previous breast screening) and attitudes (negative attitudes to AI; see online supplemental material for allocation by group). The use of these dedicated groups was designed to provide safe and inclusive spaces where participants would feel comfortable expressing their views.

Strategy 2: We contacted community group gatekeepers, shared fliers and posters at local libraries and visited community centres in southeast London. The recruitment materials were aimed at women between 45 and 70 years 'who have limited internet access. For example, if you do not own a smart phone/laptop, or if you sometimes have trouble using the internet'. Those interested in taking part were given or sent hard copies of the materials in advance.

Women were excluded from the study if they (1) were not a candidate for routine breast screening (eg, due to not having breast tissue or receiving annual surveillance due to elevated breast cancer risk), (2) had received treatment for breast cancer within the past 5 years, rendering them ineligible for routine screening, (3) had participated in a study recruited through the recruitment agency within 6 months prior to the focus groups or (4) were unable to consent to or participate in a focus group conducted in English.

Patient and public involvement
The study advisory group included five PPI contributors, recruited through National Institute for Health and Care Research (NIHR)'s People in Research, who met the eligibility criteria for regular breast screening and differed by age, location across England, ethnicity and previous PPI experience. They provided feedback on the research plan, screener questionnaire and topic guide. All study materials were reviewed by two additional PPI representatives meeting the study eligibility criteria.

Data collection
The focus groups were offered both in-person and online (via Microsoft Teams), were held in November 2022, ran for 90 minutes and were audio (in-person) or video-recorded (online). LG facilitated all groups, following a semistructured topic guide (see online supplemental material) and was supported by a notetaker for each group (SA and PM). The topic guide was informed by models

of healthcare acceptability[12] and utilisation,[36] as well as findings from past AI public perceptions research[14 22–25 37] and supplemented with feedback from expert advisors. Participants were first prompted to share their perceptions of breast screening, then their general perceptions of AI and were then shown four example scenarios based on possible uses of AI in breast screening[6] (see, figure 1 and online supplemental additional file 1).

At the beginning of each focus group, the facilitator (LG; 30-year-old, white, female) disclosed that she did not have a background in breast screening or AI. This served to position the facilitator as a non-expert whose role was to support and capture participants' thoughts and opinions without bias. No vendor or product names were shared. Recordings were transcribed verbatim by an external company and checked for accuracy by LG, SA or PM.

Data analysis
Reflexive TA was employed following Braun and Clarke's phases,[38] adopting a constructionist approach to account for participant interactions in the focus group context. Nowell et al's[39] quality criteria for TA were followed (see also online supplemental additional file 1).

RESULTS
Sample characteristics
Table 1 presents participant characteristics. We were successful in recruiting women from a range of backgrounds with varying experience and attitudes towards breast screening and AI. Participants' quotes are labelled by group (G) and participant number (P): G#P#.

Contextual factors
There were three key characteristics of the participants and the context within which the focus groups occurred. First, the topics of cost and money were overriding concerns. Participants most often reflected on the NHS being strapped for resources, being 'under pressure' and being 'sold off'. This focus on money was wrapped up with societal-level financial concerns and recent rises in living costs in the UK. Second, before being provided with any information, participants either conceptualised AI as machines and robotics or as information processing and computing. Third, feelings participants expressed towards AI in general ranged from support and enthusiasm (with participants proposing and promoting the possible benefits of AI), to wariness, scepticism and fear (with participants using descriptions like scary, creepy and frightening).
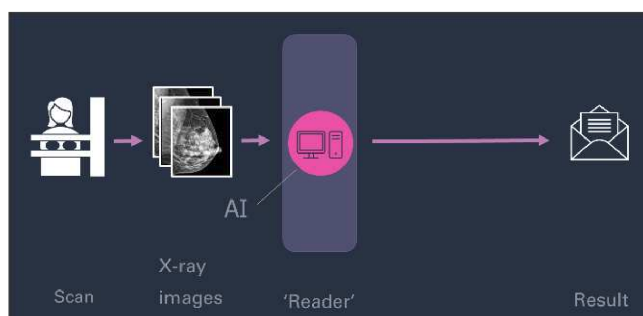
Main themes
The four themes presented in box 1 capture key patterns across participants' views and concerns surrounding AI use within the NHS Breast Screening Programme, with the perceptions captured in theme 4 tied to the issues discussed in the other three themes.
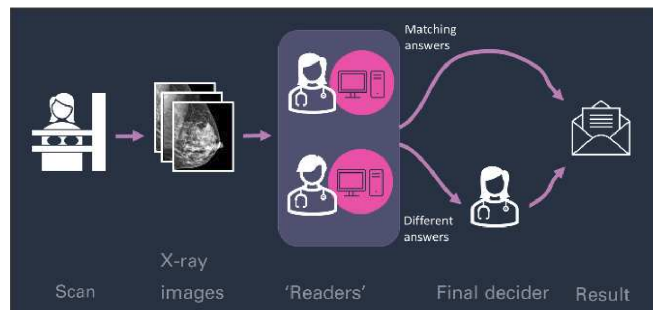
Scenario 1 – One human plus one AI reader
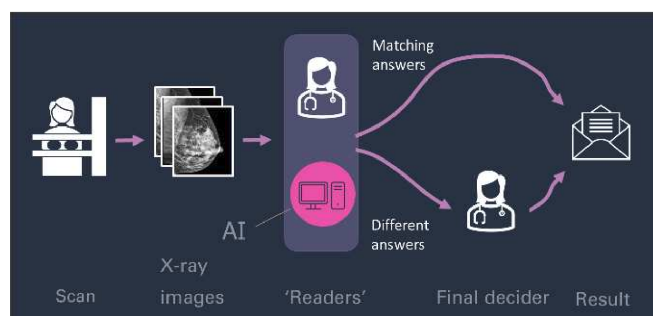
Scenario 2 – AI alone

Scenario 3 – AI support for two human readers
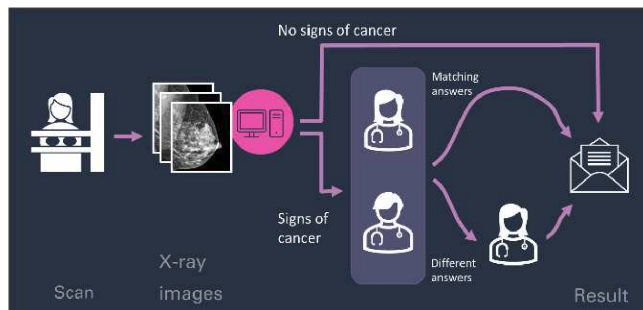

Scenario 4 – AI as triaging tool

Figure 1 Scenarios of possible uses of AI within the breast screening pathway presented to participants. AI, artificial intelligence.

Theme 1: things going wrong and being missed

There was support and interest in the use of AI in breast screening based on the premise that it could bring benefits or improvements. However, a prominent concern expressed across groups was of things going wrong or being missed while also questioning the reliability of human readers in the current breast screening process. This was tied to a perception that breast screening is a context where it is particularly important for things to 'go right', due to the intimate nature of the procedure and the significance of missing a cancer. Conversely, participants did not express concerns about false positives or overdiagnosis, which is consistent with poor public awareness of these risks.[40]

Subtheme: AI going wrong and human error

Participants frequently brought up that AI 'can go wrong' (G1P3) due to power outages, hacking, viruses, glitches, corruption of data, quality of data or programming. There was a perception that AI is reliable at doing what it is programmed to do and that limitations would be a result of who programmed it, how it was programmed and what it had been programmed to include. Only one participant mentioned the possibility of AI deviating from its programming by 'not staying within parameters' (G12P4).

There was an expectation that an AI tool should be proved to be infallible before being clinically implemented, with several participants explicitly insisting that it should be '100% accurate' (G1P3). Others, however, indicated some tolerance around this, describing AI as

being in its infancy with allowance for some errors that might prompt developers to make improvements.

Participants also regularly brought up that humans are likely to make mistakes or often do, citing examples in healthcare but, most often, by simply stating that 'humans make mistakes' (G1P3) and 'there is a certain amount of human error in everything' (G9P3). This was sometimes presented as a reason why AI might be useful or was offered in response to other participants saying AI can go wrong.

Subtheme: the importance of things not being missed

The issue of things going wrong in breast screening was inextricably tied to the fear of cancer being missed. Repeatedly, and across nearly all groups, participants brought up the concern of 'What if something got missed? A vital thing?' (G3P3) and patients falling through the gaps. Participants also often spoke about the importance and desirability of 'accuracy', mostly referring to the risk of cancers being missed.

For participants, the stakes in breast screening were high—likely because of the perceived severity of cancer. This notion of risks and stakes is captured in the following extract when discussing reliance on an AI tool in breast screening:

> G6P2 when it's people's health and lives at stake, you know, it's, um, yeah. It's scary. G6P3 If you do that with people's health, that's too risky

Across many of the groups, participants referred to health, cancer or breast screening as being particularly

Table 1 Participant characteristics across the sample (n=64)

	n
Total	64
Ages	
45–49	8
50–54	15
55–59	12
60–64	15
65–70	14
Occupation-based social grade	
AB (higher and intermediate managerial, administrative, professional)	11
C1 (supervisory, clerical and junior managerial, administrative, professional)	24
C2 (skilled manual)	20
DE (semiskilled and unskilled manual occupations, unemployed and lowest grade)	8
Highest formal education	
No formal qualifications	3
GCSE or equivalent	18
A level or equivalent	15
Undergraduate degree or equivalent	14
Postgraduate degree or equivalent	13
Ethnicity	
White	33
Black	15
Asian	11
Mixed or multiple ethnic groups	3
Other ethnic groups	1
Location	
London	18
Northwest	11
West midlands	10
Southeast	8
Yorkshire and the Humber	5
East of England	5
Southwest	4
East midlands	4
Which of the following technologies (that are supported by AI) have they used or would they use if money were not a barrier?	
Video streaming	49
Email service	56
Smart watch	30
Smart glasses	9
VR headset	14
Smart speaker	33
How many of the above technologies have they used or would they use if money were not a barrier?	
0	2
1 or 2	23
3 or 4	28

Continued

Table 1 Continued

	n
5 or 6	10
Feelings towards AI	
Mostly positive	31
Mostly negative	7
Undecided	19
Do not have strong feelings one way or the other	6
Breast screening experience	
Not been invited but intend to go	4
Not been invited but intend not to go or unsure	7
Been invited, not been but intend to go	5
Been invited, not been and intend not to go or unsure	2
Been to screening and intend to go again	36
Been to screening but do not intend to go again or are unsure	9
Have you, anyone in your family or any of your friends had cancer?	
Self	8
Partner	2
Someone in close family	27
Someone is wider family	23
Close friend	28
Acquaintance or colleague	36
None of the above	2
Prefer not to say	2
Of those who have had cancer themselves, was it breast cancer	
Yes	3
No	5
Of those with family who have had cancer, was this a blood relative	
Biological family	30
Non-biological family	14
Not sure	1

GCSE, General Certificate of Secondary Education; VR, virtual reality.

personal or important, conveying the importance placed on things going right and cancer not being missed.

Theme 2: speed of change and loss of control

A theme running through the groups was a sense of lack of control, usually overlapping with participants' descriptions of AI as being part of technological advances which were often felt to be moving too fast. There were also positive views around the potential of AI that were tied to perceptions of progress and advancements.

Subtheme: technological progress and change

Participants described AI as novel, futuristic and advanced: 'in my head it was like this super-duper amazing new technology [-] But realistically, we use AI all the time' (G12P1); '[AI is] the way forward' (G5P1). In contrast with the futuristic nature of AI, mammography (the appointment process and experience of having the mammogram taken) was considered outdated

Theme 1: things going wrong and being missed
Both artificial intelligence (AI) and human error could lead to cancer being missed, against the important and intimate nature of breast cancer screening and detecting cancer.

Theme 2: speed of change and loss of control
AI as part of technological progress and a feeling of women being left behind and in the dark, with decisions being out of their hands.

Theme 3: the importance of humans
Humans being replaced and concerns around where decision-making would lie.

Theme 4: Desire for thorough research, staggered implementation, and double-checking (which was brought up in response to and projected as necessary in light of the concerns described in themes 1–3).
AI should be shown to be effective through thorough research before being implemented, should not be used alone until shown to be effective, and scans should be double-checked to reduce risk of cancer being missed.

and old-fashioned, with participants saying that, by now, screening ought to be better than it is. Participants also expressed nervousness about AI and technological progress, and their responses suggested a feeling of technological advancement racing forward at an uncontrolled speed ('technology now is moving just so fast that we're feeling the impact a lot faster and in a lot more aggressive way' G5P3; 'I think those who are making it and producing it push it too hard and too fast, and that they just need to be reined in a bit' G6P3). Some participants felt that the rapidly increasing speed and number of technologies being produced increased the chances that problems would occur.

There was also concern about what AI might become in the future—indicated in the following exchange:

G1P3 The machine cannot do it by itself. G1P5 Well, for now –

This was tied to an idea that the future consequences of AI might not be predictable or comprehendible.

AI was described as something that is here already, in everything and an inevitability no matter how they feel about it. In some cases, this was presented positively, as well as there being 'so many possibilities' (G1P6) but, more often, was related to a sense of AI taking over and being invasive, with some participants using the term 'big brother' (G6P4 & G9P3). Some participants expressed that AI is part of the 'technological age (G5P1), inevitable and out of their control:

G3P3 it is all part of our life, and it's coming everywhere right around. G3P4 yeh there's nothing we can do.

Participants frequently brought up the topic of change ('because every day, everything changes, don't they?'; G1P3), likely prompted by thinking about AI in terms of technological progress. Participants' discussions

indicated a challenge, pressure or need to accept or get used to change associated with new technologies.

Subtheme: exclusion in healthcare and loss of agency
Participants conveyed a pervasive sense of being left in the dark about technological advances and health-related matters, with the interests of women not being placed at the centre of decisions made in healthcare or the proposed AI in breast screening. Participants reported feeling that the NHS does things without individuals' consent. Some believed that AI would be implemented without gaining their consent while many others mentioned not having a say in decisions more generally. This was against a backdrop of perceived disregard for women and their comfort by the health system, with participants stating this or giving examples of where they have felt let down or overlooked in the past. For example, 'they don't seem to put that much emphasis on doing anything, anything comfortable, um, for women.' (G3P2). There was also a perceived lack of transparency from the healthcare system, with participants saying that women are not informed adequately or would not be told about AI's role and roll-out, in breast screening.

Within this context, participants emphasised the importance of women being informed, having informed choice and being an 'informed patient' (G4P4). Some participants said that women should be told if AI was being used in breast screening although this was not a consensus view. Those who felt they did not need this information most often framed this in the context of having no knowledge of the current mammogram-reading process anyway and breast screening being worth it, regardless of the process.

Subtheme: data protection wariness and ownership considerations
Participants were aware as well as wary of data sharing rights and listed several data protection considerations and concerns, including companies profiting from their data, whether the data sharing is regulated, having to concede to share their data in exchange for healthcare benefits and having to rely on data collectors/AI users to be competent and have integrity. Vitally, participants focused on the following factors about their personal data: who owns them, where and how securely they are stored, who has access, who's sharing and who's profiting as well as confidentiality, privacy and opting in or out of such schemes. Participants indicated they felt their own data, and that of other patients, should only be shared and used to serve public interests and that private companies should not benefit from personal data unless without clear equivalent benefit to the public/patients.

Theme 3: the importance of humans
Participants expressed concern about human readers being removed from the breast screening pathway, with many asserting that humans must be included or involved. This was tied to a negative feeling that humans are rendered redundant by AI. It is likely that the idea of replacement was sparked by the presentation of AI as

taking the role of a radiologist if used in the mammogram reading pathway (as depicted in figure 1). Within this context of the potential for human redundancy, participants affirmed the importance of humans.

Also, more generally, participants' consideration of AI was framed around humans. When AI was described or considered, it was usually in comparison to humans, considering the strengths, flaws, chances of errors or mistakes, potential impacts or what they would want to be told about AI if implemented. Participants focus on humans when discussing AI indicates a tension between humans and AI in participants' thoughts.

Participants often asserted that humans must be involved at some point in the screening process, with some wanting humans to be involved at every stage. They attested to the importance of human contact and affirmed that having a human involved is reassuring. Some had concerns about AI making decisions alone: 'You have to have human beings to give the overall decision' (G10P3), and this was linked to wariness and scepticism around the capabilities of AI.

The essential role of humans was grounded in the unique strengths humans can bring to breast screening, including being able to provide opinions and emotional support; and that humans have professional experience and training. The idea that AI is more capable than humans was raised by participants, whether asserting the belief or opposing it. The following quote captures both the threat of human redundancy from AI and an assertion of the uniqueness of the medical context as to why humans are still necessary:

> G2P1: in this day and age, robotics are replacing the human, hugely. And we are basically becoming redundant [-] I do believe that that human beings still have a place in the workplace and especially in medicine, medicine and veterinary. Anything to do with life.

There was a recurring idea that humans and AI would or should be working together; sometimes, specifically the humans assisting the AI tool, other times the AI tool assisting the humans. Participants also described AI as a tool for humans to use or suggested this is how it should be approached.

Many participants raised concerns about AI taking people's jobs. Most often, there was worry, contestation and indignation but some participants were accepting of these possible impacts on radiographer and radiologists' jobs. Some also reflected on the potential positive impact AI might have within the context of a resource-strapped and struggling healthcare service, heavily related to the wider context of financial concerns.

Theme 4: desire for thorough research, staggered implementation and double-checking

This final theme captures conditions that participants deemed should be in place or assured of before AI is to be used in the mammogram reading pathway, namely research, staggered implementation, double-checking

and safeguards. Together, these demonstrate a desire for the utmost to be done to ensure that AI is accurate and trustworthy.

Across groups, participants repeatedly suggested that they would want any AI tool to be robustly trialled and tested. This was tied to participants saying an AI tool must prove that it gives the right, accurate or good quality results. Although the notion of 'accuracy' was not clearly defined by most, some participants referred to 'how many mistakes have been made' (G8P2) and 'picking up things that humans aren't' (G9P2).

The outcomes that participants focused on were nearly always how AI compared with the human readers, but also how the new system compared with the old system. Participants saw the amount of time (whether that be to develop, test, trial or implement) as a main indicator of the strength of evidence, with some participants suggesting 'a few years' (G2P2) to 'many, many years' (G4P1) is necessary to produce strong enough evidence. It is likely that participants' focus on the amount of time an AI tool ought to be studied for, was linked to the perceived newness of AI and fear of progress being too fast, but also tied to the idea that research takes a long time. To a lesser degree, the following things were also considered indicators of strong evidence: the number of trials, the number of people in trials, the diversity of trial participants and whether trials had been conducted in other countries and for other screening tests or conditions.

Participants were keen for AI to be implemented into breast screening in a staggered way and often spoke about what they thought needed to be in place 'in the beginning' (G5P1) and 'as a first step' (G8P4). This included human involvement, participation being optional and running alongside the current system. A desire for staggered implementation was tied to a perception that the AI tool was yet to be trialled or starting to be trialled. Many participants expressed that relying on AI, and without it being trialled, would carry unacceptable risk: 'To leave from where we are to the fourth scenario [where AI is triaging] seems a bit too risky to me' (G5P3); 'until it's embedded in and it's—it's widely known and it's reliable, it's—it's risky' (G8P1).

Participants' responses indicated a desire to have mammograms checked as much as possible and, with this, the inclusion of AI as another reader was viewed positively. Many times, participants brought up the importance of the AI and the human readers to check on one another. Participants also used the terms 'verifying' (G5P2) and 'confirmation' (G10P3). This was not about one checking on the other but about the fact the mammogram is being more 'thoroughly investigated' which would give people more 'confidence' (G12P5). In a similar vein, participants also wanted backup and safety measures to be in place (regarding evaluation and implementation indiscriminately), with some participants bringing up 'backup safety measures' (G5P5), having a 'safety net' (G12P5), 'audit' (G4P2), 'safeguards' (G11P2).

Box 2   Key points from the study

1. There is a growing awareness and understanding surrounding what artificial intelligence (AI) is, with many in the group correctly discussing AI as relating to information processing. However, for some participants, the term conjured up ideas of automated robots and something futuristic.
2. There was concern and a lack of confidence around the reliability of AI and, with this, a demand for AI not to be relied on without some form of human involvement.
3. Women's key concerns were around:
   – AI adversely affecting cancer outcomes.
   – AI being thoroughly trialled and tested.
   – Need for human oversight.
   – AI outcomes should be monitored and checked.
   – Patient data should be protected.
4. Careful communication will be needed to ensure public and patient acceptability prior to any implementation of AI in healthcare (such as AI reading mammograms).

DISCUSSION

This study expands knowledge about public understanding and attitudes towards AI and its possible use in the NHS Breast Screening Programme as part of the mammogram reading process, from the perspective of women eligible for breast screening in England. The key findings and implications from this study are outlined in box 2.

This study found an openness to the potential benefits of AI for patients, cancer detection and the healthcare service, as previously found in other studies in northern Europe,[21 23] but also a belief that AI decisions cannot be relied on entirely, consistent with a survey study in Italy.[24] If AI is introduced into breast screening, any evidence regarding its performance included in information materials for invitees should be shared in a manner that is comprehensible and transparent. A large survey in Germany similarly found that participants wished to see performance evidence for each healthcare AI case before making a judgement about acceptability.[41]

In line with previous quantitative and qualitative studies in the Netherlands and Sweden, we found there was greater confidence where AI accompanies the work of human readers[21–23] and there was a demand for human interaction and involvement, as was seen in a small survey study among hospital staff in England.[25] Concerns over the loss of human involvement have been a dominant finding in research on patient and public perceptions of AI in healthcare generally.[42] A key implication is that, currently, ongoing human involvement in the breast screening programme is necessary for AI to be deemed acceptable by the target population. Additionally, communication about any switch to using AI in the screening process will need to provide assurance that humans will still be involved in decision-making and interactions with patients. Previous qualitative work with stakeholders in the UK has shed light on when, and in what form, humans might need to be involved

in clinical pathways that use AI, from a legal and ethical perspective.[43]

The themes, reported here, of viewing AI as futuristic and very advanced, concerns over the speed of change and feelings of loss of control in regard to AI in healthcare have emerged previously in digital dialogue and focus groups with members of the public and other key stakeholders in the UK.[44] Future communication might wish to avoid distancing the screening-eligible population from engaging with breast screening information when introducing AI by addressing such perceptions directly.

Women in our study felt that the healthcare system disregards women's comfort and preferences. This is a challenge for the roll-out of new technologies in the NHS, given the importance of perceived benevolence and perceived integrity in establishing trust.[45] Our work highlights that screening-eligible women may feel loss of agency in relation to AI, that they have no say in how AI will change their lives and that it is futile to engage with healthcare AI decisions. Poor communication about AI risks exacerbating these negative feelings, undermining public trust and negatively impacting breast screening attendance. Additionally, it is essential that women feel their interests are placed at the centre of decision-making through continuous efforts and research to include them in these processes, to avoid disenfranchising and disempowering women and the wider public.

Participants desired strong evidence before accepting AI use in breast screening, echoing the cautious approach taken by the UK National Screening Committee.[11] This is unsurprising given the lack of supporting evidence provided during the focus groups and their uncertainty around AI capabilities. To support acceptance, future communications about AI use in breast screening should provide evidence or assurances of appropriate AI testing, as well as clearly define the intended purpose and outcomes of implementing the AI tool. Participants were focused on traditional heuristics of trial quality; that the AI tool is tested for a long time and with large samples. Communicators may wish to raise awareness about other important indicators of quality such as diversity or completeness of test data.[46 47]

Participants held the common misconception that screening is a diagnostic test, that is, it provides a definitive answer as to whether a person has cancer.[48] This was borne out in participants' desire for as much checking as possible of the mammograms to confirm the presence or absence of cancer. Thus, probability, and the uncertainty inherent in all screening, appear to be important to include in communication materials, in a manner that is accessible and understood by the public.[49]

Participants shared concerns about data security and data rights (including privacy and ownership). However, unlike other studies, participants in our focus groups did not reflect on specific regulations, regulators or accountability[21 50] suggesting these

topics may not be common frames of reference for those eligible for breast screening. This does not minimise the importance of having legal and regulatory frameworks but does suggest more research is needed to understand their impact on public acceptance of AI being used in the NHS Breast Screening Programme.

Our approach to understanding acceptability was informed by the Theoretical Framework of Acceptability[12] (TFA), ensuring we took a broad view of the concept. Although the themes we identified did not map directly onto the TFA constructs, the idea of perceived effectiveness was pervasive throughout discussions about how accurate an AI tool would be and the need for research ahead of implementation. Affective attitude and ethicality, two other TFA constructs, were implicit in participants' feelings about the potential loss of control and humans being replaced. The potential loss of human involvement could also fall within the opportunity costs construct of the TFA. Given that an AI tool would not have direct impact on women's experience of taking part in screening, self-efficacy and perceived burden were less relevant. Other more behaviour-focused theories may be useful in future work seeking to understand whether the introduction of AI would affect intentions to participate more specifically.

Strengths

This study included women from varied demographic groups, across geographical regions in England and with a range of views and experiences of breast screening and AI. We carefully considered group composition to maximise the range of attitudes we could elicit. Additionally, this is the first study to investigate acceptability of AI in radiology or healthcare using reflexive TA which allowed for an in-depth understanding of public perceptions—allowing for not only collecting what the public report to be their concerns (e.g., wanting human involvement) but also offering explanations as to why these concerns are reported (e.g., doubt over AI's reliability, feelings of security and comfort from human presence, retaliation to sense of human redundancy).

Limitations

Although our sample was diverse, we had a higher-than-expected number of participants with university-level education. Our second recruitment strategy did allow us to recruit women with lower digital literacy, including four who did not use email, although this was limited to London. Additionally, we were unable to include participants who did not speak English. Three of the women who took part did not meet the eligibility criteria as they had been treated for breast cancer within the previous 5 years (n=2) or were receiving yearly screening due to being at high risk (n=1), but our analysis suggested their views were similar to those of other participants. Future studies

should focus on the inclusion of people from unrepresented educational backgrounds and non-English speakers to ensure inclusivity and representativeness.

This study represents a snapshot of public perceptions at one point in time. It would be worthwhile capturing public perceptions over time while tracking exposure to information on AI technologies (eg, greater everyday use; news coverage, high-profile stories or controversy; technological developments). This would provide valuable insights into the evolution of public attitudes and factors most influential in this.

Although not the aim of the present study, exploring the views of radiologists and radiographers working within the screening programme is essential ahead of implementation and is currently underway by King's Technology Evaluation Centre in the context of the same commissioned evaluation. Similar work is being carried out in Portugal,[51] the USA[52] and Sweden.[53]

CONCLUSION

The findings from this study reinforce the importance of the inclusion of women as key stakeholders in decisions about the use of AI in the NHS Breast Screening Programme and the need for clear public communication surrounding the implementation of such a system. Optimal communication could include information on the role of healthcare professionals, the extent to which the AI tool has been tested and other assurances around its reliability and steps being taken to ensure data security, taking into account the diverse views and concerns expressed by our participants. It is hoped that this work will serve to amplify their voices and support decision-makers, policy-makers and healthcare professionals in selecting, communicating and implementing AI.

Author affiliations
[1]Cancer Prevention Group, School of Cancer & Pharmaceutical Sciences, King's College London, London, UK
[2]King's Technology Evaluation Centre, School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK
[3]Centre for Cancer Screening, Prevention and Early Diagnosis, Wolfson Institute of Population Health, Queen Mary University of London, London, UK
[4]School of Mental Health & Psychological Sciences, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[5]Health Services Research Unit, University of Aberdeen, Aberdeen, UK

ORCID iDs
Lauren Gatting http://orcid.org/0000-0003-3693-2013
Angie A Kehagia http://orcid.org/0000-0002-5142-7456
Jo Waller http://orcid.org/0000-0003-4025-9132

REFERENCES

1 Marmot MG, Altman DG, Cameron DA, *et al*. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* 2013;108:2205–40.
2 NHS England. Breast screening: guidance for image reading. 2024. Available: https://www.gov.uk/government/publications/breast-screening-guidance-for-image-reading/breast-screening-guidance-for-image-reading
3 Rimmer A. Radiologist shortage leaves patient care at risk, warns royal college. *BMJ* 2017;j4683.
4 The Royal College of Radiologists. Clinical radiology workforce census 2022.
5 Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013;2013:CD001877.
6 Freeman K, Geppert J, Stinton C, *et al*. Use of artificial intelligence for mammographic image analysis in breast cancer screening: rapid review and evidence map. 2022. Available: https://nationalscreening.blog.gov.uk/
7 Bahl M. Artificial Intelligence: A Primer for Breast Imaging Radiologists. *J Breast Imaging* 2020;2:304–14.
8 Zizaan A, Idri A. Machine learning based Breast Cancer screening: trends, challenges, and opportunities. *Comput Methods Biomech Biomed Eng Imaging Vis* 2023;11:976–96.
9 Schaffter T, Buist DSM, Lee CI, *et al*. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. *JAMA Netw Open* 2020;3:e200265.

10 Sharma N, Ng AY, James JJ, *et al*. Retrospective large-scale evaluation of an AI system as an independent reader for double reading in breast cancer screening. *medRxiv* 2021.
11 UK National Screening Committee. Criteria for a population screening programme. 2022. Available: https://www.gov.uk/government/publications/evidence-review-criteria-national-screening-programmes/criteria-for-appraising-the-viability-effectiveness-and-appropriateness-of-a-screening-programme
12 Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res* 2017;17:88.
13 Mandrik O, Tolma E, Zielonke N, *et al*. Systematic reviews as a 'lens of evidence': Determinants of participation in breast cancer screening. *J Med Screen* 2021;28:70–9.
14 Accelerated Access Collaborative. Public perceptions and attitudes to artificial intelligence (AI) in healthcare an exploratory study. 2022.
15 Patel R. Keeping society in the loop about data ethics and ai. Ada Lovelace Institute; 2018. Available: https://www.adalovelaceinstitute.org/blog/keeping-society-in-the-loop-about-data-and-ai/
16 Hickman S. *Evaluating artificial intelligence in breast cancer screening*. University of Cambridge, 2022.
17 Hunter B, Hindocha S, Lee RW. The Role of Artificial Intelligence in Early Cancer Diagnosis. *Cancers (Basel)* 2022;14:1524.
18 Wu C, Xu H, Bai D, *et al*. Public perceptions on the application of artificial intelligence in healthcare: a qualitative meta-synthesis. *BMJ Open* 2023;13:e066322.
19 Vo V, Chen G, Aquino YSJ, *et al*. Multi-stakeholder preferences for the use of artificial intelligence in healthcare: A systematic review and thematic analysis. *Soc Sci Med* 2023;338:116357.
20 Hemphill S, Jackson K, Bradley S, *et al*. The implementation of artificial intelligence in radiology: a narrative review of patient perspectives. *Future Healthc J* 2023;10:63–8.
21 Haan M, Ongena YP, Hommes S, *et al*. A Qualitative Study to Understand Patient Perspective on the Use of Artificial Intelligence in Radiology. *J Am Coll Radiol* 2019;16:1416–9.
22 Ongena YP, Yakar D, Haan M, *et al*. Artificial Intelligence in Screening Mammography: A Population Survey of Women's Preferences. *J Am Coll Radiol* 2021;18:79–86.
23 Jonmarker O, Strand F, Brandberg Y, *et al*. The future of breast cancer screening: what do participants in a breast cancer screening program think about automation using artificial intelligence? *Acta Radiol Open* 2019;8:2058460119880315.
24 Pesapane F, Rotili A, Valconi E, *et al*. Women's perceptions and attitudes to the use of AI in breast cancer screening: a survey in a cancer referral centre. *Br J Radiol* 2023;96:20220569.
25 Lennox-Chhugani N, Chen Y, Pearson V, *et al*. Women's attitudes to the use of AI image readers: a case study from a national breast screening programme. *BMJ Health Care Inform* 2021;28:e100293:1–8:.
26 Carter SM, Carolan L, Saint James Aquino Y, *et al*. Australian women's judgements about using artificial intelligence to read mammograms in breast cancer screening. *Digit Health* 2023;9.
27 Jermutus E, Kneale D, Thomas J, *et al*. Influences on User Trust in Healthcare Artificial Intelligence: A Systematic Review. *Wellcome Open Res* 2022;7:65.
28 Balfe N, Sharples S, Wilson JR. Understanding Is Key: An Analysis of Factors Pertaining to Trust in a Real-World Automation System. *Hum Factors* 2018;60:477–95.
29 Newlands R, Bruhn H, Díaz MR, *et al*. A stakeholder analysis to prepare for real-world evaluation of integrating artificial intelligent algorithms into breast screening (PREP-AIR study): a qualitative study using the WHO guide. *BMC Health Serv Res* 2024;24:569.
30 Douglas E, Waller J, Duffy SW, *et al*. Socioeconomic inequalities in breast and cervical screening coverage in England: are we closing the gap? *J Med Screen* 2016;23:98–103.
31 Bhola J, Jain A, Foden P. Impact of index of multiple deprivation and ethnicity on breast screening uptake in the North West of England. *Breast Cancer Res* 2015;17:24.
32 Office of National Statistics, Cooke A, Butt A, *et al*. Mortality from leading causes of death by ethnic group, England and Wales. ONS.gov.uk; 2021. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/mortalityfromleadingcausesofdeathbyethnicgroupenglandandwales
33 Cancer Research UK, National Cancer Intelligence Network. Cancer by deprivation in England: incidence, 1996-2010, mortality, 1997-2011. 2014. Available: http://www.ncin.org.uk/about_ncin/cancer_by_deprivation_in_england
34 Fry A, White B, Nagarwalla D, *et al*. Relationship between ethnicity and stage at diagnosis in England: a national analysis of six cancer sites. *BMJ Open* 2023;13:e062079.

35 Braun V, Clarke V. Is thematic analysis used well in health psychology? A critical review of published research, with recommendations for quality practice and reporting. *Health Psychol Rev* 2023;17:695–718.

36 Venkatesh V, Smith RH, Morris MG, *et al*. User Acceptance of Information Technology: toward a Unified View. *MIS Q* 2003;27:425.

37 Antes AL, Burrous S, Sisk BA, *et al*. Exploring perceptions of healthcare technologies enabled by artificial intelligence: an online, scenario-based survey. *BMC Med Inform Decis Mak* 2021;21:221.

38 Braun V, Clarke V. *Thematic analysis: a practical guide*. Los Angeles: SAGE, 2022.

39 Nowell LS, Norris JM, White DE, *et al*. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *Int J Qual Methods* 2017;16.

40 Mathioudakis AG, Salakari M, Pylkkanen L, *et al*. Systematic review on women's values and preferences concerning breast cancer screening and diagnostic services. *Psychooncology* 2019;28:939–47.

41 Safi S, Danzer G, Schmailzl KJ. Empirical Research on Acceptance of Digital Technologies in Medicine Among Patients and Healthy Users: Questionnaire Study. *JMIR Hum Factors* 2019;6:e13472.

42 McCradden MD, Sarker T, Paprica PA. Conditionally positive: a qualitative study of public perceptions about using health data for artificial intelligence research. *BMJ Open* 2020;10:e039798.

43 Redrup Hill E, Mitchell C, Brigden T, *et al*. Ethical and legal considerations influencing human involvement in the implementation of artificial intelligence in a clinical pathway: A multi-stakeholder perspective. *Front Digit Health* 2023;5:1139210.

44 Suffield M, Busby A, Roberts N, *et al*. BEIS public attitudes to science digital dialogues-wave 2 perceptions and acceptability of the use of artificial intelligence with a focus in healthcare report. 2019. Available: https://www.kantar.com/-/media/project/kantar/global/public-polls/beis-pas/beis-pas---dialogue---wave-2.pdf

45 Dietz G, Den Hartog DN. Measuring trust inside organisations. *Pers Rev* 2006;35:557–88.

46 Daneshjou R, He B, Ouyang D, *et al*. How to evaluate deep learning for cancer diagnostics - factors and recommendations. *Biochim Biophys Acta Rev Cancer* 2021;1875:188515.

47 Wu E, Wu K, Daneshjou R, *et al*. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med* 2021;27:582–4.

48 Dunlop K, Rankin NM, Smit AK, *et al*. Acceptability of risk-stratified population screening across cancer types: Qualitative interviews with the Australian public. *Health Expect* 2021;24:1326–36.

49 Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. *J Natl Cancer Inst* 2011;103:1436–43.

50 Ongena YP, Haan M, Yakar D, *et al*. Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. *Eur Radiol* 2020;30:1033–40.

51 Calisto FM, Santiago C, Nunes N, *et al*. BreastScreening-AI: Evaluating medical intelligent agents for human-AI interactions. *Artif Intell Med* 2022;127:102285.

52 Hendrix N, Lowry KP, Elmore JG, *et al*. Radiologist Preferences for Artificial Intelligence-Based Decision Support During Screening Mammography Interpretation. *J Am Coll Radiol* 2022;19:1098–110.

53 Kloub M, Gerigoorian A. A cross-sectional technology acceptance study of an AI CAD system in a breast screening unit. KTH Royal Institute of Technology in Stockholm; 2021. Available: https://kth.diva-port