BJR https://doi.org/10.1259/bjr.20220569

Received: 01 June 2022 | Accepted: 11 September 2022 | Published online: 15 November 2022

© 2023 The Authors. Published by the British Institute of Radiology under the terms of the Creative Commons Attribution-NonCommercial 4.0 Unported License http://creativecommons.org/licenses/by-nc/4.0/, which permits unrestricted non-commercial reuse, provided the original author and source are credited.

Cite this article as:
Pesapane F, Rotili A, Valconi E, Agazzi GM, Montesano M, Penco S, 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) 10.1259/bjr.20220569.

FULL PAPER

Women's perceptions and attitudes to the use of AI in breast cancer screening: a survey in a cancer referral centre

[1]FILIPPO PESAPANE, [1]ANNA ROTILI, [2]ELENA VALCONI, [2]GIORGIO MARIA AGAZZI, [1]MARTA MONTESANO, [1]SILVIA PENCO, [1]LUCA NICOSIA, [1]ANNA BOZZINI, [1]LORENZA MENEGHETTI, [1]ANTUONO LATRONICO, [1]MARIA PIZZAMIGLIO, [3]ELEONORA ROSSERO, [4]AURORA GAETA, [4]SARA RAIMONDI, [5]SILVIA FRANCESCA MARIA PIZZOLI, [6]ROBERTO GRASSO, [7]GIANPAOLO CARRAFIELLO, [5,6]GABRIELLA PRAVETTONI and [1]ENRICO CASSANO

[1]Breast Imaging Division, IEO European Institute of Oncology IRCCS, Milan, Italy
[2]Diagnostic and Interventional Radiology Unit, Department of Diagnostic and Therapeutic Advanced Technology, Azienda Socio Sanitaria Territoriale Santi Paolo and Carlo Hospital, Milan, Italy
[3]Laboratorio dei Diritti Fondamentali, Collegio Carlo Alberto, Torino ER, Turin, Italy
[4]Department of Experimental Oncology, IEO European Institute of Oncology IRCCS, Milan, Italy
[5]Dipartimento di Oncologia ed Emato Oncologia, Università degli studi di Milano, Milan, Italy
[6]Applied Research Division for Cognitive and Psychological Science, IEO European Institute of Oncology IRCCS, Milan, Italy
[7]Department of Radiology and Department of Health Sciences, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico and University of Milano, Milan, Italy

Address correspondence to: Dr Filippo Pesapane
E-mail: *Filippopesapane@gmail.com*

Objective: Although breast cancer screening can benefit from Artificial Intelligence (AI), it is still unknown whether, to which extent or under which conditions, the use of AI is going to be accepted by the general population. The aim of our study is to evaluate what the females who are eligible for breast cancer screening know about AI and how they perceive such innovation.
Methods: We used a prospective survey consisting of a 11-multiple-choice questionnaire evaluating statistical associations with Chi-Square-test or Fisher-exact-test. Multinomial-logistic-regression was performed on items with more than two response categories. Odds ratio (OR) with 95% CI were computed to estimate the probability of a specific response according to patient's characteristics.
Results: In the 800 analysed questionnaires, 51% of respondents confirmed to have knowledge of AI. Of these, 88% expressed a positive opinion about its use in medicine. Non-Italian respondents were associated with the belief of having a deep awareness about AI more often than Italian respondents (OR = 1.91;95% CI[1.10–3.33]). Higher education level was associated with better opinions on the use of AI in medicine (OR = 4.69;95% CI[1.36–16.12]). According to 94% of respondents, the radiologists should always produce their own report on mammograms, whilst 77% agreed that AI should be used as a second reader. Most respondents (52%) considered that both the software developer and the radiologist should be held accountable for AI errors.
Conclusions: Most of the females undergoing screening in our Institute approve the introduction of AI, although only as a support to radiologist, and not in substitution thereof. Yet, accountability in case of AI errors is still unsolved. advances in knowledge:
This survey may be considered as a pilot-study for the development of large-scale studies to understand females's demands and concerns about AI applications in breast cancer screening.

INTRODUCTION

Breast cancer is the most common non–skin cancer and a leading cause of cancer death in North American and European females.[1] Mammography is currently the only screening test which has shown to reduce breast cancer–related mortality.[2] However, the large amount of mammography produced every year, the consequent high proportion of false-negative and false-positive results reported and the shortage of trained radiologists capable of interpreting these exams are just part of the screening management problem, leading to additional economical costs and inequalities between low- and high-income countries.[3,4] In most countries, screening programmes, generally addressed to all females aged between 40/50 and 70/80 years, are essentially based on free periodic mammography (every two years) which are read double-blinded by two expert radiologists trained specifically for this purpose.[2,5]

Artificial Intelligence (AI), which is a field of computer science dedicated to the creation of systems performing tasks that usually require human intelligence,[6] is increasingly considered a potential solution to the limits of screening mammography, and several studies are evaluating how and when it will be successfully used in clinical practice. However, it is still unknown how the use of AI in breast cancer screening programmes is going to be accepted by general population, as only few studies have investigated this aspect so far.

The aim of our survey is to evaluate what females eligible for breast cancer screening know about AI and how they perceive this innovation to be integrated with the radiologist's workflow in the cancer prevention programmes.

METHODS AND MATERIALS

We investigated through a prospective survey what participants in a breast cancer screening programme think about the introduction of AI in mammography. An anonymous questionnaire was developed collaboratively by the authors. It was written in Italian (Supplementary Figure 1), printed out and offered, from 1 May to 30 June 2021, by a radiology fellow to females waiting for their screening mammography in our referral centre for breast cancer.

Potential respondents were provided with both oral and written information about the study and participation was anonymous and optional. Participation in the survey was considered as informed consent, as reported in the questionnaire itself. The study was performed according to the Declaration of Helsinki and approved by the local ethics committee.

The underlying screening population consisted of females who underwent mammography for prevention purposes (namely, with no symptoms).

The self-reported questionnaire (Supplementary Figure 2) was composed of 11 multiple-choice questions (Qs): six questions about the knowledge of AI and the perception of its use in breast screening mammography and five questions about basic data on demographics of respondent herself (age, nationality, education title, first time of mammography screening and previous diagnosis of breast cancer). These questions were preceded by a short introduction about breast cancer and AI Supplementary Figures 3 and 4 with few scenarios where AI could be used in mammography. The questions about AI were designed to force respondents to choose a positive or negative stance to the presented scenarios by use of Likert scales (from 4 to 7 points).[7]

The questionnaire was initially tested in a pilot study over three weeks with 100 responses (results not reported). The questionnaire underwent minor revisions during this process.

Results were analysed and discussed by a collaborative group of statisticians, psychologists, sociologists, and radiologists.

Statistical analysis

To measure the level of education, we used categories taken from the Italian educational system (*i.e.,* elementary school, middle school, high school, university degree, master's degree/PhD) because these were easiest to understand for respondents.

Statistical associations of basic and demographic variables with questionnaire items were evaluated with Chi Square test or Fisher exact test, where appropriate. For ordinal categories, Mantel-Haenszel Chi square for trend was also computed. We then performed multivariable logistic regression to evaluate the association of patients' characteristics (age, nationality, educational level, and having had a previous screening mammogram) with survey's responses. We did not include in the multivariable model the information on any previous breast cancer diagnosis because of limited data. For items with more than two response categories, multinomial logistic regression was performed. Odds ratio (OR) with 95% Confidence Intervals (CI) were computed to estimate the probability of a specific response according to patients' characteristics.

We performed multiple correspondence analysis (MCA)[8] to identify individuals with similar profiles in their responses to the questions (so as to create a profiling on AI thinking) and associations between responses to questions. We performed two multiple correspondence analyses: the first one (mca1) concerning general opinion and knowledge accuracy stated and responsibility attribution in case of errors (Q1, Q2, and Q6, see); the second one (mca2) focused on the attitude to the relative role of software and radiologist in a hypothesised diagnostic process (Q3, Q4, Q5, see). For each MCA, we selected the first two dimensions (dim1 and dim2), which were combinations of the original questions' responses, and thus incorporated and summarised the information included in the questionnaire. In both MCAs, we added demographic supplementary variables to show how they changed within the dim1 and dim2. The coordinates of the dimensions according with supplementary variables assist in interpreting the cloud of individuals. An analysis of variance is made on supplementary variables, for each categorical variable, and a Student t-test is conducted to compare the average of the individuals who possess that category with the general average (V-test). The significance of V-test indicated the patients' characteristics that were mostly associated with specific perceptions and attitudes to AI (dimensions).

RESULTS

We analysed 800 questionnaires of the 870 females who took part in the survey, because 70 of them did not complete the full questionnaire and were therefore not considered.

Demographic characteristics of the sample are reported in Supplementary Table 1 and summarised in Figure 1.

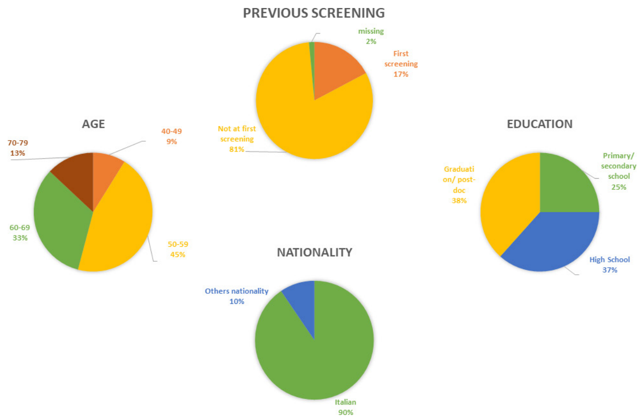


Figure 1. Baseline characteristics of the sample, including age classes, nationality (Italian, others), education level and having undergone mammograms in the past.

Females in the age group of 50–59 years (45%) were prevalent in the sample. Females with a university level education were prevalent (38%), followed by those with higher education (37%) and with only elementary education (25%). The majority of respondents were Italian (90%) and were not at their first mammography screening (81%). Only four respondents (0,5%) reported having been diagnosed with breast cancer in the past. The incidence of the above characteristics among the participants to our study was similar to and consistent with their incidence in the general screening population of our centre, with the exception of the level of education which could not be compared as it is not part of the data collected for screening patients.

Half (51%) of respondents said they had an accurate idea about the use of AI in medicine, whereas the majority of the sample expressed a positive opinion (stating that it will be useful and secure) about its use (88%). In any case, regardless of the use of AI, for 94% of females the radiologists should always provide their report on mammography, and according to 90% of respondents AI can help choosing which cases need further investigation. The vast majority of respondents (77%) agreed that AI should at least be used as a second reader.

Potential mistakes made by AI systems were in most respondents (52%) attributable to the responsibility of both the software developer and the radiologist.

Tables 1–6 show the association of the respondents' characteristics with survey's responses at multivariable analysis.

Younger females were more prone to declare not having an accurate knowledge about AI. Specifically, higher AI knowledge accuracy was reported among females aged 50–79 compared with younger females (40-49). Non-Italian respondents were associated with the belief of having a deeper awareness about AI in medicine compared to Italian respondents (OR = 1.91; 95% CI [1.10–3.33]) (Table 1). As educational title increased, a less accuracy in medical AI knowledge was reported (p-value for the trend <0.0001) (Table 1 and Supplementary Table 2). Females at their first screening reported a lower AI knowledge than females who had previous mammography screening (OR = 0.51 95% CI [0.32–0.81]) (Table 1 and Supplementary Table 4).

We reported a significant association between educational level and vision on AI (Table 2 and Supplementary Table 2): having a higher education level (*i.e.,* high school diploma or graduation) was positively associated with a positive thinking on the use of AI, although some degree of concern is observed among the more educated (graduated) who were observed to be also alarmed (Graduated or + *vs* Elementary/Middle school OR = 4.69; 95%IC [1.36–16.12]). First-time participants considered both more positive and more alarming the introduction of AI compared to the participants who already underwent screening mammography. The significance of age founded in the univariate analysis was not confirmed.

The effect of age on whether AI systems should choose which examinations to be reported was significant, showing that over

Table 1. Multivariable logistic regression on how accurate is the idea about AI in medicine

Multivariate Logistic Regression Q.1 How accurate is your idea of the use of Artificial Intelligence in medicine? (Accurate/Not Accurate)					
	Effect	**Odds Ratio**	**95% Confidence Limits**		**Pr >ChiSq**
Age	50–59 *vs* 40–49	**5.37**	**2.7**	**10.64**	**<.0001**
	60–69 *vs* 40–49	**3.68**	**1.80**	**7.55**	
	70–79 *vs* 40–49	2.11	0.96	4.63	
Nationality	Others *vs* Italian	**1.91**	**1.10**	**3.33**	**0.02**
Education title	High School *vs* Elementary/Middle school	**0.47**	**0.32**	**0.71**	**<.0001**
	Graduated or + *vs* Elementary/Middle school	**0.31**	**0.20**	**0.46**	
Mammography screening	At first screening *vs* no	**0.51**	**0.32**	**0.81**	**0.004**

Table 2. Multinominal multivariable logistic regression on participant's opinion about AI in the medical field

Multinomial Multivariate Logistic Regression Q.2 What do you think of the introduction of Artificial Intelligence systems in the medical field?						
	Effect		**Odds Ratio**	**95% Confidence Limits**		**Pr >ChiSq**
Age	50–59 *vs* 40–49	alarming *vs* useless	1.20	0.19	7.72	0.09
	50–59 *vs* 40–49	positive *vs* useless	1.26	0.37	4.28	
	60–69 *vs* 40–49	alarming *vs* useless	2.07	0.27	15.95	
	60–69 *vs* 40–49	positive *vs* useless	2.54	0.70	9.29	
	70–79 *vs* 40–49	alarming *vs* useless	1.27	0.15	10.44	
	70–79 *vs* 40–49	positive *vs* useless	0.74	0.21	2.68	
Nationality	Others *vs* Italian	alarming *vs* useless	NC	NC	NC	-
	Others *vs* Italian	positive *vs* useless	1.69	0.49	5.66	
Education title	High School *vs* Elementary/Middle school	alarming *vs* useless	2.17	0.59	8.05	**0.008**
	High School *vs* Elementary/Middle school	positive *vs* useless	**2.40**	**1.26**	**4.55**	
	Graduated or + *vs* Elementary/Middle school	alarming *vs* useless	**4.69**	**1.36**	**16.12**	
	Graduated or + *vs* Elementary/Middle school	positive *vs* useless	**2.93**	**1.48**	**5.82**	
Mammography screening	At first screening *vs* no	alarming *vs* useless	**10.99**	**1.70**	**71.12**	**0.03**
	At first screening *vs* no	positive *vs* useless	**6.66**	**1.43**	**31.05**	

50 women believed that the radiologist should always report the mammography and that AI should only chose which examinations to be reported first (Tables 3–5). Specifically, it was observed that females over 50 (with the exception of the 60–69 range) are in greater agreement than younger females in letting AI recognise mammograms that need to be reviewed by radiologists. As educational attainment increases, the reliance placed on AI in distinguishing mammograms that need medical review

Table 3. Multivariable logistic regression on the statement that the radiologist should always view the mammogram regardless of AI use

Multivariate Logistic Regression Q3. Should the radiologist always view the mammogram regardless of AI use?					
	Effect	**Odds Ratio**	**95% Confidence Limits**		**Pr >ChiSq**
Age	50–59 *vs* 40–49	2.72	0.88	8.42	0.07
	60–69 *vs* 40–49	1.81	0.56	5.82	
	70–79 *vs* 40–49	**13.96**	**1.48**	**132.20**	
Nationality	Others *vs* Italian	0.61	0.24	1.55	0.30
Education title	High School *vs* Elementary/Middle school	0.77	0.34	1.75	0.79
	Graduated or + *vs* Elementary/Middle school	0.93	0.40	2.17	
Mammography screening	At first screening *vs* no	2.05	0.70	5.97	0.19

Table 4. Multivariable logistic regression on the statement that AI systems should choose which examinations to be reported first

Multivariate Logistic Regression Q4. Should the Artificial Intelligence system choose which examinations to be reported first?					
	Effect	**Odds Ratio**	**95% Confidence Limits**		**Pr >ChiSq**
Age	50–59 *vs* 40–49	**5.68**	**2.06**	**15.71**	**<.0001**
	60–69 *vs* 40–49	2.00	0.74	5.40	
	70–79 *vs* 40–49	**7.80**	**2.12**	**28.73**	
Nationality	Others *vs* Italian	3.50	0.82	15.00	0.09
Education title	High School *vs* Elementary/Middle school	0.82	0.41	1.66	**0.03**
	Graduated or + *vs* Elementary/Middle school	**0.47**	**0.24**	**0.90**	
Mammography screening	At first screening *vs* no	**3.89**	**1.40**	**10.78**	**0.01**

decreases (Graduated or + *vs* Elementary/Middle school, OR = 0.47, 95%IC [0.24–0.90]). Those with prior screenings were more fearful of allowing the computer/AI to decide which mammograms need radiology review, as a result, those with first-time screenings place much more trust in AI (OR = 3.89, 95% CI [1.40–10.78]).

Finally, there were also significant differences in how responsibility is attributed when an error occurs (Table 6 and Supplementary Table 2-3). Non-Italian respondents indicate that neither the radiologist nor the company who developed the software is responsible for errors rather than both of them (OR = 4.26; 95% CI [2.007–8.78]). It was then observed that females with a high school level education are significantly more prone to blame both the radiologist and the software developer compared to patients with elementary/middle school level who more frequently blame only the radiologist or only the software developer. Moreover, females at their first screening were more prone to consider the radiologist responsible for possible errors in comparison with females who were not at their first mammogram (26% vs 19%), while females with previous screening more frequently attributed the responsibility to the software developer (15% vs 1%) (Supplementary Table 4).

In mca1 the two dimensions explain 33.5% of the variability in responses. With mca1, we were able to identify four groups on the basis of the coordinates of each response with respect to the two dimensions. The first group refers to patients who believe that AI is dangerous and that software developers are responsible for errors: it consists of females between 40–49 and 70–79 years old, Italians with a high level of school education at their first screening. The second group refers to patients who believe that the radiologist's role is central, and that AI is unnecessary (50–59, 70–79, primary/secondary school, not first screening). In the third group, patients think no one is responsible for errors (50–59, not Italian, high school, not first screening), while the fourth group have a positive opinion of AI (40–49 years old, high school, first screening) (Supplementary Figures 2 and 3).

In mca2 the two dimensions explain 78.3% of the variability in responses. With mca2, we were able to identify two main patterns on the use of AI as supervised or unsupervised tool. A first group of patients agrees with the use of supervised AI and is

Table 5. Multivariable logistic regression on the statement that AI should do only second reading

Multinomial Multivariate Logistic Regression Q5. Should AI do only the second reading to verify the first reading of the radiologist?					
	Effect	**Odds Ratio**	**95% Confidence Limits**		**Pr >ChiSq**
Age	50–59 *vs* 40–49	0.98	0.51	1.89	0.42
	60–69 *vs* 40–49	1.02	0.51	2.06	
	70–79 *vs* 40–49	1.61	0.70	3.71	
Nationality	Others *vs* Italian	0.68	0.40	1.18	0.17
Education title	High School *vs* Elementary/Middle school	1.15	0.74	1.80	0.26
	degree or + *vs* Elementary/ Middle school	0.83	0.54	1.28	
Mammography screening	At first her screening *vs* no	1.04	0.63	1.72	0.87

Table 6. Multinominal multivariable logistic regression on possible responsible in the event of AI error

Multinomial Multivariate Logistic Regression
Q6. Who is responsible for AI errors?

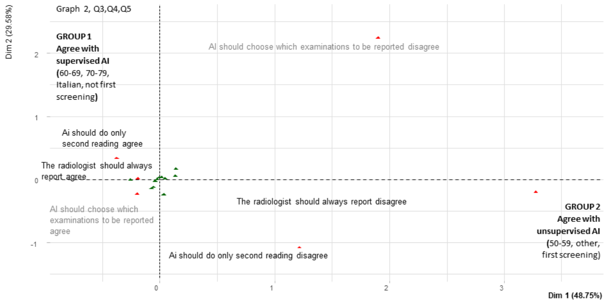
	Effect		**Odds Ratio**	**95% Confidence Limits**		**Pr >ChiSq**
Age	50–59 *vs* 40–49	The Software developer *vs* to both	0.64	0.31	1.32	0.10
	50–59 *vs* 40–49	Radiologist *vs* to both	1.04	0.46	2.36	
	50–59 *vs* 40–49	None *vs* to both	0.96	0.31	2.30	
	50–59 *vs* 40–49	Others *vs* to both	1.26	0.31	5.06	
	60–69 *vs* 40–49	The Software developer *vs* to both	0.68	0.31	1.50	
	60–69 *vs* 40–49	Radiologist *vs* to both	1.13	0.48	2.69	
	60–69 *vs* 40–49	None *vs* to both	0.55	0.16	1.92	
	60–69 *vs* 40–49	Others *vs* to both	1.29	0.29	5.85	
	70–79 *vs* 40–49	The Software developer *vs* to both	1.40	0.59	3.34	
	70–79 *vs* 40–49	Radiologist *vs* to both	0.94	0.35	2.54	
	70–79 *vs* 40–49	None *vs* to both	0.39	0.08	2.00	
	70–79 *vs* 40–49	Others *vs* to both	3.05	0.62	14.90	
Nationality	Others *vs* Italian	The Software developer *vs* to both	1.47	0.71	3.03	**0.002**
	Others *vs* Italian	Radiologist *vs* to both	1.29	0.63	2.64	
	Others *vs* Italian	None *vs* to both	**4.26**	**2.07**	**8.78**	
	Others *vs* Italian	Others *vs* to both	0.75	0.17	3.32	
	Effect		**Odds Ratio**	**95% Confidence Limits**		**Pr >ChiSq**

(Continued)

Table 6. (Continued)

Multinomial Multivariate Logistic Regression Q6. Who is responsible for AI errors?						
Education title	High School *vs* Elementary/Middle school	The Software developer *vs* to both	**0.31**	**0.18**	**0.55**	**<.0001**
	High School *vs* Elementary/Middle school	Radiologist *vs* to both	**0.46**	**0.28**	**0.76**	
	High School *vs* Elementary/Middle school	None *vs* to both	1.23	0.57	2.64	
	High School *vs* Elementary/Middle school	Others *vs* to both	1.22	0.48	3.12	
	degree or + *vs* Elementary/Middle school	The Software developer *vs* to both	0.94	0.58	1.53	
	degree or + *vs* Elementary/Middle school	Radiologist *vs* to both	0.92	0.57	1.48	
	degree or + *vs* Elementary/Middle school	None *vs* to both	0.95	0.41	2.17	
	degree or + *vs* Elementary/Middle school	Others *vs* to both	1.62	0.64	4.13	
Mammography screening	At first her screening *vs* no	The Software developer *vs* to both	1.77	0.99	3.17	0.10
	At first her screening *vs* no	Radiologist *vs* to both	0.87	0.47	1.62	
	At first her screening *vs* no	None *vs* to both	0.64	0.26	1.57	
	At first her screening *vs* no	Others *vs* to both	1.82	0.71	4.67	

Figure 2. Survey variables for questions Q3, Q4 and Q5 plotted on dimensions 1 (Dim1) and 2 (Dim2) for Multiple Correspondence Analysis. These two dimensions account for 78.3% of the variability in responses. Distance from the axis indicates the association of the variable to the dimension. In addition, two points that are close to each other have greater association with each other. A plausible interpretation of the components is provided in each quadrant, in order to visualise the characteristics of different groups of females according to the two dimensions.



characterised by females over 60 years of age who are not at their first screening. The second group agrees with the unsupervised use of AI and is characterised by females in their 50s of other nationalities at their first screening (Figure 2, Supplementary Figure 4).

DISCUSSION

Radiologists are already familiar with computer-aided detection systems, which were first introduced in the 1960s in mammography.[9] However, advances in algorithm development, combined with the ease of access to computational resources, allows AI to be applied in radiological decision-making at a higher functional level[10] achieving a sensitivity from 0.56 to 0.82 with a specificity of 0.84–0.97,[11,12] showing a cancer detection accuracy comparable to an average breast radiologist.[13] Researchers from Imperial College London and Google Health showed that DeepMind's medical AI system may outperform radiologists on identifying breast cancer from mammography,[11] paving the way for clinical trials to improve the accuracy and efficiency of breast cancer screening by AI.

AI-based algorithms may save radiologist's time scrutinising mammography screenings by detecting and characterising abnormalities on mammograms allowing radiologists to move faster through cancer-free cases and give more attention to the images with suspicious findings, while making screening cheaper and more accessible for patients.[14] Once the performance of AI in mammography is assessed in actual clinical practice (whereas current studies are based on retrospective or *in-silico* data resources, that may not be representative of real-world clinical practice),[15] this technology may prove to be the solution for accessing reliable breast cancer screening in low- and middle-income countries where cancer screening is limited due to equipment cost and the expert skill required for interpretation of mammograms, and it may help reduce existing health inequalities.[4] Moreover, AI introduction in breast cancer screening mammogram interpretation could be essential even in high-income countries to face the current (and the expected) shortage of radiologists which is increasingly putting breast cancer screening under strain.[16] Finally, while radiologists' performance tends to decrease after 70 or 80 min of reading,[17] AI never gets tired and has consistent performance. Mammography screening supported by AI can help to reduce the radiologists' overload of work, decreasing the increasing burnout rate and making radiologists less anxious during their shifts.[6,18–20]

Although such possible benefits about the introduction of new technologies in breast cancer screening, many females may not be familiar with AI. Among our respondents, we reported a fairly balanced distribution of females who claim to have a very and fairly accurate knowledge about AI in medicine and those who claim to have a fairly inaccurate and inaccurate knowledge of it. In the multivariate analysis, higher knowledge accuracy was reported among females aged 50–69 compared with younger ones and among non-Italian females compared to Italian females. Notably, we reported a less perceived accuracy in medical AI knowledge as educational title increased. This subjective evaluation of the personal knowledge about medical AI might be partially explained by Dunning-Kruger effect,[21] which described how people with limited skills or knowledge in a domain tend to overestimate their own knowledge or competence in that domain.

Our results showed that respondents who had previous mammography screening reported higher AI knowledge than females at their first screening while no role of previous diagnosis of breast cancer was found. This is consistent with results of another recent survey that showed that previous experience of mammography screening (as well as family history of breast cancer) had no association with AI knowledge.[16] Notably, Lennox-Chhugani N. et al[22] investigated females's attitudes, both current and future users of breast screening, towards the use of AI in mammogram reading. In their study, females of screening age were less likely than females under screening age to use technology apps for healthcare advice, but they were more likely to feel positive about AI used to read mammograms. This suggests that efforts to improve patients' knowledge of usefulness of technology in medicine are still needed.

Our results showed that females with higher education level (*i.e.,* high school diploma or graduation) positively evaluated the introduction of AI, although some degree of concern was observed among the more educated (graduated). This is consistent in literature, as in another survey performed by Jonmarker O. et al[23] the respondents with more than 12 years of education showed a high level of trust in AI while lower level of education was associated with a lower trust in AI.

Previous research on the attitudes toward new technologies and their perceived danger has shown that unknown hazards tend to be perceived as more risky than well-known hazards.[24,25] Our respondents declaring to have a less accurate idea about AI express a perceived lack of knowledge, and this may imply that their perceived poor familiarity with AI systems results in an increased perception of danger and tendency to overestimate their risk. Although the association of these attitudes with a higher educational achievement seems counterintuitive, literature about the impact of education on technology acceptance shows mixed results.[26] If we assume educational degree to be positively associated with the degree of scientific literacy (although we do not know the domain of the degree held by our respondents), it is worth to observe that scientific literacy does not always result in increased support for science, and it does not necessarily lead to increased scientific trust.[27]

Although most of respondents agreed with the introduction of AI in mammography, they were sceptical of having AI alone interpret their examinations. Most of respondents (particularly females over 70 years-old) expressed their strong preference for having a radiologist to confirm the diagnosis of AI and that AI never decides in total autonomy indeed.

Older females's trust in AI systems[28] seems mediated by the presence of a human radiologist. Studies showed that people who have limited knowledge about technology, such as AI systems that have not been introduced in clinical practice yet, rely on social trust to assess the risks and benefits of that technology,[28,29] trusting in experts that they perceive as trustworthy to make decisions (*e.g.,* trained radiologists that read and assess mammographic images). Also, our results are consistent with the evidence-based expectation of a positive association between perceived usefulness of a technology and trust in the technology itself.[29]

According to our respondents, AI systems in mammography should be currently limited to choose which examinations to be reported first by the radiologist and/or a second reader in the breast cancer screening. This is consistent with a recent survey study of 922 Dutch females,[3] in which authors showed that respondents did not support a fully independent use of such systems without involving a radiologist. When we summarised the attitude of our respondents towards the use of AI in assisting diagnosis, we found that older Italian females (aged over 60) with previous screenings tended to agree with the use of AI only if supervised (*i.e.,* AI should do only second reading and the radiologist should always view the exam). On the contrary, non-Italian females aged 50–59 at their first screening were more prone to accept unsupervised AI methods (*i.e.,* the radiologist does not necessarily need to view the mammography and AI should not do only second reader).

Additionally, we investigated the accountability in the case of AI-related diagnostic errors, which nowadays is still an unsolved conundrum.[3,30] In our study, both Italian and non-Italian respondents hold both software and radiologist accountable of errors, but a significant higher percentage of non-Italian than Italian respondents also did not hold anyone accountable. In our population, females with higher education more frequently did not hold anyone accountable for errors, while we did not find any correlation with age and accountability. Finally, females at first screening were more prone to consider radiologist as responsible for errors, while females with previous screening more frequently hold accountable the software developer. Overall, the public's expectations of the efficacy of screening mammography can be considered high, and diagnostic errors may have major legal consequences for the screening radiologist.[3] Delay in diagnosis of breast cancer has been reported to be a common cause for allegations of malpractice.[22] The pending clinical introduction of AI and the unanswered accountability questions underline the urgent need for policymaker to develop legal frameworks for the use of AI in screening mammography.[30,31]

Although the percentage of variance explained by mca1 is quite low (33.5%), we were able to identify possible groups of patients with different knowledge and perception of AI. The group of patients who mostly considered AI in a positive way was made by young females (40–49 years) with high school educational level at their first screening. This positive perception is somewhat decreased by a certain alarming perception for the graduated Italian youngest and oldest females (40–49 and 70–79) at their first screening: overall, they are led to consider AI as dangerous, and retained the software as the main responsible of possible errors.

Females in the middle age group (50-59) and older (70-79), with lower level of education, and with previous screening resulted more prone to consider AI useless and the radiologist the main responsible of possible errors.

Our study has some limitations because it was performed in a high-income country where screening mammograms are independently interpreted by at least two radiologists and the costs for screening mammography are completely covered by the public healthcare system. In addition, our country currently has no lack of breast cancer screening radiologists, and our Institute is considered a referral centre for breast cancer care.[32,33] The results of this study may have been different in low- and medium-income countries or rural areas, and in countries where the costs of breast cancer screening are not routinely paid by the government and where not every citizen has a health insurance. Furthermore, attitude toward AI may also be different in other countries due to cultural differences. Therefore, further research is necessary to verify whether our results are also applicable to other countries. Finally, the results are only applicable to a specific screening setting and not in clinical settings (*i.e.,* in symptomatic patients) and not necessary in another Institutes.

CONCLUSIONS

Despite recent breakthroughs in the diagnostic performance of AI in mammography, the general population seems not to support a fully independent use of computers without involving human assessment. Most of the screening females in our survey approve the introduction of AI although only as an adjunct to the radiologist's judgement.

Accountability in case of AI-related diagnostic errors in screening mammography is still an unresolved issue.

Therefore, the key question for radiology (alongside the rest of medicine) is how to convince females to embrace AI innovation in such a delicate matter as prevention of breast cancer: patients, radiologists, healthcare providers and policymakers must work together starting from the consideration of patients' values and preferences.[34,35]

Our survey may be considered as a pilot study in this topic but there is a need for the development of comprehensive and international large-scale studies to understand females' demands, expectations, and concerns when it comes to AI applications in breast cancer screening. The results will be essential to develop a successful and patient-centred medical AI innovation pathway.

FUNDING

Open access funding provided by BIBLIOSAN.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; **71**: 209–49. https://doi.org/10.3322/caac.21660
2. NCCN.org. Breast Cancer Screening and Diagnosis Version 3.2018. In: *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines)*. ; 2018.
3. Ongena YP, Yakar D, Haan M, Kwee TC. Artificial intelligence in screening mammography: a population survey of women's preferences. *J Am Coll Radiol* 2021; **18**: S1546-1440(20)30989-3: 79–86: . https://doi.org/10.1016/j.jacr.2020.09.042
4. Peek ME, Han JH. Disparities in screening mammography. current status, interventions and implications. *J Gen Intern Med* 2004; **19**: 184–94. https://doi.org/10.1111/j.1525-1497.2004.30254.x
5. Tyagi NK, Dhesy-Thind S. Clinical practice guidelines in breast cancer. *Curr Oncol* 2018; **25**: S151–60. https://doi.org/10.3747/co.25.3729
6. Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? radiologists again at the forefront of innovation in medicine. *Eur Radiol Exp* 2018; **2**: 35. https://doi.org/10.1186/s41747-018-0061-6
7. Xiao Y, Liu H, Li H. Integration of the forced-choice questionnaire and the likert scale: a simulation study. *Front Psychol* 2017; **8**: 806. https://doi.org/10.3389/fpsyg.2017.00806
8. Sourial N, Wolfson C, Zhu B, Quail J, Fletcher J, Karunananthan S, et al. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. *J Clin Epidemiol* 2010; **63**: 638–46. https://doi.org/10.1016/j.jclinepi.2009.08.008
9. Lee J-G, Jun S, Cho Y-W, Lee H, Kim GB, Seo JB, et al. Deep learning in medical imaging: general overview. *Korean J Radiol* 2017; **18**: 570–84. https://doi.org/10.3348/kjr.2017.18.4.570
10. Pesapane F. How scientific mobility can help current and future radiology research: a radiology trainee's perspective. *Insights Imaging* 2019; **10**(1): 85. https://doi.org/10.1186/s13244-019-0773-z
11. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature* 2020; **577**: 89–94. https://doi.org/10.1038/s41586-019-1799-6
12. Salim M, Wåhlin E, Dembrower K, Azavedo E, Foukakis T, Liu Y, et al. External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. *JAMA Oncol* 2020; **6**: 1581–88. https://doi.org/10.1001/jamaoncol.2020.3321
13. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, et al. Stand-Alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. *J Natl Cancer Inst* 2019; **111**: 916–22. https://doi.org/10.1093/jnci/djy222
14. Conant EF, Toledano AY, Periaswamy S, Fotin SV, Go J, Boatsman JE, et al. Improving accuracy and efficiency with concurrent use of artificial intelligence for digital breast tomosynthesis. *Radiol Artif Intell* 2019; **1**: e180096. https://doi.org/10.1148/ryai.2019180096
15. Lehman CD, Topol EJ. Readiness for mammography and artificial intelligence. *Lancet* 2021; **398**: : S0140-6736(21)02484-3. https://doi.org/10.1016/S0140-6736(21)02484-3
16. de Vries CF, Morrissey BE, Duggan D, Staff RT, Lip G. Screening participants' attitudes to the introduction of artificial intelligence in breast screening. *J Med Screen* 2021; **28**: 221–22. https://doi.org/10.1177/09691413211001405
17. Bernstein MH, Baird GL, Lourenco AP. Digital breast tomosynthesis and digital mammography recall and false-positive rates by time of day and reader experience. *Radiology* 2022; **303**: 210318: 63–68: . https://doi.org/10.1148/radiol.210318
18. Pesapane F, Abbate F, Bozzini A, Dominelli V, Farina M, Ferrari F, et al. What breast radiologists have learned from the COVID-19 pandemic. *J Public Health Emerg* 2021; **6**: 7. https://doi.org/10.21037/jphe-21-84
19. Lehman CD, Lamb LR, D'Alessandro HA. Mitigating the impact of coronavirus disease (COVID-19) vaccinations on patients undergoing breast imaging examinations: a pragmatic approach. *American Journal of Roentgenology* 2021; **217**: 584–86. https://doi.org/10.2214/AJR.21.25688
20. Chen RC, Haynes K, Du S, Barron J, Katz AJ. Association of cancer screening deficit in the United States with the COVID-19 pandemic. *JAMA Oncol* 2021; **7**: 878–84. https://doi.org/10.1001/jamaoncol.2021.0884
21. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* 1999; **77**: 1121–34. https://doi.org/10.1037//0022-3514.77.6.1121
22. Lennox-Chhugani N, Chen Y, Pearson V, Trzcinski B, James J. 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**(1): e100293. https://doi.org/10.1136/bmjhci-2020-100293
23. Jonmarker O, Strand F, Brandberg Y, Lindholm P. 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. https://doi.org/10.1177/2058460119880315
24. Slovic P. Perception of risk. *Science* 1987; **236**: 280–85. https://doi.org/10.1126/science.3563507
25. Siegrist M, Keller C, Kiers HAL. A new look at the psychometric paradigm of perception of hazards. *Risk Anal* 2005; **25**: 211–22. https://doi.org/10.1111/j.0272-4332.2005.00580.x

26. ConnorMS. Factors influencing people's acceptance of gene technology: the role of knowledge. *Health Expectations, Naturalness, and Social Trust Science Communication* 2010. https://doi.org/10.1177/1075547009358919
27. Roberts MR, Reid G, Schroeder M, Norris SP. Causal or spurious? the relationship of knowledge and attitudes to trust in science and technology. *Public Underst Sci* 2013; **22**: 624–41. https://doi.org/10.1177/0963662511420511
28. Siegrist M, Cvetkovich G. Perception of hazards: the role of social trust and knowledge. *Risk Anal* 2000; **20**: 713–19. https://doi.org/10.1111/0272-4332.205064
29. Steinke F, Bading N, Fritsch T, Simonsen S. Factors influencing trust in ambient assisted living technology: a scenario-based analysis. *Gerontechnology* 2014; **12**(2). https://doi.org/10.4017/gt.2013.12.2.002.00
30. Pesapane F, Volonté C, Codari M, Sardanelli F. Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. *Insights Imaging* 2018; **9**: 745–53. https://doi.org/10.1007/s13244-018-0645-y
31. Pesapane F, Bracchi DA, Mulligan JF, Linnikov A, Maslennikov O, Lanzavecchia MB, et al. Legal and regulatory framework for AI solutions in healthcare in Eu, us, China, and Russia: new scenarios after a pandemic. *Radiation* 2021; **1**: 261–76. https://doi.org/10.3390/radiation1040022
32. Newsweek. World's Best Specialized Hospitals 2022. 2022. Available from: https://www.newsweek.com/worlds-best-specialized-hospitals-2022/oncology2022
33. Pesapane F, Penco S, Rotili A, Nicosia L, Bozzini A, Trentin C, et al. How we provided appropriate breast imaging practices in the epicentre of the COVID-19 outbreak in Italy. *Br J Radiol* 2020; **93**(1114): 20200679. https://doi.org/10.1259/bjr.20200679
34. Kruskal JB, Berkowitz S, Geis JR, Kim W, Nagy P, Dreyer K. Big data and machine learning-strategies for driving this bus: a summary of the 2016 intersociety summer conference. *J Am Coll Radiol* 2017; **14**: S1546-1440(17)30199-0: 811–17: . https://doi.org/10.1016/j.jacr.2017.02.019
35. Chichua M, Brivio E, Mazzoni D, Pravettoni G. Shared decision-making and the lessons learned about decision regret in cancer patients. *Support Care Cancer* 2022; **30**: 4587–90. https://doi.org/10.1007/s00520-021-06725-5