Artificial Intelligence in Screening Mammography: A Population Survey of Women's Preferences

Yfke P. Ongena, PhD[a], Derya Yakar, MD, PhD[b], Marieke Haan, PhD[c], Thomas C. Kwee, MD, PhD[d]

Abstract

Objective: To investigate the general population's view on the use of artificial intelligence (AI) for the diagnostic interpretation of screening mammograms.

Methods: Dutch women aged 16 to 75 years were surveyed using the Longitudinal Internet Studies for the Social sciences panel, representative for the Dutch population. Attitude toward AI in mammography screening was measured by means of five items: necessity of a human check; AI as a selector for second reading; AI as a second reader; developer is responsible for error; and radiologist is responsible for error.

Results: Of the 922 participants included, 77.8% agreed with the necessity of a human check, whereas the item AI as a selector for a second reading was more heterogeneously answered, with 41.7% disagreement, 31.5% agreement, and 26.9% responding with "neither agree nor disagree." The item AI as a second reader was mostly responded with "neither agree nor disagree" (37.1%) and "agree" (37.6%), whereas the two last items on developer's and radiologist' responsibilities were mostly answered with "neither agree nor disagree" (44.6% and 39.2%, respectively).

Discussion: Despite recent breakthroughs in the diagnostic performance of AI algorithms for the interpretation of screening mammograms, the general population currently does not support a fully independent use of such systems without involving a radiologist. The combination of a radiologist as a first reader and an AI system as a second reader in a breast cancer screening program finds most support at present. Accountability in case of AI-related diagnostic errors in screening mammography is still an unresolved conundrum.

Key Words: Artificial intelligence, breast cancer, mammography, mass screening, surveys and questionnaires

J Am Coll Radiol 2021;18:79-86.

[a]Assistant Professor and Chair Faculty Council Faculty of Arts, Center of Language and Cognition, University of Groningen, Groningen, The Netherlands.

[b]Radiologist, Department of Radiology, Medical Imaging Center, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

[c]Assistant Professor, Department of Sociology, University of Groningen, Groningen, The Netherlands.

[d]Vice-Chair, Department of Radiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

Corresponding author and reprints: Yfke P. Ongena, University of Groningen, Centre for Language and Cognition, Discourse and Communication Group, Oude Kijk in 't Jatstraat 26, 9712 EK Groningen, The Netherlands; e-mail: y.p.ongena@rug.nl.

The authors state that they have no conflict of interest related to the material discussed in this article. Dr Ongena, Dr Yakar, Dr Haan, and Dr Kwee are nonpartner, non-partnership track employees.

INTRODUCTION

Breast cancer is the most common malignancy in women, and one of the three most common cancers worldwide [1]. Early breast cancer is considered potentially curable [1], and a substantial reduction in mortality from breast cancer can be achieved with screening mammography [2]. However, there is a shortage of radiologists to interpret screening mammograms in many regions such as in the United Kingdom and rural areas in the United States [3-5]. Furthermore, the interpretation of screening mammograms by radiologists is not perfect [6]. False-positive mammography results are common (reported estimates varying between 65.2 and 121.2 per 1,000 women, with higher rates at younger age), and the proportion of false-negatives is not negligible

1546-1440/20/$36.00 ■ https://doi.org/10.1016/j.jacr.2020.09.042

either (reported estimates varying between 1.0 and 1.3 per 1,000, without significant differences by age) [6].

Artificial intelligence (AI) algorithms, particularly deep learning, have demonstrated remarkable progress in image-recognition tasks [7]. Recent studies have shown that cancer detection by radiologists can be improved when using an AI system for support [8,9] and that AI systems can achieve a diagnostic performance that is comparable to that of a breast radiologist in evaluating screening mammography examinations [9,10]. In an even more recent study, it was reported that an AI system was capable of surpassing radiologists in breast cancer prediction on screening mammography [11].

Based on its diagnostic potential [8,9,11], the question is perhaps not if but when and how AI systems will be used in the standard practice of screening mammography. Importantly, this clinical implementation not only depends on the diagnostic performance of an AI system, but ethical, legal, and societal issues will also have to be taken into account [12]. The voice of the population who will undergo AI-based diagnostic tests is crucial in this context, because it is a determining factor for the boundaries within which an AI system is allowed to operate. Moreover, the success of a breast cancer screening program depends on the willingness of subjects to participate, and this willingness may be affected if AI systems are used without taking into account the populations' wishes, concerns, and objections. Therefore, it is important to determine the view of the population on the use of AI in mammography screening programs.

It is currently unknown how the introduction of AI in mammography screening programs would be received by the general population and which variables are associated with a more favorable approach toward AI implementation in this setting. We hypothesized that the majority of the population has a positive view about the use of AI in screening mammography, because evidence about its diagnostic potential has also reached mainstream media [13]. We also anticipated younger and higher educated subjects to have more trust in AI systems than older generations because of more affinity with new technology [14]. Furthermore, we expected a difference in attitudes to be related to experience with mammography screening, and we hypothesized the attitude toward AI in general to be positively associated with the attitude toward AI in mammography screening.

The purpose of this study was to investigate the general population's view on the use of AI for the diagnostic interpretation of screening mammograms.

METHODS

Study Design and Subjects

We used data from the LISS (Longitudinal Internet Studies for the Social sciences) panel, a nationally representative household panel study for people aged 16 years and above in the Netherlands. The LISS panel uses an informed consent procedure that ensured double consent, via a reply card and an Internet login; see Scherpenzeel and Das [15] for details. Ethical approval for the procedures in the LISS Panel was given by the board of overseers [16]. All data are available from the LISS panel data archive [17].

In the LISS panel, the same pool of respondents is asked different questions at each separate data collection (ie, a wave). We combined data from two waves: one fielded in April 2020 including items on the attitudes toward AI in mammography screening and general medicine, and one fielded in December 2018 including items on health characteristics (eg, experience with mammography screening and diagnosis of cancer). In the Netherlands, all women aged 50 to 75 years are invited biennially to undergo screening mammography. All screening mammograms are independently interpreted by two radiologists, and a third radiologist is involved in case of discrepancies between the first two readers. In this study, only responses from female respondents aged below 75 years were included. Women aged 16 to 49 years were also included because they constitute the future target population for mammography screening. All statistical analyses were conducted in R version 3.6.1 [18].

Measurement and Analysis Attitude Toward AI in Mammography Screening

Attitude toward AI in mammography screening was measured by means of five items (Table 1), using a 5-point agree-disagree scale. Because the five single items on attitude toward AI in mammography screening were measured on an ordinal scale and the data did not meet the requirement of proportional odds for an ordinal logistic regression, we conducted multinomial logistic regression analyses. For a more concise presentation of multinomial regression results, and to account for an experimental manipulation of end-point versus fully labeled scales (see e-only appendix), we converted the 5-point scale into a 3-point scale, combining the options "strongly disagree" and "disagree" and combining the options "strongly agree" and "agree."

Measurement Predictor Variables

To measure the level of education, we used the LISS panel item of highest earned degree and categories taken from the Dutch educational system (easiest to understand for respondents), which were converted into international categories, ranging from lower education (ie, primary education or lower vocational education), high school (pre-university education or mediate vocational education), college (university or higher vocational education), and other (no degree, or degree not included among response options).

Table 1. Items attitude toward artificial intelligence in mammography screening

Variable Name	Item Wording
1. Necessity of a human check	When a computer examines a mammogram, I think it is necessary that a radiologist also takes a look at the study.
2. AI as a selector for second reading	The computer should decide for which mammograms it is required to have a radiologist as a second reader.
3. AI as a second reader	Instead of a second radiologist, the computer should be used to check the judgment of the first radiologist.
4. Developer is responsible for error	When a computer gives the wrong result, the developer of the computer program is responsible.
5. Radiologist is responsible for error	When a computer gives the wrong result, the radiologist is responsible.

AI = artificial intelligence.

General attitude toward AI was measured using items developed from a scale by selecting three previously validated factors [14] (see e-only appendix for descriptive statistics in the current sample), namely, Trust and Accountability (trust in AI in taking over diagnostic interpretation tasks of the radiologist, both with regard to accuracy, communication, and confidentiality), Personal Interaction (preference of personal interaction over AI-based communication), Efficiency (belief in whether AI will improve diagnostic workflow), and a newly developed scale measuring the attitude toward AI in general medicine (Table 2) (Cronbach's $\alpha = 0.87$, mean = 3.6, SD = 0.78; a higher score means a more positive attitude toward AI).

RESULTS

Subjects

For the April 2020 wave, 3,117 LISS panel members were contacted (for sampling and recruitment details, see Scherpenzeel and Das [15]), and 2,411 completed the full questionnaire (77.4% response rate). For the December 2018 wave, 6,466 panel members were contacted, and 5,455 completed the full questionnaire (84.4% response rate). Within the combined sample, a total of 922 female participants (70.1% participated in both waves). Demographic characteristics of the sample are summarized in Table 3.

AI in Mammography Items

A descriptive analysis of the five AI in mammography items is shown in Figure 1. The item "Necessity of a human check" showed that a vast majority agreed with the statement (77.8% agreed or strongly agreed), whereas the item "AI as a selector for a second reading" showed a much more diverse distribution of respondents, with 41.7% disagreement, 31.5% agreement, and 26.9% responding with "neither agree nor disagree." The item "AI as a second reader" was mostly responded with "neither agree nor disagree" (37.1%) and "agree" (37.6%), whereas the two last items on developer's and radiologist' responsibilities were mostly answered with "neither agree nor disagree" (44.6% and 39.2%, respectively).

Results Multivariable Analyses

Table 4 summarizes the results of the multinomial logistic regression, including the relative risk ratio for agree versus disagree and for neutral versus disagree. The increase of the relative risk ratio should be interpreted for a 1-unit increase of the variable for continuous variables (age and the factors Trust and Accountability, Personal Interaction, Efficiency, General Attitude Toward AI). So, for example, for "Necessity of a human check," the relative risk ratio for a 1-unit increase in the factor Personal Interaction is 2.53 for agreeing with the statement versus disagreeing, and 0.34 for neither agreeing nor disagreeing versus disagreeing. For categorical variables, such as education, being in the screening population and experience with mammography, the relative risk ratio should be interpreted against the reference category mentioned. For example, for "Necessity

Table 2. Newly developed scale to measure attitude toward artificial intelligence in general medicine

Item number	Item Wording
A111*	I find using computers to perform medical tasks a bad idea.
A112	I find it safe to use computers to perform medical tasks.
A113	I find it helpful to use computers to perform medical tasks.
A114	I find it useful to use computers to perform medical tasks.

*Item marked was reverse scored to match the direction of the other items.

Table 3. Demographic characteristics of the sample

Variable	n (%)
Age (y)	
Below 50	513 (55.6%)
Between 50 and 75	409(44.4%)
Previous experience with mammography screening	
Yes (total sample)	443 (48.1%)
Yes (within screening age)	323 (79.0%)
No (total sample)	479 (51.9%)
No (within screening age)	86 (21.0%)
Diagnosed with cancer* within the past 3 years	
Yes	17 (1.8%)
No	905 (98.2%)
Level of education	
Low (elementary school)	234 (25.3%)
High school or lower vocational	341 (36.9%)
College (BA, MA, MSc, MD, or PhD)	347 (37.7%)
Immigration background†	
Dutch	683 (74.08%)
Western immigration background	75 (8.13%)
Non-Western immigration background	49 (5.31%)
Unknown	115 (12.47%)

*Nonmelanoma skin cancer was excluded since the question only asked respondents about cancer.

†Immigration background was asked in terms of country of birth of the respondent and both parents. First- and second-generation immigrants are collapsed in this table, and countries were recoded into Western and non-Western countries. Immigrants from Western countries include Europe (Turkey excluded), North America, Oceania (including Australia and New Zealand), Japan and Indonesia, the latter because of main immigration from former Dutch colonies. Non-Western immigrants in the Netherlands concern mostly Turkey, Morocco, Surinam, and the Dutch Antilles.

of a human check," in education the relative risk ratio for switching from "low" to "college education" is 0.32 for agreeing versus disagreeing, and the relative risk ratio for switching from "low" to "high school" education is 0.36 for agreeing with the statement versus disagreeing.

Necessity of a Human Check

For the item "When a computer examines a mammogram I think it is necessary that a radiologist also takes a look at the study," the multinomial logistic regression showed significant effects for the factors Personal Interaction, Efficiency, and the respondents' level of education (Table 4 and Fig. 1a, b, and c in the e-only appendix). These findings showed that respondents who find a human check of mammograms necessary tend to be persons who find personal interaction in discussing results of a scan important, find AI less efficient, and are lower educated. Respondents who were neutral with respect to the necessity of a human check tend to be persons who find personal interaction in discussing results of a scan less important.

AI as a Selector for Second Reading

The multinomial logistic regression of the item "The computer should decide for which mammograms it is required to have a radiologist as a second reader" showed significant effects for the factors Personal Interaction, Efficiency, and General Attitude (Table 4 and Fig. 2a, b, and c in the e-only appendix). These findings showed that respondents who tend to agree with or are neutral toward using AI as a selector for second reading tend to be persons who find personal interaction in discussing results of a scan less important and find AI more efficient.

AI as a Second Reader

The item "Instead of a second radiologist, the computer should be used to check the judgment of the first radiologist" showed significant effects for the factors Personal Interaction, Efficiency, and General Attitude (Table 4 and Fig. 3a, b, and c in the e-only appendix). Respondents who favored using AI as a second reader tend to be persons who find personal interaction in discussing results of a scan less important, find AI more efficient, and have a more positive attitude toward AI in general medicine. Respondents with a neutral opinion toward using AI as a second reader tend to be persons who find personal interaction in discussing results of a scan less important.

Developer Is Responsible for Error

The item "When a computer gives the wrong result, the developer of the computer program is responsible" showed significant effects for the factor General Attitude and the respondents' age and level of education (Table 4 and Fig. 4a and b in the e-only appendix). Respondents who agreed that the developer is responsible tend to be respondents with a negative attitude toward AI in general medicine and tend to have lower education.

Radiologist Is Responsible for Error

The item "When a computer gives the wrong result, the radiologist is responsible" showed significant effects for the factor Personal Interaction and for whether respondents' age between 50 and 75 years (matching that of the target

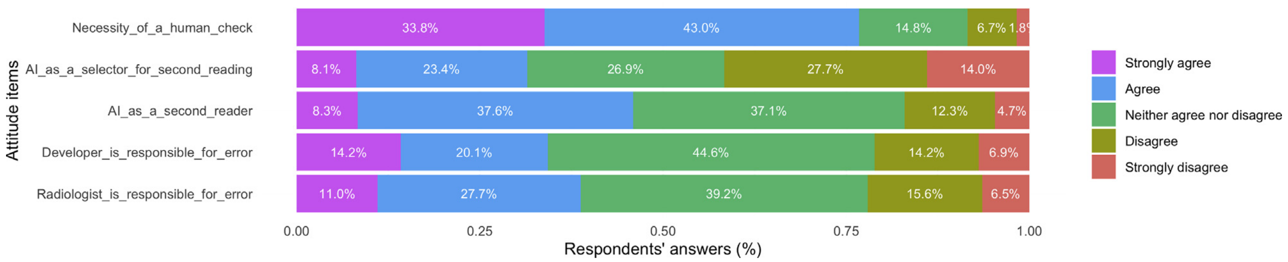

Fig 1. Descriptive artificial intelligence (AI) in mammography items.

screening population) or not (Table 4 and Fig. 5a in the e-only appendix). Findings in this case were only significant for respondents who chose the option neutral with regard to the responsibility of the radiologist; they tend to be respondents who find personal interaction less important, and they tend to be below 50 years of age.

DISCUSSION

In this study we presented the findings of a survey conducted in the LISS panel, representative for the Dutch population. Our results show that most women (77.8%) do not support a fully independent use of AI-based diagnostics in screening mammography without involvement of a radiologist. These findings are somewhat surprising, because recent work has shown AI to be of great promise for mammography screening, even outperforming radiologists [11], as was also published in the lay press [13,19]. Therefore, from the population's perspective, it is too premature to leave the interpretation of screening mammograms completely up to independently operating AI algorithms. Participants in the present study were also asked if AI could play a role in selecting cases that require second reading or as an independent second reader, because such strategies may also decrease radiologists' workload. Overall, respondents were slightly more optimistic about the use of AI as such. However, a considerable proportion (41.7%) still opposed the idea of using AI as a tool to select patients for second reading, which indicates the importance women attach to having a second reading in any case. On the other hand, a much smaller proportion (17.0%) explicitly objected against using AI as an actual second reader. Therefore, the combination of a radiologist as a first reader and an AI

Table 4. Relative risk ratio for selecting agree and neutral versus disagree on five mammography items (n = 922)

	Necessity of a Human Check		AI as a Selector for Second Reading		AI as a Second Reader		Developer Is Responsible for Error		Radiologist Is Responsible for Error	
Ref: disagree	Agree	Neutral	Agree	Neutral	Agree	Neutral	Agree	Neutral	Agree	Neutral
Age	1.01	0.93	0.99	0.99	0.98	1.01	1.02	1.03†	0.99	1.00
Education (ref: low)										
High school	0.36*	0.59	0.82	1.07	0.69	0.63	0.79	0.74	1.13	1.04
College	0.32†	0.47	1.07	1.13	1.39	1.14	0.74†	0.57*	1.06	0.77
Trust and Accountability	0.99	0.83	0.94	0.59	0.99	0.35	1.08	0.65	0.87	0.52
Personal Interaction	2.53‡	0.34‡	0.62†	0.29‡	1.16	0.64*	1.35	0.71	0.82	0.44‡
Efficiency	0.49‡	0.77	3.48‡	1.67‡	5.32‡	2.22‡	1.07	1.04	1.00	0.83
General Attitude Toward AI	0.75	0.78	1.42*	0.96	1.68†	0.97	0.71*	0.70*	0.83	0.98
In screening population (ref: no)	0.93	1.34	0.87	0.78	1.37	0.70	0.82	0.34†	0.78	0.38*
Mammography experience (ref: no)	0.68	0.64	1.15	1.35	0.75	0.66	0.97	1.24	1.00	1.59

AI = artificial intelligence; ref = reference.
$^{*}P < .05$ (two-tailed).
$^{\dagger}P < .01$ (two-tailed).
$^{\ddagger}P < .001$ (two-tailed).

system as a second reader seems to be the most acceptable approach to the population at present, although still not fully embraced by the entire population. Of interest, our multivariable analyses show that respondents who find personal interaction in discussing results of a radiologic examination important, who think AI to be less efficient, who have a more negative attitude toward AI in medicine, and who are lower educated to be particularly less supportive about AI taking over the tasks of a radiologist in screening mammography. Improved information supply and education about the development, possibilities, and limitations of AI algorithms in screening mammography may potentially overcome some of the perceived obstacles and increase acceptance of this new technique in clinical practice.

Most respondents (44.6%) were ambiguous as to whether the developer of the AI system should be held responsible when diagnostic errors are made. Similarly, a large proportion of the population (39.2%) had no clear opinion when asked about the responsibility of the radiologist in case diagnostic errors are made by the AI system. Our multivariable analysis shows that respondents who have a negative attitude toward AI in medicine and who have a lower education more frequently believe the AI developer to be responsible for diagnostic errors than respondents with the opposite characteristics. There were no variables that were significantly associated with an outspoken opinion (ie, agree or disagree) about the responsibility of the radiologist in case of AI-related diagnostic errors in screening mammography. The implications of these multivariable results on the public's opinion about accountability are not completely clear. Nevertheless, 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 [20]. Delay in diagnosis of breast cancer has been reported to be a common cause for allegations of malpractice [21]. The pending clinical introduction of AI and the unanswered accountability questions underline the urgent need for governing bodies and lawmakers to develop legal frameworks for the use of AI in screening mammography [22].

Previous research has shown that patients are generally not overly optimistic about AI systems taking over diagnostic interpretations that are currently performed by radiologists [23]. In a survey study by Ongena et al [14] that included 155 subjects who underwent CT, MRI, or conventional radiography in an outpatient setting, patients indicated a general need to be well and completely informed on all aspects of the diagnostic process, both when it comes to how and which of their imaging data are acquired and processed [14]. A strong need of patients to keep human interaction also emerged, particularly when communicating the results of their imaging examinations [14]. Although these data apply to patients' views on AI in general radiology, they resonate with the findings of the present study [23]. Other survey studies on the population's view on AI in radiology are currently lacking. Research on this topic outside of the field of radiology has also been very limited so far. However, one recent qualitative survey study by Nelson et al [24] aimed to explore how patients perceive the use of AI for skin cancer screening [24]. This clinical scenario has some parallel with mammography screening. Their study included 48 patients who visited a general dermatology clinic [24]. Their key finding was that 75% of patients would recommend the use of AI for skin cancer screening to friends and family members [24]. However, 94% of patients expressed the importance of symbiosis between humans and AI [24]. Human-computer symbiosis was referred to as a form of teamwork in which humans provide strategic input while computers provide depth of analysis [24,25]. Rather than replacing a physician, patients envisioned AI referring to a physician and providing a second opinion for a physician [24]. These results are completely in line with those of the present study. Nelson et al [24] also asked participants to identify entities responsible for AI accuracy. Patients most often named the technology company (52%) and the physician (42%), followed by the collective (25%) and the health care institution (23%) [24]. These heterogeneous answers on accountability also match those of the present study.

The present study had some limitations. First, it was performed in a Western European country where screening mammography is offered to all women aged 50 to 75 years, and all screening mammograms are independently interpreted by at least two radiologists. The costs for screening mammography are completely covered by the Dutch government, and any necessary subsequent investigations are covered by the patient's health insurance (all citizens of the Netherlands are obliged to have health insurance). In addition, our country has no lack of breast cancer screening radiologists, and the health care system is considered as one of the best in Europe [26]. The results of this study may have been different in more resource-constrained countries or regions, and in countries such as the United States where the costs of breast cancer screening are not routinely paid by the government unless insured under a governmental program such as Medicare and where not every citizen has a health insurance. Attitude toward AI may also be different in other countries because of cultural differences. Therefore, further research is necessary to verify whether are results are also applicable to other countries. Second, the results are only applicable to a screening setting and not to clinical settings in which mammography serves another purpose.

Third, before conducting the survey, we did not inform the participants about older computer-aided detection (CAD) systems that have already been used in clinical practice. Doing so may have influenced their sentiment about the use of newer AI systems in mammography screening. However, older CAD systems have a relatively poor diagnostic performance (particularly suffering from a considerable number of false-positives [27]) and were not used in a completely autonomous way, whereas recently developed AI systems do have that potential. Our survey items solely focused on the autonomous use of AI in mammography screening. Comparing the projected autonomous use of AI to the previous situation in which older CAD systems were used as an adjunct to the radiologist without autonomous use would be less meaningful. Moreover, informing participants about these outdated CAD systems and their use in clinical practice would be confusing and may have undermined the validity of our survey.

In conclusion, despite recent breakthroughs in the diagnostic performance of AI algorithms for the interpretation of screening mammograms, the general population currently does not support a fully independent use of such systems without involving a radiologist. The combination of a radiologist as a first reader and an AI system as a second reader in a breast cancer screening program finds most support at present. Accountability in case of AI-related diagnostic errors in screening mammography is still an unresolved conundrum.

TAKE-HOME POINTS

- There is a shortage of radiologists to interpret screening mammograms in many regions.
- Despite recent breakthroughs in the diagnostic performance of AI algorithms for the interpretation of screening mammograms, the general population currently does not support a fully independent use of such systems without the involvement of a radiologist.
- The combination of a radiologist as a first reader and an AI system as a second reader in a breast cancer screening program finds most support at present.
- Accountability in case of AI-related diagnostic errors in screening mammography is still an unresolved conundrum.

ACKNOWLEDGMENTS

This research is part of a project funded by a grant of the Open Data Infrastructure for Social Science and Economic Innovations (ODISSEI) in the Netherlands.

ADDITIONAL RESOURCES

Additional resources can be found online at: https://doi.org/10.1016/j.jacr.2020.09.042.

REFERENCES

1. Harbeck N, Gnant M. Breast cancer. Lancet 2017;389:1134-50.
2. Dibden A, Offman J, Duffy SW, Gabe R. Worldwide review and meta-analysis of cohort studies measuring the effect of mammography screening programmes on incidence-based breast cancer mortality. Cancers (Basel) 2020;12:976.
3. Torres-Mejía G, Smith RA, Carranza-Flores Mde L, et al. Radiographers supporting radiologists in the interpretation of screening mammography: a viable strategy to meet the shortage in the number of radiologists. BMC Cancer 2015;15:410.
4. Rimmer A. Radiologist shortage leaves patient care at risk, warns royal college. BMJ 2017;359:j4683.
5. Wing P, Langelier MH. Workforce shortages in breast imaging: impact on mammography utilization. AJR Am J Roentgenol 2009;192:370-8.
6. Nelson HD, O'Meara ES, Kerlikowske K, Balch S, Miglioretti D. Factors associated with rates of false-positive and false-negative results from digital mammography screening: an analysis of registry data. Ann Intern Med 2016;164:226-35.
7. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018;18:500-10.
8. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, et al. Detection of breast cancer with mammography: effect of an artificial intelligence support system. Radiology 2019;290:305-14.
9. Wu N, Phang J, Park J, et al. Deep neural networks improve radiologists' performance in breast cancer screening. IEEE Trans Med Imaging 2020;39:1184-94.
10. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, et al. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 Radiologists. J Natl Cancer Inst 2019;111:916-22.
11. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. Nature 2020;577:89-94.
12. Geis JR, Brady A, Wu CC, et al. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. Insights Imaging 2019;10:101.
13. Grady D. A.I. is learning to read mammograms. Available at: https://www.nytimes.com/2020/01/01/health/breast-cancer-mammogram-artificial-intelligence.html. Published 2020. Accessed October 12, 2020.
14. Ongena YP, Haan M, Yakar D, Kwee TC. Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. Eur Radiol 2020;30:1033-40.
15. Scherpenzeel AC, Das M. "True" longitudinal and probability-based internet panels: evidence from the Netherlands. In: Das M, Ester P, Kaczmirek L, eds. Social and behavioral research and the internet: advances in applied methods and research strategies. New York, NY: Routledge; 2010:77-104.
16. LISS Panel. Board of overseers. Available at: https://www.lissdata.nl/organization/board-overseers. Accessed October 12, 2020.
17. Liss Panel. Data archive. Available at: https://www.dataarchive.lissdata.nl/. Accessed October 12, 2020.
18. R. The R project for statistical computing. Available at: https://www.r-project.org/. Accessed October 12, 2020.
19. Korteweg N. De software die slimmer is dan de dokter. Available at: https://www.nrc.nl/nieuws/2020/01/31/de-software-die-slimmer-is-dan-de-dokter-a3988893. Published 2020. Accessed October 12, 2020.

20. van Breest Smallenburg V, Setz-Pels W, Groenewoud JH, et al. Malpractice claims following screening mammography in The Netherlands. Int J Cancer 2012;131:1360-6.
21. Ward CJ, Green VL. Risk management and medico-legal issues in breast cancer. Clin Obstet Gynecol 2016;59:439-46.
22. Carter SM, Rogers W, Win KT, Frazer H, Richards B, Houssami N. The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. Breast 2020;49:25-32.
23. Haan M, Ongena YP, Hommes S, Kwee TC, Yakar D. A qualitative study to understand patient perspective on the use of artificial intelligence in radiology. J Am Coll Radiol 2019;16:1416-9.
24. Nelson CA, Pérez-Chada LM, Creadore A, et al. Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study. JAMA Dermatol 2020;156:1-12.
25. Licklider JCR. Man-computer symbiosis. IRE Trans Hum Factors Electron 1960;HFE-1:4-11.
26. Björnberg A, Phang AY. Euro Health Consumer Index 2018. Available at: https://healthpowerhouse.com/publications/. Accessed October 12, 2020.
27. Kim SJ, Moon WK, Seong MH, Cho N, Chang JM. Computer-aided detection in digital mammography: false-positive marks and their reproducibility in negative mammograms. Acta Radiol 2009;50:999-1004.