

Original Paper

Patient Reactions to Artificial Intelligence—Clinician Discrepancies: Web-Based Randomized Experiment

Farrah Madanay¹, PhD; Laura S O'Donohue², MD; Brian J Zikmund-Fisher³, PhD

¹Center for Bioethics and Social Sciences in Medicine, University of Michigan—Ann Arbor, Ann Arbor, MI, United States

²Department of Radiology, University of Michigan Medicine, University of Michigan—Ann Arbor, Ann Arbor, MI, United States

³Health Behavior and Health Equity, Internal Medicine, Center for Bioethics and Social Sciences in Medicine, University of Michigan—Ann Arbor, Ann Arbor, MI, United States

Corresponding Author:

Farrah Madanay, PhD

Center for Bioethics and Social Sciences in Medicine

University of Michigan—Ann Arbor

2800 Plymouth Rd., Bldg. 14, G016

Ann Arbor, MI, 48109

United States

Phone: 1 808 352 4196

Email: madanafi@med.umich.edu

Abstract

Background: As the US Food and Drug Administration (FDA)–approved use of artificial intelligence (AI) for medical imaging rises, radiologists are increasingly integrating AI into their clinical practices. In lung cancer screening, diagnostic AI offers a second set of eyes with the potential to detect cancer earlier than human radiologists. Despite AI's promise, a potential problem with its integration is the erosion of patient confidence in clinician expertise when there is a discrepancy between the radiologist's and the AI's interpretation of the imaging findings.

Objective: We examined how discrepancies between AI-derived recommendations and radiologists' recommendations affect patients' agreement with radiologists' recommendations and satisfaction with their radiologists. We also analyzed how patients' medical maximizing-minimizing preferences moderate these relationships.

Methods: We conducted a randomized, between-subjects experiment with 1606 US adult participants. Assuming the role of patients, participants imagined undergoing a low-dose computerized tomography scan for lung cancer screening and receiving results and recommendations from (1) a radiologist only, (2) AI and a radiologist in agreement, (3) a radiologist who recommended more testing than AI (ie, radiologist overcalled AI), or (4) a radiologist who recommended less testing than AI (ie, radiologist undercalled AI). Participants rated the radiologist on three criteria: agreement with the radiologist's recommendation, how likely they would be to recommend the radiologist to family and friends, and how good of a provider they perceived the radiologist to be. We measured medical maximizing-minimizing preferences and categorized participants as maximizers (ie, those who seek aggressive intervention), minimizers (ie, those who prefer no or passive intervention), and neutrals (ie, those in the middle).

Results: Participants' agreement with the radiologist's recommendation was significantly lower when the radiologist undercalled AI (mean 4.01, SE 0.07, $P<.001$) than in the other 3 conditions, with no significant differences among them (radiologist overcalled AI [mean 4.63, SE 0.06], agreed with AI [mean 4.55, SE 0.07], or had no AI [mean 4.57, SE 0.06]). Similarly, participants were least likely to recommend ($P<.001$) and positively rate ($P<.001$) the radiologist who undercalled AI, with no significant differences among the other conditions. Maximizers agreed with the radiologist who overcalled AI ($\beta=0.82$, SE 0.14; $P<.001$) and disagreed with the radiologist who undercalled AI ($\beta=-0.47$, SE 0.14; $P=.001$). However, whereas minimizers disagreed with the radiologist who overcalled AI ($\beta=-0.43$, SE 0.18, $P=.02$), they did not significantly agree with the radiologist who undercalled AI ($\beta=0.14$, SE 0.17, $P=.41$).

Conclusions: Radiologists who recommend less testing than AI may face decreased patient confidence in their expertise, but they may not face this same penalty for giving more aggressive recommendations than AI. Patients' reactions may depend in part on whether their general preferences to maximize or minimize align with the radiologists' recommendations. Future research should test communication strategies for radiologists' disclosure of AI discrepancies to patients.

(*J Med Internet Res* 2025;27:e68823) doi: [10.2196/68823](https://doi.org/10.2196/68823)

KEYWORDS

artificial intelligence; medical maximizing-minimizing; early detection of cancer; decision making; radiologists; patient satisfaction; patient-physician relationship; communication

Introduction

Radiologists are trained to detect lung cancer from low-dose computerized tomography (LDCT) scans with high accuracy [1,2] and medical artificial intelligence (AI) promises to improve this accuracy detection by acting as a “second pair of eyes” [1,3,4]. As of August 2024, the US Food and Drug Administration (FDA) published a list of 950 authorized AI-enabled medical devices, 723 (76%) of which were in the field of radiology. The developers of a medical AI model meant for lung cancer screening, called Sybil, claim the AI can detect future lung cancer in CT scans up to 6 years before lesions become visible to human radiologists [5]. In a 2020 survey of American College of Radiology members, 33% of radiologists reported currently using AI, while 20% of those not currently using AI planned to use it in the next 1 to 5 years [6].

Despite this rise in medical AI among radiologists, patient acceptance lags. Previous research has found patients’ resistance to AI stems from beliefs that AI performs inferior to humans and cannot be held accountable for errors, lower objective and subjective understanding of AI versus human decision-making, and fear that AI does not meet patients’ unique needs [7-10].

In part due to concerns about AI accuracy, US federal policies emphasize including a “human in the loop” (HITL) when it comes to AI development and use. In 2021, the FDA, along with Health Canada and the United Kingdom’s Medicines and Healthcare Products Regulatory Agency, published 10 guiding principles for machine learning development [11]. One of the principles describes focusing on the performance of the “human-AI team,” with AI performing with a HITL rather than in isolation.

Despite these guidelines, little is known about how patients will respond to the HITL. Whereas previous research has focused on patients’ trust or mistrust of medical AI [7,8,10], few studies have examined patients’ trust in the clinician of the human-AI team. Research shows patients’ trust in their clinician is related to patient satisfaction, compliance, and health outcomes [12-14]. Yet, medical AI could fundamentally change the patient-clinician relationship, and erode patients’ trust in clinician expertise [15,16].

In this study, we inquired how human-AI discrepancies affect patients’ agreement with radiologist recommendations and satisfaction with their radiologist. We examined 2 directions of human-AI discrepancy. First, the radiologist can “overcall AI,” in which the radiologist identifies higher cancer risk than the AI and recommends more testing. Second, the radiologist can “undercall AI,” in which the radiologist identifies lower cancer risk than the AI and recommends less testing. We hypothesize that patients will least agree with the recommendation of the radiologist who undercalls AI.

We anticipate patients’ agreement with their radiologist’s recommendation will vary based on individual differences, like

medical maximizing-minimizing (MMM) preferences. Whereas medical maximizers typically seek more aggressive and optional approaches to care, minimizers prefer no or passive medical intervention unless deemed completely necessary [17]. Researchers have shown that MMM is associated with preferences for cancer screening and surveillance [18,19], treatment of incidental findings on imaging tests [20], concern with stopping medication [21], pursuit of appropriate care [22], and avoidance of health care [23]. In this context, we hypothesize that measuring MMM might help to identify opportunities to tailor communication and guide shared decision-making about medical AI to patients with different underlying care preferences.

Methods**Participants and Procedure**

This study was preregistered at Open Science Framework [24]. We programmed the web-based quantitative experimental questionnaire in Qualtrics (Silver Lake) and collected data in April and May 2024. A total of 1828 adult, English-speaking participants were recruited by Dynata [25]. Dynata facilitates data collection for research surveys through its panel of potential web-based participants. We provided Dynata with quotas for age, gender, and race and ethnicity based on the US population composition. Inclusion criteria were at least 18 years of age and residents of the United States. These quotas and inclusion criteria enabled us to receive a diverse range of participant reactions to our experiment.

Study Design

The study was designed as a 1-way, between-subjects, randomized experiment. Participants assumed the role of a patient advised to undergo an LDCT scan for lung cancer screening. After reading the hypothetical scenario, participants were randomized to receive 1 of 4 results conditions, in which (1) a radiologist alone identified low risk of cancer and recommended a repeat screening CT in 6 months, (2) AI and a radiologist both identified low risk of cancer and recommended a repeat screening CT in 6 months, (3) AI identified medium risk of cancer and recommended immediate additional testing with a nuclear medicine examination (positron emission tomography-computed tomography [PET-CT]) but a radiologist identified low risk and recommended a repeat screening CT in 6 months (ie, radiologist undercalled AI), or (4) AI identified low risk of cancer and recommended a repeat screening CT in 6 months but a radiologist identified medium risk and recommended immediate additional testing (PET-CT; ie, radiologist overcalled AI). We used the randomizer within Qualtrics to randomly assign participants to conditions. Participants read the results both in an electronic health record (EHR) system report and in a transcript of an oral follow-up. **Textbox 1** shows the LDCT scan results received in the 4 experimental conditions.

Textbox 1. Hypothetical results of the low-dose computed tomography scan, received by participants in the four experimental conditions.

Radiologist only

- Impression:
 - There is a 7-mm solid nodule in the left lung.
 - Lung Imaging Reporting and Data System (Lung-RADS) 3: Probably benign. Low risk of malignancy (cancer) 1%-2%.
- Recommendation:
 - Repeat low-dose computed tomography (LDCT) in 6 months to ensure no interval growth.

Radiologist-AI (artificial intelligence) agreement

- Impression
 - LungDetect AI: There is a 7-mm solid nodule in the left lung. Lung-RADS 3: Probably benign. Low risk of malignancy (cancer) 1%-2%.
 - Radiologist: There is a 7-mm solid nodule in the left lung. Lung-RADS 3: Probably benign. Low risk of malignancy (cancer) 1%-2%.
- Recommendation
 - Repeat low-dose CT in 6 months to ensure no interval growth, based on AI and radiologist agreement.

Radiologist overcalls AI

- Impression
 - LungDetect AI: There is a 7-mm solid nodule in the left lung. Lung-RADS 3: Probably benign. Low risk of malignancy (cancer) 1%-2%.
 - Radiologist: There is a 9-mm solid nodule in the left lung. Lung-RADS 4A: Suspicious. Medium risk of malignancy (cancer) 5%-15%.
- Recommendation
 - LungDetect AI: Repeat low-dose CT in 6 months to ensure no interval growth.
 - Radiologist: Undergo additional imaging immediately.

Radiologist undercalls AI

- Impression
 - LungDetect AI: There is a 9-mm solid nodule in the left lung. Lung-RADS 4A: Suspicious. Medium risk of malignancy (cancer) 5%-15%.
 - Radiologist: There is a 7-mm solid nodule in the left lung. Lung-RADS 3: Probably benign. Low risk of malignancy (cancer) 1%-2%.
- Recommendation
 - LungDetect AI: Undergo additional imaging immediately.
 - Radiologist: Repeat low-dose CT in 6 months to ensure no interval growth.

Participants chose their preferred next examination (repeat CT in 6 months or PET-CT now) and rated how much they agreed with the radiologist's recommendation. They also rated their satisfaction with the radiologist. Participants then answered questions about their attitudes toward medical AI and their MMM preferences. Refer to [Multimedia Appendix 1](#) for the hypothetical scenario and questionnaire measures.

The main dependent variable was a single, 6-point Likert scale measure asking participants how much they agreed or disagreed with the radiologist's recommendation (1=strongly disagree to 6=strongly agree). We also measured participants' satisfaction with their radiologist using questions adapted from the Agency for Healthcare Research and Quality's Consumer Assessment of Healthcare Providers and Systems patient-experience surveys [26]. These included two 10-point scales asking participants

how likely they would be to recommend the radiologist to family and friends and how good of a provider they perceived the radiologist.

Covariates included mean composite scores of participants' attitudes toward medical AI and of participants' MMM preferences. Attitudes toward medical AI were measured by asking participants to rate their agreement with 7 statements on a 6-point Likert scale (1=strongly disagree to 6=strongly agree). Of these statements, 5 were adapted from previous literature [27,28], and 2 were added by the authors. We reverse-coded statements 2 through 6 before creating the composite so that higher scores represented greater affinity for medical AI. Cronbach α for the 7 measures was 0.62, indicating acceptable internal consistency.

We measured MMM using 2 previously published questions [17,29]. The measures asked, in medical situations, whether participants: (1) tended to lean toward waiting and seeing or taking action (1=I strongly lean toward waiting and seeing to 6=I strongly lean toward taking action) [17], and (2) tended to lean toward doing only what is necessary or everything possible (1=I strongly lean toward doing only what is necessary to 6=I strongly lean toward doing everything possible) [29]. Higher scores on these measures indicated a greater preference for medical maximizing. Cronbach α for the 2 measures was 0.74, indicating good internal consistency. We also categorized participants as minimizers (1-2.5), neutrals (3-4), and maximizers (4.5-6), in line with previous research [29,30], for ease of interpretation.

Analysis

We conducted ordinary least squares regressions to test whether receiving a recommendation from a radiologist alone, or in agreement or disagreement with AI, affected participants' agreement with their radiologist's recommendation. Two-sided $P < .05$ was considered significant. In exploratory analyses, we also tested how the experimental condition affected participants' follow-up test choice and participants' satisfaction with their radiologist.

In sensitivity analyses, we conducted Kruskal-Wallis H tests, Dunn tests, and ordinal logistic regression models to check the robustness of the effects on our main dependent variable, agreement with the radiologist's recommendation. We also conducted regressions with the inclusion of medical AI attitudes, MMM preferences, and participant demographics as covariates to check the sensitivity of the findings to the inclusion of controls. We ran moderation analyses with participants' MMM preferences first as a continuous variable and second as a categorical variable.

Ethical Considerations

The study was declared exempt from review by the University of Michigan Health Sciences and Behavioral Sciences institutional review board.

Results

Descriptive Data

The average time to completion was 8 minutes 32 seconds after excluding 222 speeder participants who completed the questionnaire in 3 minutes or less. A total of 1606 (88%) participants (51% female, mean age 51.49 years) were used in the analyses. Participants in each condition were similar in mean age and sex, indicating successful randomization (Table 1).

Table 1. Descriptive statistics for the sample (N=1606).

Characteristic	Radiologist only (n=413)	Radiologist-AI ^a agreement (n=401)	Radiologist over-calls AI (n=408)	Radiologist under-calls AI (n=384)	Total (N=1606)	P value ^b
Age (years), mean	52.15	50.79	51.29	51.72	51.49	.69
Gender, n (%)						.67
Women	215 (52)	206 (51)	219 (54)	183 (48)	823 (51)	
Men	194 (47)	192 (48)	187 (46)	199 (52)	772 (48)	
Other	4 (1)	3 (1)	2 (0)	2 (1)	11 (1)	
Highest level of completed education, n (%)						.28
High school or less	80 (19)	97 (24)	90 (22)	77 (20)	344 (21)	
Some postsecondary	133 (32)	127 (32)	117 (29)	137 (36)	514 (32)	
Bachelor's	125 (30)	105 (26)	127 (31)	108 (28)	465 (29)	
Graduate	73 (18)	67 (17)	74 (18)	60 (16)	274 (17)	
Other	2 (0)	5 (1)	0 (0)	2 (1)	9 (1)	
Race, n (%)						.76
White	314 (76)	286 (71)	301 (74)	290 (76)	1191 (74)	
Black	41 (10)	55 (14)	50 (12)	50 (13)	196 (12)	
Asian	22 (5)	22 (5)	24 (6)	17 (4)	85 (5)	
Multiracial or other	36 (9)	38 (9)	33 (8)	27 (7)	134 (8)	
Hispanic or Latino, n (%)						.42
Not Hispanic	332 (80)	314 (78)	344 (84)	317 (83)	1307 (81)	
Hispanic	78 (19)	82 (20)	60 (15)	63 (16)	283 (18)	
Other	3 (1)	5 (1)	4 (1)	4 (1)	16 (1)	

^aAI: artificial intelligence.

^bANOVA for age, chi-square tests for gender, education, race, and Hispanic/Latino. Some percentages may total >100 due to rounding.

Main Results

Participants' Agreement With the Radiologist's Recommendation

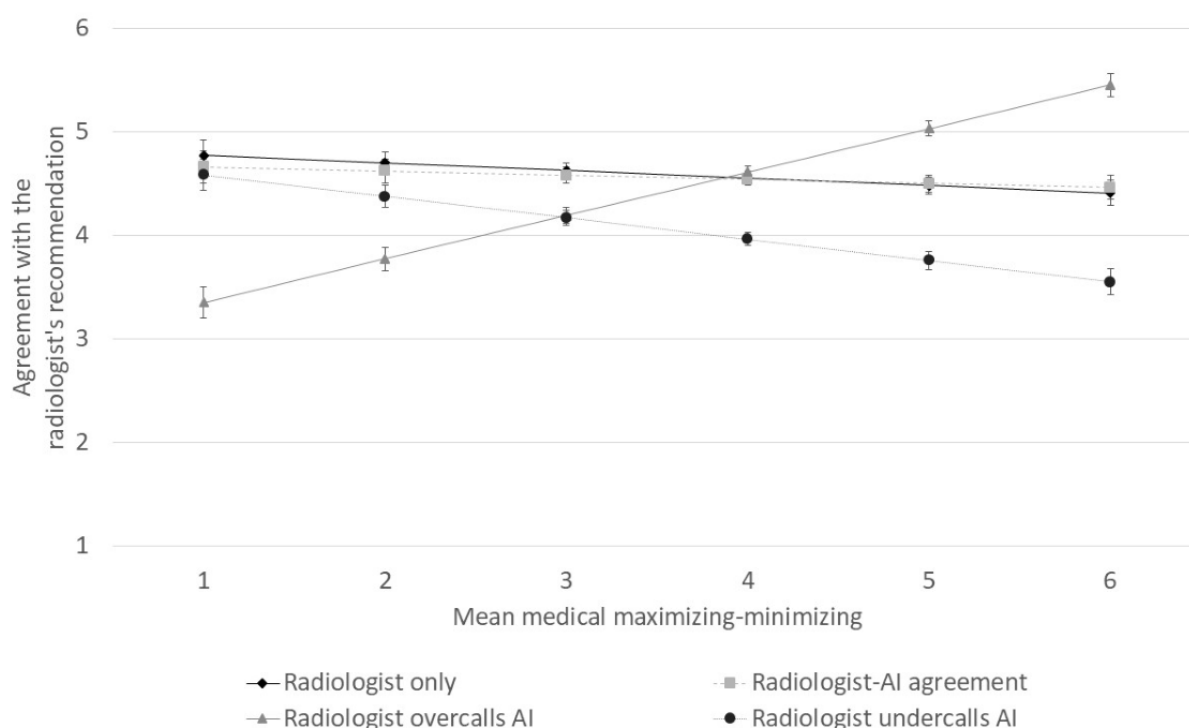
Participants were less likely to agree with the radiologist who undercalled AI (mean 4.01, SE 0.07), which was significantly different than the other conditions ($P<.001$). Participants' agreement with their radiologist's recommendation was not significantly different among the radiologist-only (mean 4.57, SE 0.06), radiologist-AI agreement (mean 4.55, SE 0.07), and radiologist-overcalls-AI (mean 4.63, SE 0.06) conditions. Results were robust to the inclusion of covariates, including medical AI attitudes, MMM preferences, and participant demographics (Table S1 in [Multimedia Appendix 1](#)), and

remained consistent when using alternative model specifications, including the Kruskal-Wallis H tests and Dunn tests and ordinal logistic regressions (Tables S2 and S3 [Multimedia Appendix 1](#)).

Moderation by Medical Maximizing-Minimizing Preferences

MMM scores had a strong positive effect on participants' agreement in the radiologist-overcalls-AI condition ($\beta=0.42$, SE 0.05, $P<.001$) and a negative effect on the radiologist-undercalls-AI condition ($\beta=-0.21$, SE 0.05, $P<.001$). There was no significant effect on the radiologist-only ($\beta=-0.07$, SE 0.05, $P=.13$) and radiologist-AI agreement ($\beta=-0.04$, SE 0.05, $P=.42$) conditions ([Figure 1](#)).

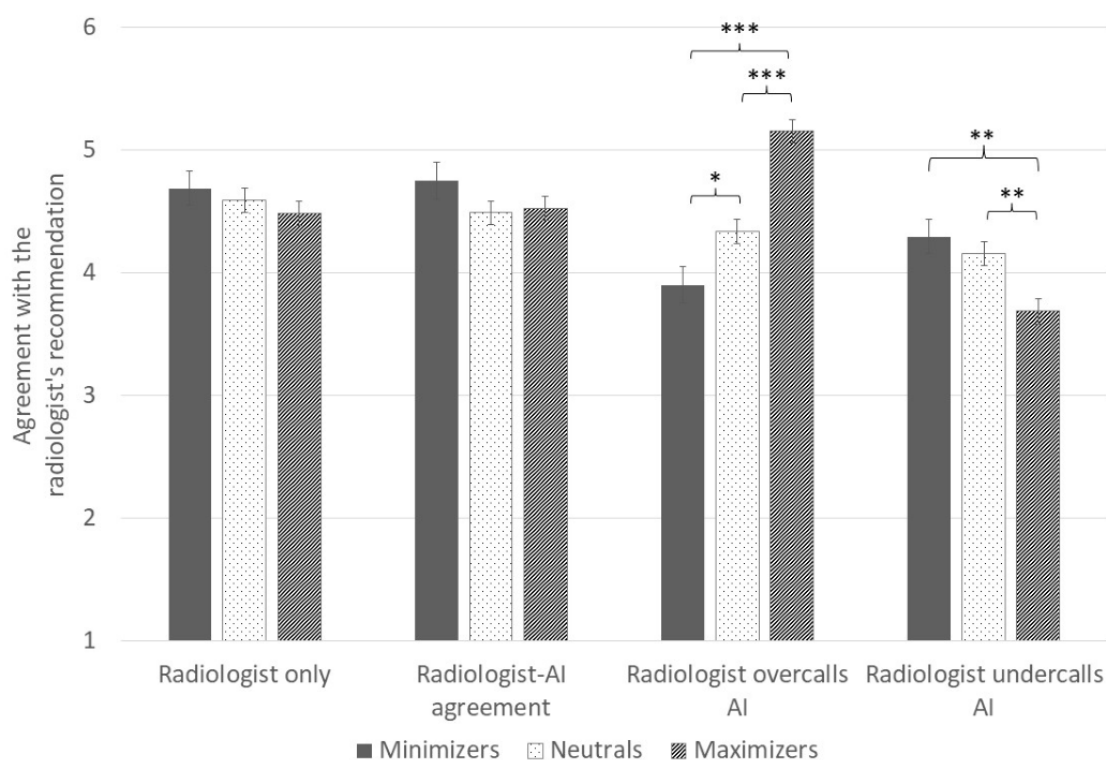
Figure 1. Effect of participants' medical maximizing-minimizing preferences on agreement with the radiologist's recommendation, by condition. Error bars depict SEs. AI: artificial intelligence.



In our sample, MMM scores were left-skewed; 302 (18.8%) participants were minimizers, 654 (40.72%) were neutrals, and 650 (40.47%) were maximizers. When analyzing by MMM category, we found, relative to neutrals, minimizers disagreed more strongly with the radiologist who overcalled AI ($\beta=-0.43$, SE 0.18, $P=.02$) but did not agree more strongly with the radiologist who undercalled AI ($\beta=0.14$, SE 0.17, $P=.41$).

Relative to neutrals, maximizers agreed more strongly with the radiologist who overcalled AI ($\beta=0.82$, SE 0.14, $P<.001$) and disagreed more strongly with the radiologist who undercalled AI ($\beta=-0.47$, SE 0.14, $P=.001$; [Figure 2](#)). The pattern of this moderation analysis with MMM as a categorical variable matched that of the analysis using MMM as a continuous variable.

Figure 2. Effect of condition on participants' agreement with the radiologist's recommendation, by participants' MMM category. Means and standard error bars are depicted. Stars represent significant differences between groups (* $P<.05$, ** $P<.01$, *** $P<.001$). AI: artificial intelligence; MMM: medical maximizing-minimizing.



Secondary Results

Patients' Choice Between Follow-Up Screening Computed Tomography in 6 Months or Additional Testing (PET-CT) Immediately

Participants were least likely to follow the recommendation of the radiologist who undercalled AI (184/384, 47.9%), which was significantly different than the other conditions ($P<.001$). Participants' choice to follow their radiologist's recommendation was not significantly different among the radiologist-only (260/413, 63%), radiologist-AI agreement (254/401, 63.3%), and radiologist-overcalls-AI conditions (263/408, 64.5%; Figure S1 in [Multimedia Appendix 1](#)). This pattern aligned with the results of participants' agreement with their radiologist's recommendation.

Patients' Likelihood to Recommend the Radiologist and Radiologist Rating

In exploratory analyses, we examined participants' likelihood to recommend their radiologist and the rating of their radiologist after receiving the results of their LDCT scan. [Table 2](#) presents means and SEs for all dependent variables. Participants were both least likely to recommend (mean 6.47, SE 0.13) and positively rate (mean 6.82, SE 0.11) the radiologist who undercalled AI, which was significantly different from the other conditions ($P<.001$). There were no significant differences among the radiologist-only, radiologist-AI agreement, and radiologist-overcalls-AI conditions in participants' recommendation likelihood (mean 7.24, SE 0.12 vs mean 7.40, SE 0.12 vs mean 7.43, SE 0.11) or rating of the radiologist (mean 7.45, SE 0.10 vs mean 7.55, SE 0.10 vs mean 7.58, SE=0.10). Results were robust to the inclusion of covariates, including AI attitudes, MMM preferences, and participant demographics. (Tables S4 and S5 in [Multimedia Appendix 1](#)).

Table 2. Means and standard errors of the dependent variables.

Dependent variable	Radiologist only (n=413)	Radiologist-AI ^a agreement (n=401)	Radiologist overcalls AI (n=408)	Radiologist undercalls AI (n=384)
Agreement with radiologist's recommendation ^b	4.57 (0.06)	4.55 (0.07)	4.63 (0.06)	4.01 (0.07)
Likelihood to recommend radiologist ^b	7.24 (0.12)	7.40 (0.12)	7.43 (0.11)	6.47 (0.13)
Rating of radiologist ^b	7.45 (0.10)	7.55 (0.10)	7.58 (0.10)	6.82 (0.11)

^aAI: artificial intelligence.

^bThe radiologist-undercalls-AI group significantly differed from the three other groups on this dependent variable ($P<.001$). Agreement with the radiologist's recommendation measured from 1=strongly disagree to 6=strongly agree. The likelihood to recommend the radiologist measured from 1=definitely would not recommend to 10=definitely would recommend. The rating of radiologist measured from 1=worst provider possible to 10=best provider possible.

Moderation by Medical Maximizing-Minimizing Preferences

MMM scores had a positive effect on participants' likelihood to recommend the radiologist in the radiologist-overcalls-AI condition ($\beta=0.71$, SE 0.09, $P<.001$). Relative to neutrals, maximizers in the radiologist-overcalls-AI condition were more likely to recommend the radiologist ($\beta=1.49$, SE 0.26, $P<.001$).

MMM scores also had a positive effect on participants' ratings of the radiologist in the radiologist-AI agreement ($\beta=0.22$, SE 0.08, $P=.005$) and the radiologist-overcalls-AI ($\beta=0.61$, SE 0.07, $P<.001$) conditions. Relative to neutrals, minimizers in the radiologist-overcalls-AI condition gave lower radiologist ratings ($\beta=-0.63$, SE 0.28, $P=.03$) whereas maximizers gave higher ratings ($\beta=1.28$, SE 0.22, $P<.001$). In addition, relative to neutrals, maximizers in the radiologist-undercalls-AI condition gave lower radiologist ratings ($\beta=-0.49$, SE 0.23, $P=.03$) (Figures S2 and S3 in [Multimedia Appendix 1](#)).

Discussion

Principal Findings

In this study, we found that participants penalized radiologists who undercalled AI in favor of less testing but not radiologists who overcalled AI in favor of more testing. Maximizers especially disagreed with radiologists who undercalled AI. However, whereas maximizers agreed with radiologists who overcalled AI, minimizers strongly disagreed. This suggests people's confidence in the expertise of radiologists who use AI may depend in part on whether their general preferences to maximize or minimize align with the provided recommendations.

Previous research has shown that patients demonstrate algorithm aversion and superior trust in their human clinicians versus AI [8,31-33]. Our research suggests this may not always be the case. When AI and human clinicians are at odds, patients may trust AI over the clinician when AI detects something more than the clinician and recommends more aggressive treatment.

Although radiologists were penalized for undercalling AI, they were not penalized for overcalling AI. This overall attitude toward the radiologist who overcalled AI masked significant variability based on participants' MMM preferences. Whereas

minimizers penalized the radiologist, maximizers rewarded the radiologist. The averaging of these competing attitudes resulted in a level of agreement that did not significantly differ from that in the radiologist-only condition. Thus, radiologists who overcall AI may in fact face significantly lowered trust from patients who have general minimizing preferences.

We did not observe a similar aggregate averaging effect in the radiologist-undercalls-AI condition. This asymmetry may be explained by the perceptions of minimizers. Whereas maximizers demonstrated stable preferences for pursuing more aggressive treatment [30], regardless of whether the radiologist or AI recommended it, minimizers penalized the radiologist who overcalled AI but did not reward the radiologist who undercalled AI.

Minimizers typically prefer not to receive treatment unless considered essential [30,34]. However, previous research has shown minimizers are more responsive to information than maximizers and will change their attitudes in response to evidence [18]. Thus, 1 theory for this asymmetry is that minimizers perceived AI as more accurate than the radiologist and therefore deferred to the AI's recommendation both when AI recommended more and less testing. This suggests patients, like clinicians, may be prone to automation bias when medical AI is involved [35,36]. Another theory is that the emotional salience of the AI's cancer results softened minimizers' attitudes toward more testing. Importantly, patients demonstrated higher agreement in the radiologist-only and radiologist-AI agreement conditions, despite the radiologist recommending less testing just like in the radiologist-undercalls-AI condition. This suggests participants were opposed to the radiologist's recommendation only when in contrast with the more aggressive AI recommendation. Research suggests the emotional salience of a cancer diagnosis, particularly fear of cancer, influences patients' decision-making [37-39]. More research is needed to test these theories, including examining patients' general and context-specific perceptions of the accuracy of AI compared with that of clinicians [31].

Radiologists' confirmatory "second opinion" of AI did not boost participants' agreement with them. This finding adds a new dimension to the literature examining patient perceptions of human- versus model-assisted decisions. In 1 study, patients' trust in AI increased when AI, acting as a second opinion, confirmed the human clinician's diagnosis, but decreased when

the AI was disconfirming [32]. In another study, participants devalued clinicians who consulted a computer-based diagnostic aid for a second opinion, but not clinicians who consulted a human expert [40]. In this study, we found radiologists who agreed with the AI were neither rewarded nor penalized compared with the radiologist alone.

Research shows the human-AI team augments the diagnostic performance of the radiologist [41,42]. However, we show that the human-AI team may not benefit, and could even hurt, the radiologist in terms of patient agreement or satisfaction. This study demonstrated that even small discrepancies (ie, 2 mm) between AI and radiologist interpretations can elicit strong patient reactions. We anticipate larger discrepancies would result in more pronounced negative reactions toward clinicians. As a result, clinicians may feel compelled to overcall or agree with AI to avoid negative patient reactions, as well as potential legal consequences [43,44]. This may lead to incorrect clinician decisions [45], as well as unnecessary testing, biopsies, and procedures. More research is needed to probe clinicians' attitudes toward the consequences of using AI.

Future research is also needed to better understand patient attitudes toward the human-AI team in various clinical contexts. Research should also interrogate what and how to communicate AI results with patients, especially when the human-AI team disagrees. Offering 2 recommendations to patients may eschew clinicians' professional agency and burden patients with a choice between following clinician or AI expertise [46]. However, there are ethical considerations to smoothing over discrepancies with a single recommendation that sides with either the clinician or the AI. If both recommendations are medically reasonable, a single clinician recommendation could be seen as too paternalistic [47] and may go against patient preference [48].

Limitations

This study has several limitations. First, we intentionally recruited a diverse sample of US adults to receive the broadest range of participants; however, this sample is not reflective of the patients who typically would receive LDCT scans. Future research should consider replicating this study with a narrowed sample of participants, particularly smokers, or real patients undergoing cancer screening. For example, we anticipate participants with a higher pretest probability for lung cancer would find a radiologist undercalling AI particularly difficult to accept in this scenario.

Second, we instructed participants that they were eligible to receive an LDCT scan because of "personal risk factors" but did not specify further. Risk factors could range from genetic to environmental to behavioral, and how we develop these risk factors influences our acceptance of test results [49,50]. If participants had different mental models for how they would be eligible for lung cancer screening, these differences could have impacted their responses to the AI and radiologist's recommendations.

Acknowledgments

This study was funded by a pilot grant awarded by the University of Michigan Center for Bioethics and Social Sciences in Medicine.

Third, although studies show patients want to know and consent to when AI is used [27,51], we do not yet have a template for how health care systems currently report, will report in the future, or should report AI use to patients. In our study, the EHR results and oral follow-up provided both the AI and radiologist's interpretations and recommendations. A different presentation of AI's involvement, such as AI's interpretation but not a separate recommendation, or simply that AI was consulted but not its interpretation, could elicit different patient perceptions of their radiologist. We also do not provide measures of accuracy for either AI or radiologists. Although radiologists' accuracy is not typically disclosed to patients who receive imaging, it is possible that patients may inquire about the AI's accuracy, especially vis-à-vis the radiologists'. How patients understand and interpret measures of accuracy, including more advanced measures, such as sensitivity, specificity, precision, and recall, requires further investigation.

Furthermore, future research is needed to develop validated instruments for measuring attitudes toward medical AI. Although there are validated scales measuring attitudes toward AI in general [52,53], these were insufficient for our experimental context. Thus, the scale we deployed, for which we created a composite score from custom-designed questions and achieved a low Cronbach α , was neither validated nor highly reliable, a limitation of our study. As a result, we may not have adequately controlled for confounding medical AI attitudes, weakening our overall findings.

Finally, the radiologist-patient relationship differs from the classic clinician-patient relationship, as patients might not choose, directly communicate with, or have a preexisting relationship with their radiologist [54,55]. The nature of the radiologist-patient relationship may therefore not generalize to other medical specialties. Despite these limitations, exploring patients' attitudes toward cancer screening recommendations and the radiologist-patient relationship are understudied yet increasingly important topics in the era of medical AI [56].

Conclusion

As radiologists begin to integrate AI into cancer detection, discrepancies within the human-AI team may influence patients' reactions toward their radiologist. People may penalize radiologists who undercall AI in lung cancer screening yet may not be more confident in radiologists who overcall or agree with AI. Patients' MMM preferences moderate this effect and accounting for patients' MMM preferences may give insight to clinicians into what and how much information to give about AI's role in their decision-making. Our findings highlight the complexity of the patient-AI-clinician relationship and have implications for clinical practice, communication, and shared decision-making. Future research is needed to determine how radiologists should communicate AI discrepancies to patients in a way that builds trust and maintains their relevance.

Data Availability

The dataset generated for this study is available from the corresponding author on request.

Authors' Contributions

FM, LSO, and BJZ conceptualized the study design. FM conducted data collection and formal analysis. FM, LSO, and BJZ drafted and edited the manuscript. All authors read and approved of the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional material.

[\[DOCX File , 752 KB-Multimedia Appendix 1\]](#)

References

1. Wang TW, Hong JS, Chiu HY, Chao HS, Chen YM, Wu YT. Standalone deep learning versus experts for diagnosis lung cancer on chest computed tomography: a systematic review. *Eur Radiol*. 2024;34(11):7397-7407. [doi: [10.1007/s00330-024-10804-6](https://doi.org/10.1007/s00330-024-10804-6)] [Medline: [38777902](https://pubmed.ncbi.nlm.nih.gov/38777902/)]
2. Gierada DS, Pinsky P, Nath H, Chiles C, Duan F, Aberle DR. Projected outcomes using different nodule sizes to define a positive CT lung cancer screening examination. *J Natl Cancer Inst*. 2014;106(11):dju284. [FREE Full text] [doi: [10.1093/jnci/dju284](https://doi.org/10.1093/jnci/dju284)] [Medline: [25326638](https://pubmed.ncbi.nlm.nih.gov/25326638/)]
3. Tam MDBS, Dyer T, Dissez G, Morgan TN, Hughes M, Illes J, et al. Augmenting lung cancer diagnosis on chest radiographs: positioning artificial intelligence to improve radiologist performance. *Clin Radiol*. 2021;76(8):607-614. [doi: [10.1016/j.crad.2021.03.021](https://doi.org/10.1016/j.crad.2021.03.021)] [Medline: [33993997](https://pubmed.ncbi.nlm.nih.gov/33993997/)]
4. Rubin DL. Artificial intelligence in imaging: The radiologist's role. *J Am Coll Radiol*. 2019;16(9 Pt B):1309-1317. [FREE Full text] [doi: [10.1016/j.jacr.2019.05.036](https://doi.org/10.1016/j.jacr.2019.05.036)] [Medline: [31492409](https://pubmed.ncbi.nlm.nih.gov/31492409/)]
5. Mikhael PG, Wohlwend J, Yala A, Karstens L, Xiang J, Takigami AK, et al. Sybil: a validated deep learning model to predict future lung cancer risk from a single low-dose chest computed tomography. *J Clin Oncol*. 2023;41(12):2191-2200. [FREE Full text] [doi: [10.1200/JCO.22.01345](https://doi.org/10.1200/JCO.22.01345)] [Medline: [36634294](https://pubmed.ncbi.nlm.nih.gov/36634294/)]
6. Allen B, Agarwal S, Coombs L, Wald C, Dreyer K. 2020 ACR data science institute artificial intelligence survey. *J Am Coll Radiol*. 2021;18(8):1153-1159. [doi: [10.1016/j.jacr.2021.04.002](https://doi.org/10.1016/j.jacr.2021.04.002)] [Medline: [33891859](https://pubmed.ncbi.nlm.nih.gov/33891859/)]
7. Longoni C, Bonezzi A, Morewedge CK. Resistance to medical artificial intelligence. *Journal of Consumer Research*. 2019;46(4):629-650. [doi: [10.1093/jcr/ucz013](https://doi.org/10.1093/jcr/ucz013)]
8. Cadario R, Longoni C, Morewedge CK. Understanding, explaining, and utilizing medical artificial intelligence. *Nat Hum Behav*. 2021;5(12):1636-1642. [doi: [10.1038/s41562-021-01146-0](https://doi.org/10.1038/s41562-021-01146-0)] [Medline: [34183800](https://pubmed.ncbi.nlm.nih.gov/34183800/)]
9. Promberger M, Baron J. Do patients trust computers? *Behavioral Decision Making*. 2006;19(5):455-468. [doi: [10.1002/bdm.542](https://doi.org/10.1002/bdm.542)]
10. Gaczek P, Pozharliev R, Leszczyński G, Zieliński M. Overcoming consumer resistance to AI in general health care. *Journal of Interactive Marketing*. 2023;58(2-3):321-338. [doi: [10.1177/10949968221151061](https://doi.org/10.1177/10949968221151061)]
11. Food U, Administration D. Good machine learning practice for medical device development: guiding principles. Washington, DC, USA. The US Food and Drug Administration; 2021.
12. Birkhäuser J, Gaab J, Kossowsky J, Hasler S, Krummenacher P, Werner C, et al. Trust in the health care professional and health outcome: A meta-analysis. *PLoS One*. 2017;12(2):e0170988. [FREE Full text] [doi: [10.1371/journal.pone.0170988](https://doi.org/10.1371/journal.pone.0170988)] [Medline: [28170443](https://pubmed.ncbi.nlm.nih.gov/28170443/)]
13. Lee YY, Lin JL. The effects of trust in physician on self-efficacy, adherence and diabetes outcomes. *Soc Sci Med*. 2009;68(6):1060-1068. [doi: [10.1016/j.socscimed.2008.12.033](https://doi.org/10.1016/j.socscimed.2008.12.033)] [Medline: [19162386](https://pubmed.ncbi.nlm.nih.gov/19162386/)]
14. Thom DH, Hall MA, Pawlson LG. Measuring patients' trust in physicians when assessing quality of care. *Health Aff (Millwood)*. 2004;23(4):124-132. [doi: [10.1377/hlthaff.23.4.124](https://doi.org/10.1377/hlthaff.23.4.124)] [Medline: [15318572](https://pubmed.ncbi.nlm.nih.gov/15318572/)]
15. Zondag AGM, Rozestraten R, Grimmelikhuijsen SG, Jongsma KR, van Solinge WW, Bots ML, et al. The effect of artificial intelligence on patient-physician trust: Cross-sectional vignette study. *J Med Internet Res*. 2024;26:e50853. [FREE Full text] [doi: [10.2196/50853](https://doi.org/10.2196/50853)] [Medline: [38805702](https://pubmed.ncbi.nlm.nih.gov/38805702/)]
16. Derevianko A, Pizzoli SFM, Pesapane F, Rotili A, Monzani D, Grasso R, et al. The use of artificial intelligence (AI) in the radiology field: What is the state of doctor-patient communication in cancer diagnosis? *Cancers (Basel)*. 2023;15(2):470. [FREE Full text] [doi: [10.3390/cancers15020470](https://doi.org/10.3390/cancers15020470)] [Medline: [36672417](https://pubmed.ncbi.nlm.nih.gov/36672417/)]
17. Scherer LD, Zikmund-Fisher BJ. Eliciting medical maximizing-minimizing preferences with a single question: development and validation of the MM1. *Med Decis Making*. 2020;40(4):545-550. [doi: [10.1177/0272989X20927700](https://doi.org/10.1177/0272989X20927700)] [Medline: [32522094](https://pubmed.ncbi.nlm.nih.gov/32522094/)]

18. Scherer LD, Kullgren JT, Caverly T, Scherer AM, Shaffer VA, Fagerlin A, et al. Medical maximizing-minimizing preferences predict responses to information about prostate-specific antigen screening. *Med Decis Making*. 2018;38(6):708-718. [doi: [10.1177/0272989X18782199](https://doi.org/10.1177/0272989X18782199)] [Medline: [29916784](https://pubmed.ncbi.nlm.nih.gov/29916784/)]
19. Evron JM, Reyes-Gastelum D, Banerjee M, Scherer LD, Wallner LP, Hamilton AS, et al. Role of patient maximizing-minimizing preferences in thyroid cancer surveillance. *J Clin Oncol*. 2019;37(32):3042-3049. [FREE Full text] [doi: [10.1200/JCO.19.01411](https://doi.org/10.1200/JCO.19.01411)] [Medline: [31573822](https://pubmed.ncbi.nlm.nih.gov/31573822/)]
20. Kang SK, Scherer LD, Megibow AJ, Higuera LJ, Kim N, Braithwaite RS, et al. A randomized study of patient risk perception for incidental renal findings on diagnostic imaging tests. *AJR Am J Roentgenol*. 2018;210(2):369-375. [FREE Full text] [doi: [10.2214/AJR.17.18485](https://doi.org/10.2214/AJR.17.18485)] [Medline: [29140116](https://pubmed.ncbi.nlm.nih.gov/29140116/)]
21. Vordenberg SE, Zikmund-Fisher BJ. Characteristics of older adults predict concern about stopping medications. *J Am Pharm Assoc* (2003). 2020;60(6):773-780. [doi: [10.1016/j.japh.2020.01.019](https://doi.org/10.1016/j.japh.2020.01.019)] [Medline: [32115393](https://pubmed.ncbi.nlm.nih.gov/32115393/)]
22. Scherer LD, Shaffer VA, Caverly T, DeWitt J, Zikmund-Fisher BJ. Medical maximizing-minimizing predicts patient preferences for high- and low-benefit care. *Med Decis Making*. 2020;40(1):72-80. [doi: [10.1177/0272989X19891181](https://doi.org/10.1177/0272989X19891181)] [Medline: [31975657](https://pubmed.ncbi.nlm.nih.gov/31975657/)]
23. Smith KT, Monti D, Mir N, Peters E, Tipirneni R, Politi MC. Access Is necessary but not sufficient: factors influencing delay and avoidance of health care services. *MDM Policy Pract*. 2018;3(1):2381468318760298. [FREE Full text] [doi: [10.1177/2381468318760298](https://doi.org/10.1177/2381468318760298)] [Medline: [30288438](https://pubmed.ncbi.nlm.nih.gov/30288438/)]
24. Madanay F, O'Donohue L, Zikmund-Fisher B. AI-physician discrepancy. OSF Registries. URL: <https://osf.io/j3yf2> [accessed 2025-05-13]
25. dynata. URL: <https://www.dynata.com/> [accessed 2025-05-13]
26. AHRQ. CAHPS patient experience surveys and guidance. Agency for Healthcare Research and Quality. 2023. URL: <https://www.ahrq.gov/cahps/surveys-guidance/index.html> [accessed 2025-04-28]
27. Khullar D, Casalino LP, Qian Y, Lu Y, Krumholz HM, Aneja S. Perspectives of patients about artificial intelligence in health care. *JAMA Netw Open*. 2022;5(5):e2210309. [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.10309](https://doi.org/10.1001/jamanetworkopen.2022.10309)] [Medline: [35507346](https://pubmed.ncbi.nlm.nih.gov/35507346/)]
28. Fritsch SJ, Blankenheim A, Wahl A, Hetfeld P, Maassen O, Deffge S, et al. Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. *Digit Health*. 2022;8:20552076221116772. [FREE Full text] [doi: [10.1177/20552076221116772](https://doi.org/10.1177/20552076221116772)] [Medline: [35983102](https://pubmed.ncbi.nlm.nih.gov/35983102/)]
29. Dossett LA, Mott NM, Bredbeck BC, Wang T, Jobin CT, Hughes TM, et al. Using tailored messages to target overuse of low-value breast cancer care in older women. *J Surg Res*. 2022;270:503-512. [FREE Full text] [doi: [10.1016/j.jss.2021.10.005](https://doi.org/10.1016/j.jss.2021.10.005)] [Medline: [34801801](https://pubmed.ncbi.nlm.nih.gov/34801801/)]
30. Mott N, Wang T, Miller J, Berlin NL, Hawley S, Jagsi R, et al. Medical maximizing-minimizing preferences in relation to low-value services for older women with hormone receptor-positive breast cancer: A qualitative study. *Ann Surg Oncol*. 2021;28(2):941-949. [FREE Full text] [doi: [10.1245/s10434-020-08924-6](https://doi.org/10.1245/s10434-020-08924-6)] [Medline: [32720038](https://pubmed.ncbi.nlm.nih.gov/32720038/)]
31. Robertson C, Woods A, Bergstrand K, Findley J, Balser C, Slepian MJ. Diverse patients' attitudes towards artificial intelligence (AI) in diagnosis. *PLOS Digit Health*. 2023;2(5):e0000237. [FREE Full text] [doi: [10.1371/journal.pdig.0000237](https://doi.org/10.1371/journal.pdig.0000237)] [Medline: [37205713](https://pubmed.ncbi.nlm.nih.gov/37205713/)]
32. Juravle G, Boudouraki A, Terziyska M, Rezlescu C. Trust in artificial intelligence for medical diagnoses. *Prog Brain Res*. 2020;253:263-282. [doi: [10.1016/bs.pbr.2020.06.006](https://doi.org/10.1016/bs.pbr.2020.06.006)] [Medline: [32771128](https://pubmed.ncbi.nlm.nih.gov/32771128/)]
33. York T, Jenney H, Jones G. Clinician and computer: a study on patient perceptions of artificial intelligence in skeletal radiography. *BMJ Health Care Inform*. 2020;27(3):e100233. [FREE Full text] [doi: [10.1136/bmjhci-2020-100233](https://doi.org/10.1136/bmjhci-2020-100233)] [Medline: [33187956](https://pubmed.ncbi.nlm.nih.gov/33187956/)]
34. Scherer LD, Caverly TJ, Burke J, Zikmund-Fisher BJ, Kullgren JT, Steinley D, et al. Development of the medical maximizer-minimizer scale. *Health Psychol*. 2016;35(11):1276-1287. [doi: [10.1037/hea0000417](https://doi.org/10.1037/hea0000417)] [Medline: [27617512](https://pubmed.ncbi.nlm.nih.gov/27617512/)]
35. Jabbour S, Fouhey D, Shepard S, Valley TS, Kazerooni EA, Banovic N, et al. Measuring the impact of AI in the diagnosis of hospitalized patients: A randomized clinical vignette survey study. *JAMA*. 2023;330(23):2275-2284. [FREE Full text] [doi: [10.1001/jama.2023.22295](https://doi.org/10.1001/jama.2023.22295)] [Medline: [38112814](https://pubmed.ncbi.nlm.nih.gov/38112814/)]
36. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc*. 2012;19(1):121-127. [FREE Full text] [doi: [10.1136/amiainl-2011-000089](https://doi.org/10.1136/amiainl-2011-000089)] [Medline: [21685142](https://pubmed.ncbi.nlm.nih.gov/21685142/)]
37. Mazzocco K, Masiero M, Carriero MC, Pravettoni G. The role of emotions in cancer patients' decision-making. *Ecancermedicalscience*. 2019;13:914. [FREE Full text] [doi: [10.3332/ecancer.2019.914](https://doi.org/10.3332/ecancer.2019.914)] [Medline: [31123497](https://pubmed.ncbi.nlm.nih.gov/31123497/)]
38. Nold RJ, Beamer RL, Helmer SD, McBoyle MF. Factors influencing a woman's choice to undergo breast-conserving surgery versus modified radical mastectomy. *Am J Surg*. 2000;180(6):413-418. [doi: [10.1016/s0002-9610\(00\)00501-8](https://doi.org/10.1016/s0002-9610(00)00501-8)] [Medline: [11182389](https://pubmed.ncbi.nlm.nih.gov/11182389/)]
39. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Risky feelings: why a 6% risk of cancer does not always feel like 6%. *Patient Educ Couns*. 2010;81 Suppl:S87-S93. [FREE Full text] [doi: [10.1016/j.pec.2010.07.041](https://doi.org/10.1016/j.pec.2010.07.041)] [Medline: [20739135](https://pubmed.ncbi.nlm.nih.gov/20739135/)]
40. Probst CA, Shaffer VA, Lambdin C, Arkes HR, Medow MA, Wichita KS, editor. *Ratings of Physicians Relying on Experts Versus Physicians Relying on Decision Aids*. USA. Wichita State University; 2008.

41. Li D, Pehrson LM, Lauridsen CA, Tøttrup L, Fraccaro M, Elliott D, et al. The added effect of artificial intelligence on physicians' performance in detecting thoracic pathologies on CT and chest X-ray: A systematic review. *Diagnostics* (Basel). 2021;11(12):2206. [FREE Full text] [doi: [10.3390/diagnostics11122206](https://doi.org/10.3390/diagnostics11122206)] [Medline: [34943442](https://pubmed.ncbi.nlm.nih.gov/34943442/)]
42. Bennani S, Regnard NE, Ventre J, Lassalle L, Nguyen T, Ducarouge A, et al. Using AI to improve radiologist performance in detection of abnormalities on chest radiographs. *Radiology*. 2023;309(3):e230860. [doi: [10.1148/radiol.230860](https://doi.org/10.1148/radiol.230860)] [Medline: [38085079](https://pubmed.ncbi.nlm.nih.gov/38085079/)]
43. Lawton T, Morgan P, Porter Z, Hickey S, Cunningham A, Hughes N, et al. Clinicians risk becoming 'liability sinks' for artificial intelligence. *Future Healthc J*. 2024;11(1):100007. [FREE Full text] [doi: [10.1016/j.fhj.2024.100007](https://doi.org/10.1016/j.fhj.2024.100007)] [Medline: [38646041](https://pubmed.ncbi.nlm.nih.gov/38646041/)]
44. Smith H, Fotheringham K. Artificial intelligence in clinical decision-making: Rethinking liability. *Medical Law International*. 2020;20(2):131-154. [doi: [10.1177/0968533220945766](https://doi.org/10.1177/0968533220945766)]
45. Bernstein MH, Atalay MK, Dibble EH, Maxwell AWP, Karam AR, Agarwal S, et al. Can incorrect artificial intelligence (AI) results impact radiologists, and if so, what can we do about it? A multi-reader pilot study of lung cancer detection with chest radiography. *Eur Radiol*. 2023;33(11):8263-8269. [FREE Full text] [doi: [10.1007/s00330-023-09747-1](https://doi.org/10.1007/s00330-023-09747-1)] [Medline: [37266657](https://pubmed.ncbi.nlm.nih.gov/37266657/)]
46. Kilbride MK, Joffe S. The new age of patient autonomy: Implications for the patient-physician relationship. *JAMA*. 2018;320(19):1973-1974. [FREE Full text] [doi: [10.1001/jama.2018.14382](https://doi.org/10.1001/jama.2018.14382)] [Medline: [30326026](https://pubmed.ncbi.nlm.nih.gov/30326026/)]
47. Quill TE, Brody H. Physician recommendations and patient autonomy: finding a balance between physician power and patient choice. *Ann Intern Med*. 1996;125(9):763-769. [doi: [10.7326/0003-4819-125-9-199611010-00010](https://doi.org/10.7326/0003-4819-125-9-199611010-00010)] [Medline: [8929011](https://pubmed.ncbi.nlm.nih.gov/8929011/)]
48. Gurmankin AD, Baron J, Hershey JC, Ubel PA. The role of physicians' recommendations in medical treatment decisions. *Med Decis Making*. 2002;22(3):262-271. [doi: [10.1177/0272989X0202200314](https://doi.org/10.1177/0272989X0202200314)] [Medline: [12058783](https://pubmed.ncbi.nlm.nih.gov/12058783/)]
49. Bostrom A. Mental models of risk. *Oxford Research Encyclopedia of Communication*. 2017. URL: <https://tinyurl.com/5pd6n3rf> [accessed 2017-08-22]
50. Holtrop JS, Scherer LD, Matlock DD, Glasgow RE, Green LA. The importance of mental models in implementation science. *Front Public Health*. 2021;9:680316. [FREE Full text] [doi: [10.3389/fpubh.2021.680316](https://doi.org/10.3389/fpubh.2021.680316)] [Medline: [34295871](https://pubmed.ncbi.nlm.nih.gov/34295871/)]
51. Moy S, Irannejad M, Manning SJ, Farahani M, Ahmed Y, Gao E, et al. Patient perspectives on the use of artificial intelligence in health care: A scoping review. *J Patient Cent Res Rev*. 2024;11(1):51-62. [FREE Full text] [doi: [10.17294/2330-0698.2029](https://doi.org/10.17294/2330-0698.2029)] [Medline: [38596349](https://pubmed.ncbi.nlm.nih.gov/38596349/)]
52. Schepman A, Rodway P. Initial validation of the general attitudes towards artificial intelligence scale. *Comput Hum Behav Rep*. 2020;1:100014. [FREE Full text] [doi: [10.1016/j.chbr.2020.100014](https://doi.org/10.1016/j.chbr.2020.100014)] [Medline: [34235291](https://pubmed.ncbi.nlm.nih.gov/34235291/)]
53. Grassini S. Development and validation of the AI attitude scale (AIAS-4): a brief measure of general attitude toward artificial intelligence. *Front Psychol*. 2023;14:1191628. [FREE Full text] [doi: [10.3389/fpsyg.2023.1191628](https://doi.org/10.3389/fpsyg.2023.1191628)] [Medline: [37554139](https://pubmed.ncbi.nlm.nih.gov/37554139/)]
54. Koney N, Roudenko A, Ro M, Bahl S, Kagen A. Patients want to meet with imaging experts. *J Am Coll Radiol*. 2016;13(4):465-470. [doi: [10.1016/j.jacr.2015.11.011](https://doi.org/10.1016/j.jacr.2015.11.011)] [Medline: [26857485](https://pubmed.ncbi.nlm.nih.gov/26857485/)]
55. Gunn AJ, Mangano MD, Choy G, Sahani DV. Rethinking the role of the radiologist: enhancing visibility through both traditional and nontraditional reporting practices. *Radiographics*. 2015;35(2):416-423. [doi: [10.1148/rg.352140042](https://doi.org/10.1148/rg.352140042)] [Medline: [25763726](https://pubmed.ncbi.nlm.nih.gov/25763726/)]
56. Kitts AB. Patient perspectives on artificial intelligence in radiology. *J Am Coll Radiol*. 2023;20(9):863-867. [doi: [10.1016/j.jacr.2023.05.017](https://doi.org/10.1016/j.jacr.2023.05.017)] [Medline: [37453601](https://pubmed.ncbi.nlm.nih.gov/37453601/)]

Abbreviations

AI: artificial intelligence
CT: computerized tomography
EHR: electronic health record
FDA: US Food and Drug Administration
HITL: human in the loop
LDCT: low-dose computerized tomography
MMM: medical maximizing-minimizing
PET: positron emission tomography

Edited by J Sarvestan, T Leung; submitted 14.11.24; peer-reviewed by CM Moody, JCL Chow; comments to author 03.02.25; revised version received 24.02.25; accepted 03.04.25; published 22.05.25

Please cite as:

Madanay F, O'Donohue LS, Zikmund-Fisher BJ

Patient Reactions to Artificial Intelligence–Clinician Discrepancies: Web-Based Randomized Experiment

J Med Internet Res 2025;27:e68823

URL: <https://www.jmir.org/2025/1/e68823>

doi: [10.2196/68823](https://doi.org/10.2196/68823)

PMID:

©Farrah Madanay, Laura S O'Donohue, Brian J Zikmund-Fisher. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 22.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.