

Key information influencing patient decision-making about AI in healthcare: A survey experiment

Xuan Zhu, Austin M. Stroud, Sarah A. Minter, Dong Whi Yoo, Jennifer L. Ridgeway, Maryam Mooghali, Jennifer E. Miller, Barbara A. Barry

Submitted to: Journal of Medical Internet Research
on: April 07, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
---------------------------------	----------

Preprint
JMIR Publications

Key information influencing patient decision-making about AI in healthcare: A survey experiment

Xuan Zhu¹ PhD; Austin M. Stroud¹ MA; Sarah A. Minter¹ PhD; Dong Whi Yoo² PhD; Jennifer L. Ridgeway¹ PhD; Maryam Mooghali³ MD, MSc; Jennifer E. Miller³ PhD; Barbara A. Barry¹ PhD

¹ Mayo Clinic Rochester US

² Kent State University Kent US

³ Yale University New Haven US

Corresponding Author:

Xuan Zhu PhD

Mayo Clinic
200 First Street SW
Rochester
US

Abstract

Background: Artificial Intelligence (AI)-enabled devices are increasingly used in healthcare. However, there has been limited research on patients' informational preferences, including which elements of AI device labeling enhance patient understanding, trust, and acceptance. Clear and effective patient-facing communication is essential to address patient concerns and support informed decision-making regarding AI-enabled care.

Objective: Using simulated AI device labels in a cardiovascular context, we evaluated three aims. First, we identified key information elements that influence patient trust and acceptance of an AI device. Second, we examined how these effects varied based on patient characteristics. Third, we explored how patients evaluated informational content of AI labels and their perceived effectiveness of the AI labels in informing decision-making about the use of AI device, building trust in the device, and shaping their intention to use it in their healthcare.

Methods: We recruited 340 US patients from ResearchMatch.org to participate in a web-based survey that contained two experiments. In the discrete choice experiment (DCE), participants indicated preferences in terms of trust and acceptance regarding 16 pairs of simulated AI device labels that varied across eight types of information needs identified in our previous qualitative work. In the single profile factorial experiment (SPFE), participants evaluated four randomly assigned label prototypes regarding the label's legibility, comprehensibility, information overload, credibility, and perceived effectiveness in informing about the AI device, as well as participants' trust in the AI device and intention to use the device in their healthcare. Data was analyzed using mixed effects binary or ordinal logistic regression.

Results: The DCE showed that information about regulatory approval, high device performance, provider oversight, and AI's value added to usual care significantly increased the likelihood of patient trust by 14.1-19.3% and acceptance by 13.3-17.9%. Subgroup analyses revealed variations based on patient characteristics such as familiarity with AI, health literacy, and recency of last medical checkup. The SPFE showed that patients reported good label comprehension, and that information about provider oversight, regulatory approval, device performance, and AI's added value improved perceived credibility and effectiveness of the AI label (odds ratios [ORs] range 1.35-2.05), reduced doubts in the AI device (ORs range 0.61- 0.77), and increased trust and intention to use the AI device (ORs range 1.47-1.73). However, information about data privacy and safety management protocols are less influential.

Conclusions: Patients value information about an AI device's performance, provider oversight, regulatory status, and added value during decision-making. Providing transparent, easily understandable information about these aspects is critical to support patient determinations of trust and acceptance of AI-enabled healthcare. Information elements' impact on patient trust and acceptance varies by patient characteristics, highlighting the need for a tailored approach to address the concerns of diverse patient groups about AI in healthcare. Clinical Trial: N/A

(JMIR Preprints 07/04/2025:75615)

DOI: <https://doi.org/10.2196/preprints.75615>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in a JMIR publication, my title and abstract will remain visible to all users.

No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

Key information influencing patient decision-making about AI in healthcare: A survey experiment

Xuan Zhu, PhD¹; Austin M. Stroud, MA²; Sarah A. Minter, PhD^{1,3}; Dong Whi Yoo, PhD⁴;
Jennifer L. Ridgeway, PhD^{1,5}; Maryam Mooghali, MD, MSc⁶; Jennifer E. Miller, PhD⁶;
Barbara A. Barry, PhD^{1,5}

¹ Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA

² Biomedical Ethics Research Program, Mayo Clinic, Rochester, 200 First Street SW, Rochester, MN 55905, USA

³ Department of Physical Medicine and Rehabilitation Research, 200 First Street SW, Rochester, MN 55901

⁴ School of Information, Kent State University, 1125 Risman Drive, Office 337, Kent, Ohio 44242

⁵ Division of Health Care Delivery Research, Mayo Clinic, Rochester, 200 First Street SW, Rochester, MN 55905, USA

⁶ Department of Internal Medicine, Yale School of Medicine, 330 Cedar Street, New Haven, CT 06510

Corresponding author:

Xuan Zhu, PhD

Email: zhu.xuan@mayo.edu

Address: Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA

Abstract

Background: Artificial Intelligence (AI)-enabled devices are increasingly used in healthcare. However, there has been limited research on patients' informational preferences, including which elements of AI device labeling enhance patient understanding, trust, and acceptance. Clear and effective patient-facing communication is essential to address patient concerns and support informed decision-making regarding AI-enabled care.

Objective: Using simulated AI device labels in a cardiovascular context, we evaluated three aims. First, we identified key information elements that influence patient trust and acceptance of an AI device. Second, we examined how these effects varied based on patient characteristics. Third, we explored how patients evaluated informational content of AI labels and their perceived effectiveness of the AI labels in informing decision-making about the use of AI device, building trust in the device, and shaping their intention to use it in their healthcare.

Methods: We recruited 340 US patients from ResearchMatch.org to participate in a web-based survey that contained two experiments. In the discrete choice experiment (DCE), participants indicated preferences in terms of trust and acceptance regarding 16 pairs of simulated AI device labels that varied across eight types of information needs identified in our previous qualitative work. In the single profile factorial experiment (SPFE), participants evaluated four randomly assigned label prototypes regarding the label's legibility, comprehensibility, information overload, credibility, and perceived effectiveness in informing about the AI device, as well as participants' trust in the AI device and intention to use the device in their healthcare. Data was analyzed using mixed effects binary or ordinal logistic regression.

Results: The DCE showed that information about regulatory approval, high device performance, provider oversight, and AI's value added to usual care significantly increased the likelihood of patient trust by 14.1-19.3% and acceptance by 13.3-17.9%. Subgroup analyses revealed variations based on patient characteristics such as familiarity with AI, health literacy, and recency of last medical checkup. The SPFE showed that patients reported good label comprehension, and that information about provider oversight, regulatory approval, device performance, and AI's added value improved perceived credibility and effectiveness of the AI label (odds ratios [ORs] range 1.35-2.05), reduced doubts in the AI device (ORs range 0.61- 0.77), and increased trust and intention to use the AI device (ORs range 1.47-1.73). However, information about data privacy and safety management protocols are less influential.

Conclusions: Patients value information about an AI device's performance, provider oversight, regulatory status, and added value during decision-making. Providing transparent, easily understandable information about these aspects is critical to support patient determinations of trust and acceptance of AI-enabled healthcare. Information elements' impact on patient trust and acceptance varies by patient characteristics, highlighting the need for a tailored approach to address the concerns of diverse patient groups about AI in healthcare.

Keywords: Artificial Intelligence, Health Communication, Health Decision Making, Patient Preference, Patient-Centered Care, AI Labeling



Introduction

Artificial Intelligence and machine learning (AI/ML) are increasingly being integrated into healthcare products and services due to their potential to enhance diagnostic accuracy, improve treatment planning, increase efficiency of healthcare systems, reduce costs, and ultimately improve patient outcomes.¹⁻³ Cardiology, in particular, has seen a surge in AI/ML-enabled clinical tools, aiding clinical decision-making across the care spectrum.⁴⁻⁶ While there are many promising applications of AI/ML in healthcare, effective approaches to inform patients about these technologies have not been established. This hinders informed patient decision-making and public trust in AI/ML-enabled medical devices.

Recent research on patient perspectives regarding AI in healthcare indicates that patients are generally receptive to the use of AI in their healthcare, but that certain conditions must be met.^{7,8} Patients want clear and accessible information about how AI is used in their care, its benefits and limitations, how decisions are made, and the roles of AI and healthcare providers in AI informed decisions.⁹⁻¹¹ Patients value human oversight over AI and want to ensure that healthcare decisions are ultimately made by a human.¹²⁻¹⁴ Patients also want assurance that their data is protected and used responsibly.¹⁵⁻¹⁷ Lastly, they want AI to be unbiased in its training data and outputs and desire equitable access to AI applications to prevent against underrepresentation of minoritized and disadvantaged population groups in AI research.¹⁸⁻²⁰ These nuances underscore the complexity of information needed to support informed patient use of AI and effective communication between patients, clinicians, and health systems regarding the use of AI in care.

Recent efforts to understand how to communicate transparently about AI in healthcare have focused on clinical and technical audiences. As a result, there is growing research on best practices for transparent and standardized documentation, reporting, and communication of AI/ML models for clinicians.²¹⁻²⁵ However, research efforts focusing on identifying and prioritizing patients' specific information needs remain limited.²⁶⁻²⁸ To ensure patient safety, autonomy, and informed decision-making, it is essential to understand patients' core information needs and present information in a transparent and accessible manner. Thoughtful organization and prioritization of information are key to avoiding information overload and making it easier for patients to grasp the implications of AI in their care.

The present study is part of a broader research project aimed at bridging the knowledge gap in effective communication about AI in healthcare by identifying the essential components of patient-facing information labels for these technologies. Information labels are structured summaries that provide users with key details about a device or technology, including its purpose, functionality, benefits, limitations, and risks.²⁹ The overarching goal of this project is to provide empirical evidence to inform strategies and regulatory policies that facilitate patient-centered adoption of AI in healthcare. For the purposes of our study, an AI device refers to a cardiology device equipped with ML algorithms designed to assist in decision-making across the continuum of care. Prior to this study, we conducted qualitative research, including a rapid literature review and three qualitative studies with patients and clinicians, to identify core information needs that contribute to patient and clinician trust in AI devices in healthcare, focusing on their use in diagnosing, treating, and monitoring cardiovascular conditions.^{30,31} This work identified eight elements reported by patients as influential to their trust in the AI device, including 1) data privacy and security, 2) performance, 3) AI's added value compared to usual care, 4) regulatory approval, 5) expert

endorsement, 6) generalizability and limitations, 7) device safety, and 8) healthcare provider (HCP) oversight. Building on these findings, the present study utilized two survey experiments with AI label prototype vignettes to achieve the following specific aims:

- 1) Determine the relative importance of various information elements on patient trust and acceptance of an AI device in cardiology.
- 2) Examine how the effects of information elements vary due to differences in patient characteristics, including familiarity with AI, medical mistrust, health literacy, and sociodemographic characteristics.
- 3) Assess patients' evaluation of the informational content of the AI label prototypes, perceived effectiveness of the AI label prototypes in informing decision-making about the use of the AI device, trust in the device, and intention to use the device in their healthcare.

Methods

Experimental Design

We evaluated patient's perceptions of AI information label prototypes by administering a web-based survey using two types of experimental designs: Discrete Choice Experiment (DCE) and Single Profile Factorial Experiment (SPFE). For these experiments, we created AI label prototypes based on a hypothetical cardiovascular device that varied on the following eight informational elements 1) data privacy and security, 2) performance, 3) AI's added value, 4) regulatory approval, 5) expert endorsement, 6) generalizability and limitations, 7) device safety, and 8) HCP oversight. We varied the informational content along these 8 elements based on prior qualitative research conducted with clinicians and patients, which identified these as important information needs influencing patient trust in AI.^{30,31} For experimental design efficiency, each of these elements has two levels (i.e., two variations): either in two different versions or as either present or absent. See **Table 1** for additional details about these 8 elements.

Table 1. Label information elements and levels

	Elements and definitions	Levels
	Elements with varying levels	
X1	Data privacy and security: what patient data is collected by AI; how are the patient data being collected, stored, and shared.	1. Opt-in data sharing 2. Opt-out data sharing
X2	Performance: True-positive: device accuracy when the patient DOES have A-fib True-negative: device accuracy when the patient DOES NOT have A-fib	1. High performance 2. Low performance
X3	AI's added value: improvement in patient care due to the AI-enabled device; effectiveness of the AI-enabled device in comparison to non-AI-enabled tests or conventional health care.	1. Info absent 2. Info present
X4	Regulatory approval: whether the AI device received clearance, approval, or	1. Info absent 2. Info present

	certification from a regulatory body regarding its safety and/or effectiveness.	
X5	Expert endorsement: endorsement of the device for safety and/or effectiveness issued by medical experts such as health care providers.	1. Info absent 2. Info present
X6	Generalizability and limitations: how generalizable is the device to patients with varying demographics and characteristics; what are the conditions or contexts where the device should not be used.	1. Internal validation 2. External validation
X7	Device safety: what the risk of malfunction, bugs, errors from the AI device are (both the AI algorithm and the supporting software)? How will these risks be managed?	1. Proactive auditing 2. Reactive auditing
X8	HCP oversight: whether the results or decisions from the AI-enabled device have been verified by your health care provider.	1. Info absent 2. Info present
	Elements to be displayed in the same way across all labels	
1	Purpose: what is the purpose of the AI device, for what types of patients and conditions, when and in what context should the device be used	"This AI-enabled smartwatch and smartphone app is designed to identify a potential episode of atrial fibrillation (A-fib). A-fib is an irregular and often very rapid heart rhythm that can lead to blood clots in the heart."
2	Directions: how to use the AI device, recommended actions for patients, how to interpret results (e.g., what is "normal" for pt. like me); what warrants discussion with provider; next steps by AI if any (e.g., AI will automatically alert HCP)	" • Wear your smartwatch • It will provide an alert when it detects a potential episode of A-fib • Talk to your doctor if you received an alert from the app "

Discrete Choice Experiment (DCE)

The DCE aimed to 1) determine the relative importance of different information elements influencing patient trust and acceptance of an AI device and to 2) examine how the effects of information elements vary due to differences in patient characteristics. DCEs are being increasingly used in medical and health services research to examine people's preferences regarding health care services by assessing how much they value the attributes of the service.³²⁻³⁵ DCEs ask participants to choose between two or more discrete hypothetical alternatives (e.g., Treatment A or Treatment B) consisting of multiple attributes with varying values. The choices made by participants are then analyzed to estimate the relative

influence of each attribute value on decision-making. We used a fractional factorial design to rate a selection of possible alternatives because a full factorial design (2^8 alternatives) would be infeasible to implement. To minimize the cognitive burden for participants and maximize statistical efficiency, each choice task consisted of 2 alternatives, and each participant was asked to consider 16 choice sets, comprising 32 unique label prototypes in total.³⁶ We used the R package *idfix*³⁷ to select an optimal design based on D-efficiency criterion, a commonly used metric for efficient experimental design construction to maximize statistical efficiency and precision.³⁸⁻⁴⁰ **Appendix 1** summarizes the choice sets. **Figure 1** shows an example choice set.

Figure 1. Example choice set for the discrete choice experiment

Name	A-fib Watch
Purpose	This AI-enabled smartwatch and smartphone app is designed to identify a potential episode of atrial fibrillation (A-fib). A-fib is an irregular and often very rapid heart rhythm that can lead to blood clots in the heart.
Directions	• Wear your smartwatch • It will provide an alert when it detects a potential episode of A-fib • Talk to your doctor if you received an alert from the app

	Device A	Device B
Data Privacy and Security	The information you provide in the app will be used to improve the tool. Your deidentified data will be shared unless you visit the settings menu to turn off the sharing option.	The information you provide in the app will be used to improve the tool. Your deidentified data will be shared unless you visit the settings menu to turn off the sharing option.
Performance <u>True-positive:</u> device accuracy when the patient DOES have A-fib <u>True-negative:</u> device accuracy when the patient DOES NOT have A-fib	 80% (8 out of 10) people with A-fib will be correctly identified as having A-fib.  70% (7 out of 10) people <u>without</u> A-fib will be correctly identified as NOT having A-fib.	 90% (9 out of 10) people with A-fib will be correctly identified as having A-fib.  80% (8 out of 10) people <u>without</u> A-fib will be correctly identified as NOT having A-fib.
Added value	Appropriate use of the tool, response to alerts, and treatment may reduce risk of heart failure and stroke.	
Regulatory approval	This device acquired regulatory authorization, which means that this device is as safe and effective as other regulated device(s).	
Expert endorsement	A panel of doctors has reviewed and endorsed this device for use.	A panel of doctors has reviewed and endorsed this device for use.
Validation	An independent research lab unaffiliated with the clinic where you receive care has tested and evaluated this device.	An independent research lab unaffiliated with the clinic where you receive care has tested and evaluated this device.
Tool safety	If doctors or patients report any issues with the device, the developer team will assess and fix them promptly.	This device will be audited by the developer team every 4 weeks.
Oversight from your doctor		Your doctor will review the AI results along with your medical history.

Single Profile Factorial Experiment (SPFE)

The SPFE aimed to assess how different information elements influence patients' evaluation of the AI label, their perception of its effectiveness in supporting decision-making about the AI device, their trust in the device, and their likelihood of allowing it to be used in their healthcare. This design was chosen for two key reasons. First, the single profile design – where patients evaluate AI device profiles one at a time, mirrors actual patient decision-making regarding AI device use in healthcare more closely than DCE. Realistically,

patients are more likely to be facing the decision of whether to trust one given AI device rather than comparing multiple AI devices side by side that are designed for the same purpose. Second, the SPFE allows rating-based outcome measures (e.g., 1=Very unlikely to accept to 5=Very likely to accept) that provide more direct and fine-grained insight into participants' evaluation of each alternative. The SPFE used a 2^{8-3} resolution IV fractional factorial design to evaluate the label prototypes. We used the R package FrF2⁴¹ to generate the 32 experimental conditions (**Appendix 2**). Each participant was randomly assigned to evaluate four label prototypes displayed in a randomized order. After viewing each label prototype, participants answered a series of questions on label evaluation, AI trust, and behavioral intention using 5-point Likert-style scales. **Figure 2** shows an example label prototype.

Figure 2. Example label prototype for the single profile factorial experiment

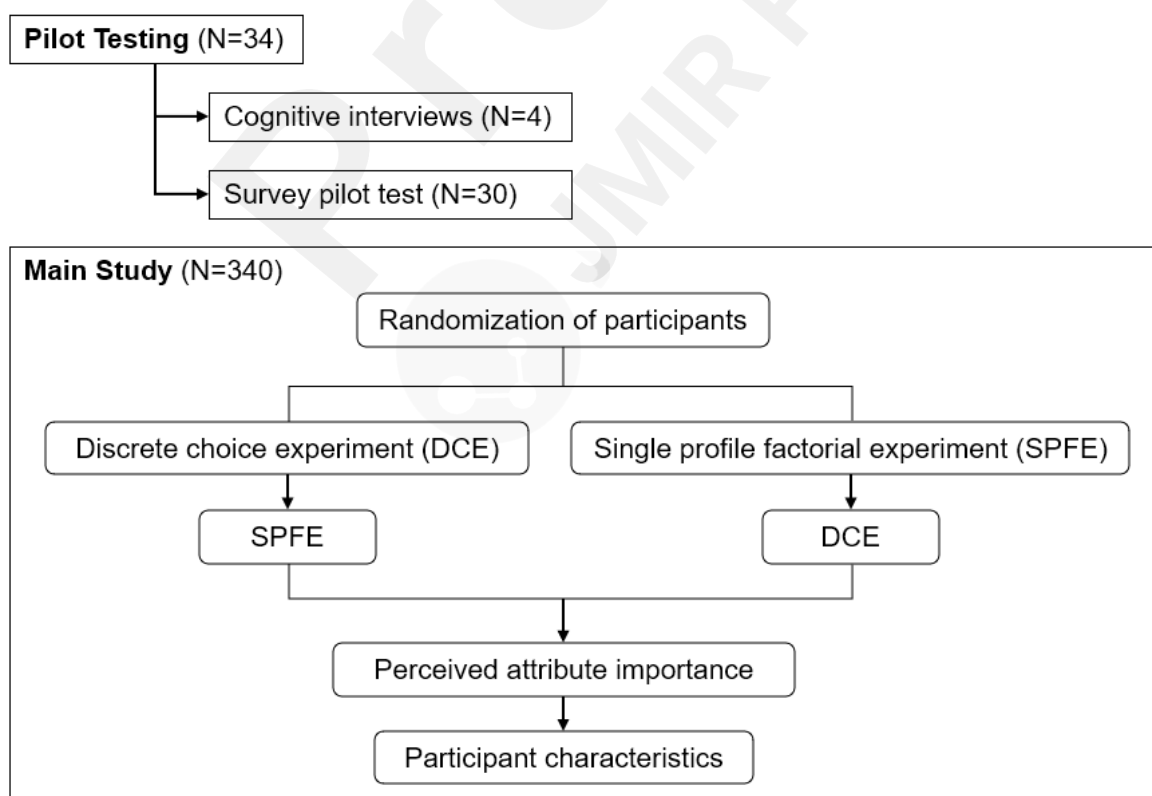
Name	A-fib Watch
Purpose	This AI-enabled smartwatch and smartphone app is designed to identify a potential episode of atrial fibrillation (A-fib). A-fib is an irregular and often very rapid heart rhythm that can lead to blood clots in the heart.
Directions	• Wear your smartwatch • It will provide an alert when it detects a potential episode of A-fib • Talk to your doctor if you received an alert from the app
Data Privacy and Security	The information you provide in the app will be used to improve the tool. Your deidentified data will be shared unless you visit the settings menu to turn off the sharing option.
Performance <u>True-positive:</u> device accuracy when the patient DOES have A-fib <u>True-negative:</u> device accuracy when the patient DOES NOT have A-fib	 90% (9 out of 10) people with A-fib will be correctly identified as having A-fib.  80% (8 out of 10) people <u>without</u> A-fib will be correctly identified as NOT having A-fib.
Added value	Appropriate use of the tool, response to alerts, and treatment may reduce risk of heart failure and stroke.
Regulatory approval	This device acquired regulatory authorization, which means that this device is as safe and effective as other regulated device(s).
Expert endorsement	A panel of doctors has reviewed and endorsed this device for use.
Validation	An independent research lab unaffiliated with the clinic where you receive care has tested and evaluated this device.
Tool safety	If doctors or patients report any issues with the device, the developer team will assess and fix them promptly.
Oversight from your doctor	

Pilot Testing

We conducted cognitive interviews (via Zoom, 30-45 minutes in length) with four participants to assess whether the information elements, variations, presentation, and survey questions were easy to understand and meaningful to participants. The interviewer shared the web-based survey with participants via shared screen and asked participants to verbally answer the survey questions and share their thought processes. The interviewer also asked participants to explain the information elements in their own words. The interviews were audio-recorded, and the interviewer took notes on reflections and observations. We revised the survey questionnaire based on interviewer's notes.

Following the cognitive interviews, we pilot tested the revised Qualtrics web-based survey with 30 participants to examine the feasibility and data quality of the two survey experiments. The survey consisted of four sections: the DCE, the SPFE, a ranking of information elements, and participant characteristics. The presentation order of the two experiments were randomized to mitigate order effect. For the DCE, participants first completed a warm-up task to familiarize them with the DCE procedure. They then proceeded to the main DCE on AI information labeling, where they completed 18 choice tasks, including 16 experimental choice tasks and two validity check tasks. The order of the 16 experimental choice tasks was randomized. In the SPFE, participants were randomly assigned to evaluate four of 32 label prototypes, presented in a random order. For each label prototype, participants answered questions on label evaluation, trust in the AI device, and intention to use the device in healthcare. After completing the experiments, participants ranked the importance of the aforementioned information elements. Finally, participants provided information on their sociodemographic and healthcare characteristics. **Figure 3** summarizes study procedure.

Figure 3. Study procedure



Main Study

For the main study, participants completed the web-based Qualtrics survey we had piloted.

Measures

Label preference was measured through participants selecting their preferred label prototype in a choice task based on trust ("Which device would you trust more?") and acceptance ("If you were given the option, which device would you be more likely to use in your health care?").

Label evaluation included six items on a 5-point Likert scale (1=Strongly disagree to 5=Strongly agree) that assessed 1) legibility ("This label is easy to read"), 2) comprehensibility ("This label is easy to understand"), 3) information overload ("Reading this label is too mentally demanding for me"), 4) credibility ("I trust the information on this label"), and 5) effectiveness of the label in informing patients about the AI device ("This label gives me all the information I need about the AI device" and "This label helps me decide whether the AI device should be used in my care").

Participants' *trust in the AI device* was measured with three items on a 5-point Likert scale (1=Strongly disagree to 5=Strongly agree), including 1) "I would trust the results from this AI device," 2) "I would have doubts about this AI device," and 3) "I would seek a second opinion." Participants' likelihood of using the device was assessed on a 5-point Likert-type scale (1=Very unlikely to 5=Very likely). Additionally, participants provided feedback on each label prototype through an optional open-ended question.

Participant characteristics were measured in terms of the need for cognition, which reflects motivation to process complex information and engage in effortful thinking;^{42,43} familiarity with AI; medical mistrust;⁴⁴ health literacy;⁴⁵⁻⁴⁷ last routine medical checkup; health insurance coverage; and demographics (e.g., age, gender, race/ethnicity, education level, and household financial status).⁴⁸

Sample Size

We conducted an a priori power analysis to estimate the smallest sample size required for the main study. Because the DCE is more statistically efficient than the SPFE, we based our estimation on the SPFE design. The analysis showed that a sample of 288 participants would allow us to detect a main effect with a standardized regression coefficient of 0.167 or higher with 80% power at an alpha level of 0.05, which is sufficient for the DCE.⁴⁹ The total sample size needed for the pilot and main studies is 323. We planned to recruit 350 participants in total to account for attrition and incomplete responses.

Recruitment

We recruited participants from ResearchMatch⁵⁰, a national health volunteer registry that was created by several academic institutions and supported by the U.S. National Institutes of Health as part of the Clinical Translational Science Award (CTSA) program. ResearchMatch has a large population of volunteers who have consented to be contacted

by researchers about health studies for which they may be eligible. Inclusion criteria for this study include being an adult aged 18 or older, being proficient in reading and writing in English, and having had a primary care or cardiology care visit within the past three years. We oversampled racial and ethnic minority patients to ensure their representation was comparable to that of non-Hispanic white patients. This supported our ability to draw meaningful conclusions about diverse patient populations while accounting for historical underrepresentation in health research. Participants received a message describing our study via the ResearchMatch platform, where they could then decide to opt-in to be contacted and receive a link to our survey. ResearchMatch provided contact information (i.e., name and email address) for participants who opted in. Participants who opted-in received an email invitation with a link to complete the online survey. Respondents received a \$10 gift card as remuneration for completing the survey. Review and approval for this study and all procedures were obtained from the Mayo Clinic Institutional Review Board.

Statistical Analysis

We used mixed effects binary logistic regression to analyze the data from the DCE. Our analysis focused on how different information factors influenced the probability of individuals trusting and accepting the AI device in their healthcare. We included individual-specific intercepts to account for heterogeneity in individual preferences. We reported average marginal effect (AME), which shows the average change in predicted probabilities (percentage point increase or decrease) of the outcome variable across all participants when moving from one level of the information factor to another, while holding all other variables constant. Additionally, we conducted subgroup analyses to explore how the effect of top preferred information factors on patient trust and acceptance of the AI device varied based on participants' AI familiarity, openness toward medical AI, health literacy, medical mistrust, and sociodemographic characteristics.

We used mixed effects ordinal logistic regression (cumulative link mixed model) to analyze the data from the SPFE. Our focus was on how different information factors influenced participants' evaluation of the AI label's legibility, comprehensibility, information overload, information credibility, perceived effectiveness in informing about the AI device, trust in the AI device, and intention to use it in healthcare. We included individual-specific intercepts to account for repeated measurements within individuals. For models on AI device trust and intention to use, we adjusted for participants' evaluation of the AI label's legibility, comprehensibility, information overload, information credibility, and perceived effectiveness in informing about the AI device. We used "flexible" thresholds in our models to allow the distance between the four cut-points of the 5-point scales to vary freely. We reported odds ratios (OR) for the effect of information factors.

All analyses were performed in R version 4.3.1.⁵¹ "tidyverse"⁵² was used for data wrangling and visualization, "lme4"⁵³ was used for fitting mixed effects binary logistic models, "ordinal"⁵⁴ was used for fitting mixed effects ordinal logistic models, and "marginaleffects"⁵⁵ was used for calculating AME.

Results

A total of 340 participants who completed at least 75% of the survey questions were included in the analysis. **Table 2** summarizes participant characteristics. Participants were

aged 18-34 (37%), 35-54 (34.75%), and 55 or older (28.05%). Nearly half were women (48.47%), 48.77% were men, and 2.76% identifying as non-binary or another gender. The majority were non-Hispanic white (63%), followed by Black/African American (23.24%), Hispanic/Latinx (13.46%), Asian/Asian American (9.17%), and Native American/Indigenous (3.98%). Regarding education, 49.09% had a bachelor's degree, 28.96% held a master's degree or higher, and 20.73% had some college or an associate degree. Financially, 53.54% reported having disposable income after paying bills, 29.85% had little spare money, and 16.62% struggled to pay bills.

Most participants had private health insurance (65.51%), while 32.59% had public coverage. Additionally, 66.88% had a routine medical checkup within the past year.

Table 2. Participant characteristics

	Total (N=340)
	N (%)
Age in years	
18-24	18 (5.49)
25-34	104 (31.71)
35-44	72 (21.95)
45-54	42 (12.80)
55-64	52 (15.85)
65 and above	40 (12.20)
Missing	12
Gender	
Man	159 (48.77)
Woman	158 (48.47)
Non-binary/other	9 (2.76)
Missing	14
Race and ethnicity ^a	
White / Caucasian	206 (63.00)
Black / African American	76 (23.24)
Hispanic / Latina / Latino	44 (13.46)
Asian / Asian American	30 (9.17)
Native American / American Indian / Alaska Native / Indigenous	13 (3.98)
Middle Eastern / North African / Arab American	8 (2.45)
Native Hawaiian / Pacific Islander	2 (0.61)
Other race	6 (1.83)
Missing	13
Education level	
High school or less	4 (1.22)
Some college or Associate's degree	68 (20.73)
Bachelor's degree	161 (49.09)
Master's degree or higher	95 (28.96)
Missing	12
Perceived household financial status	
After paying the bills, I still have enough money for special things that I want	174 (53.54)
Have enough money to pay the bills, but little spare money to	97 (29.85)

buy extra or special things	
Have enough money to pay the bills, but only because I have cut back on things	28 (8.62)
Having difficulty paying the bills, no matter what I do	26 (8.00)
Missing	15
Health insurance coverage	
Private	207 (65.51)
Public	103 (32.59)
No insurance	6 (1.90)
Missing	24
Last routine checkup	
Less than a year ago	212 (66.88)
More than 1 year, less than 2 years	53 (16.72)
More than 2 years, less than 5 years	41 (12.93)
5 or more years	6 (1.89)
Never had routine checkup	5 (1.58)
Missing	23

a The percentages may add over 100% because participants can select multiple races.

DCE Results

Importance of the information factors for patients' trust in the AI device

Figure 4 visualizes the relative importance of the information factors for participants' trust in the AI device. The four information factors that produced the greatest increase in participants' trust in AI devices were inclusion of information about regulatory approval (AME=19.34%, 95% Confidence Interval [CI]: 15.97%, 22.72%), information about high versus low performance (AME=16.62%, 95% CI: 14.40%, 18.85%), inclusion of information about HCP oversight (AME=15.50%, 95% CI: 12.56%, 18.44%), and inclusion of information about the AI's added value compared to usual care (14.06% increase, 95% CI: 12.16%, 15.96%). Including information about opt-in (versus opt-out) data privacy protocol, expert endorsement (versus information absent), and information about external (versus internal) validation less strongly increased the probability of trusting the AI device (9%, 7.32%, and 2.76%, respectively). Effect of information about proactive versus reactive device safety management protocol on participants' trust was not statistically significant (**Table 3**).

Figure 4. Difference in probability of AI device being trusted by attribute level

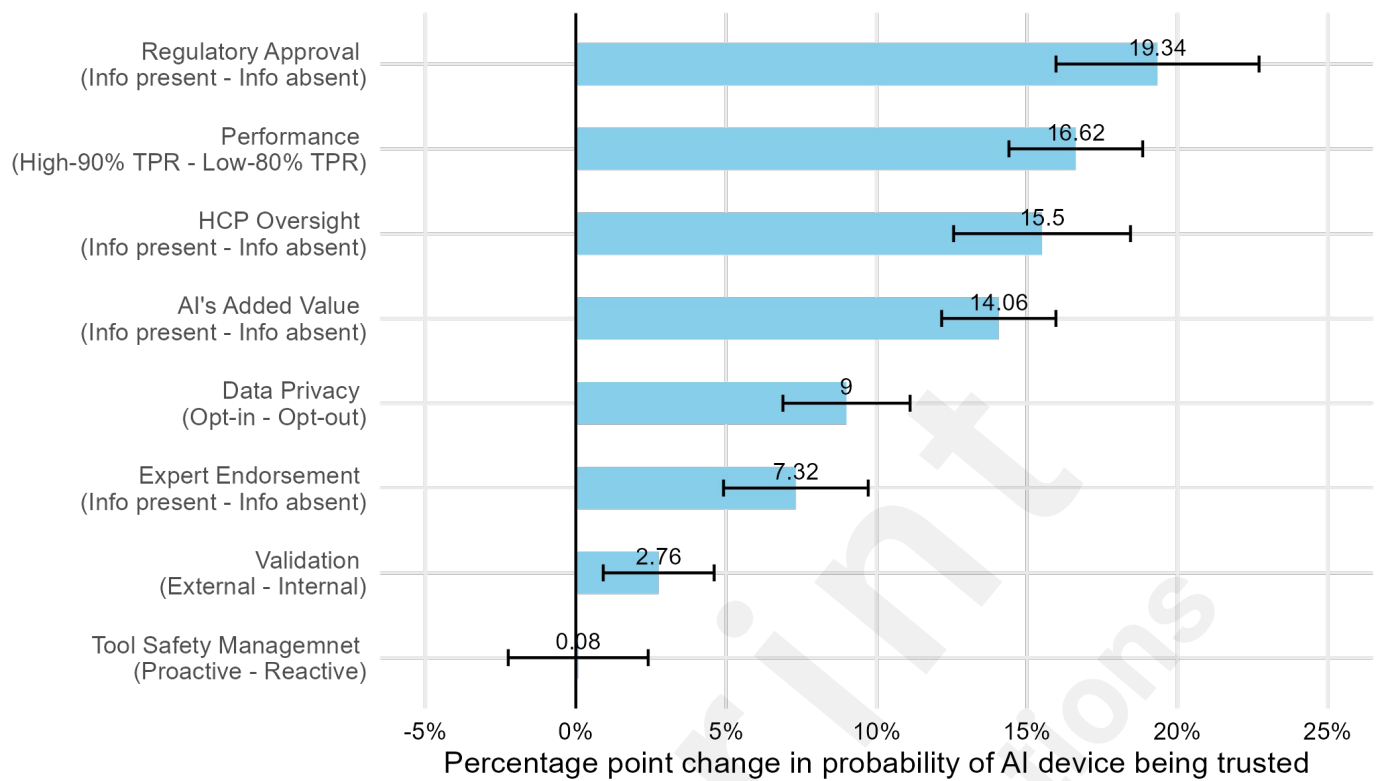


Table 3. Effects of information elements on the probability of the AI device being trusted and accepted

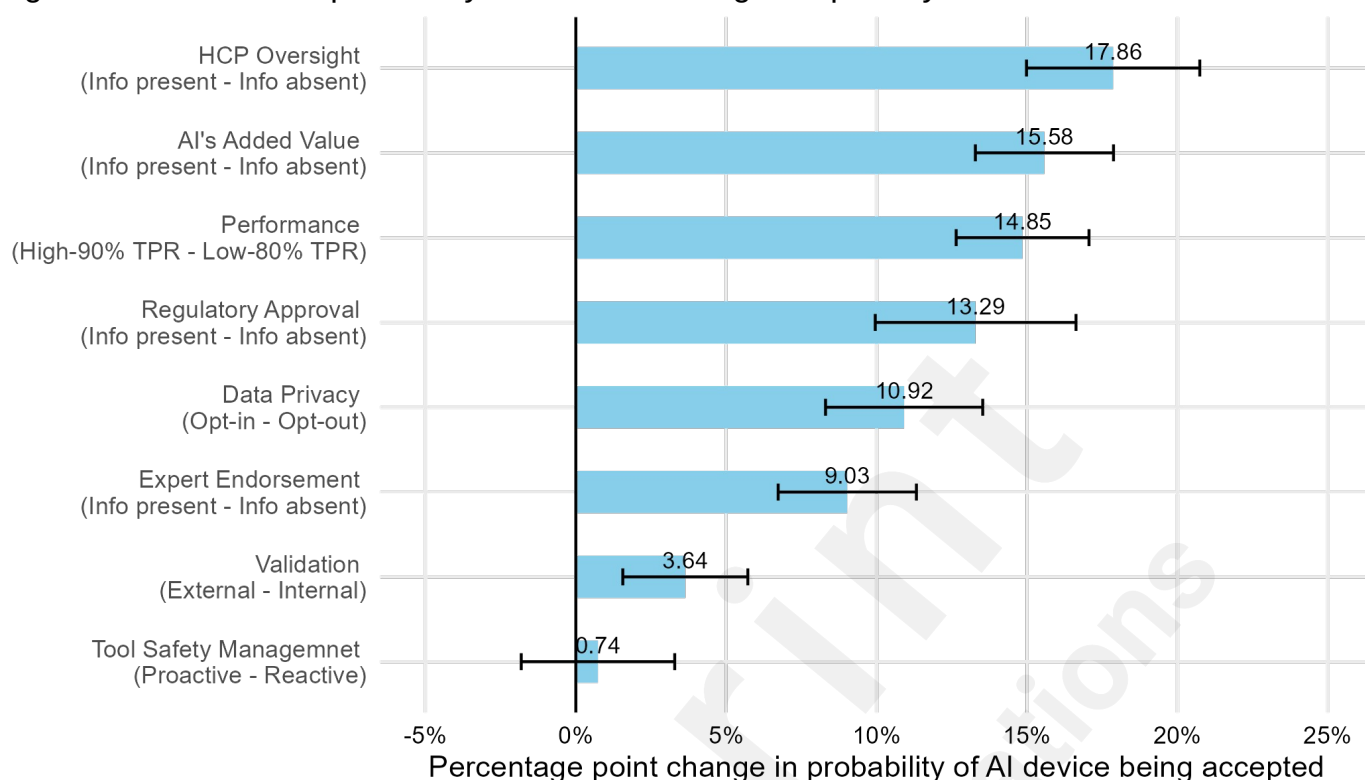
	Trust	Acceptance
	AME ^a (95% CI)	AME ^a (95% CI)
	Percentage points	Percentage points
Added value (Info present – info absent)	14.06 (12.16, 15.96)	15.58 (13.28, 17.88)
Data privacy (Opt-in – opt-out)	9.00 (6.89, 11.11)	10.92 (8.30, 13.53)
Expert endorsement (Info present – info absent)	7.32 (4.91, 9.72)	9.03 (6.73, 11.32)
HCP oversight (Info present – info absent)	15.50 (12.56, 18.44)	17.86 (14.98, 20.75)
Performance (High – low)	16.62 (14.40, 18.85)	14.85 (12.65, 17.06)
Regulatory approval (Info present – info absent)	19.34 (15.97, 22.72)	13.29 (9.96, 16.63)
Device safety (Proactive – reactive)	0.08 (-2.25, 2.41)	0.74 (-1.82, 3.29)
Validation (External – internal)	2.76 (0.91, 4.60)	3.64 (1.56, 5.72)

a AME: average marginal effect. It is the average change in the predicted probabilities (percentage point increase or decrease) of the AI device being trusted/accepted across all participants when moving from one level of the information factor to the other, keeping all other variables in the model constant.

Importance of the information factors for patients' acceptance of the AI device

Figure 5 visualizes the relative importance of the information factors for participants' willingness to use the AI device in their healthcare (i.e., acceptance). The four information factors that produced the greatest increase in participants' acceptance in using AI devices were inclusion of information about HCP oversight (17.86% increase, 95% CI: 4.98%, 20.75%), inclusion of information about the device's added value (15.58% increase, 95% CI: 13.28%, 17.88%), information about high versus low device performance (14.85% increase, 95% CI: 12.65%, 17.06%), and inclusion of information about regulatory approval (13.29% increase, 95% CI: 9.96%, 16.63%). Including information about opt-in (versus opt-out) data privacy protocol, expert endorsement (versus information absent), and information about external (versus internal) validation less strongly increased the probability of the AI device being accepted (10.92%, 9.03%, and 3.64%, respectively). Effect of information about proactive versus reactive device safety management protocol on device acceptance was not statistically significant (**Table 3**).

Figure 5. Difference in probability of AI device being accepted by attribute level



Subgroup differences in the effect of top preferred information factors on patient trust

Our analyses revealed subgroup differences in how the top four most important information factors affected trust in AI devices, based on participants' level of familiarity with AI, level of reading health literacy, level of need for cognition, recency of last routine medical checkup, age group, and gender (**Appendix 3**).

AI's Added Value

Specifically, having information about AI's added value (versus information absent) had a stronger positive effect on participants' trust in AI devices with a high (versus low) level of need for cognition (AME=16.21% versus 12.23%, $p<0.05$).

HCP Oversight

Having information about HCP oversight (versus information absent) had a stronger positive effect on the probability of AI device being trusted for participants who reported not at all to somewhat (versus very or extremely) familiar with AI (AME=19.27% versus 11.30%, $p<0.05$), for participants whose last routine checkup was less than a year ago (versus one or more years ago) (AME=18.27% versus 11.63%, $p<0.05$), and for participants with age 55 or older (versus age 35-54 or age 18-34) (AME=23.44% versus 10.34% or 15.25%, p -values <0.05).

High device performance

Information about high (versus low) device performance had a stronger positive effect on the probability of AI device being trusted for participants with a high (versus low) level of reading health literacy (AME=22.02% versus 11.06%, $p<0.05$), for participants whose last

routine checkup was less than a year ago (versus one or more years ago) (AME=20.09% versus 10.38%, $p<0.05$), and for participants who identified as women (versus men) (AME=19.38% versus 12.85%, $p<0.05$).

Regulatory Approval

Having information about regulatory approval (versus information absent) had a stronger positive effect on the probability of AI device being trusted for participants with a high (versus low) level of reading health literacy (AME=25.07% versus 13.63%, $p<0.05$) and for participants whose last routine checkup was less than a year ago (versus one or more years ago) (AME=23.29% versus 12.92%, $p<0.05$).

Subgroup differences in the effect of top preferred information factors on patient acceptance

There were subgroup differences in how the top four most important information factors affected acceptance of AI device, based on participants' reported level of familiarity with AI, level of reading health literacy, level of numeracy, recency of last routine medical checkup, gender, and race/ethnicity (**Appendix 4**).

AI's Added Value

Specifically, having information about the device's added value (versus information absent) had a stronger positive effect on the probability of the AI being accepted for participants who reported being not at all to somewhat familiar (versus very or extremely) with AI (AME=18.06% versus 11.49%, $p<0.05$), for participants with a high (versus low) level of reading health literacy (AME=19.34% versus 11.96%, $p<0.05$), for participants whose last routine checkup was within the last year (versus one or more years ago) (AME=19.19% versus 8.81%, $p<0.05$), for participants who identified as women (versus men) (AME=18.94% versus 11.38%, $p<0.05$), and for participants who identified as a person of color (versus Non-Hispanic white) (AME=19.64% versus 12.01%, $p<0.05$).

HCP oversight

Providing information about HCP oversight (versus information absent) had a stronger positive effect on the probability of AI device being accepted for participants who reported not at all to somewhat familiar (versus very or extremely) with AI (AME=25.34% versus 9.59%, $p<0.01$), for participants whose last routine checkup was within the last year (versus one or more years ago) (AME=21.92% versus 13.31%, $p<0.05$), and for participants with age 55 or older (versus age 35-54 or age 18-34) (AME=30.13% versus 13.05% or 13.77%, p -values <0.05).

High device performance

Information about high (versus low) device performance had a stronger positive effect on the probability of AI device being accepted for participants who reported not at all to somewhat familiar (versus very or extremely) with AI (AME=18.18% versus 10.58%, $p<0.05$), for participants with a high (versus low) level of reading health literacy (AME=21.41% versus 7.66%, $p<0.05$), for participants with a high (versus low) level of numeracy (AME=19.38% versus 9.77%, $p<0.05$), for participants whose last routine checkup was within the last year (versus one or more years ago) (AME=18.91% versus 7.03%, $p<0.05$), for participants who identified as women (versus men) (AME=18.71% versus 9.57%, $p<0.05$), and for participants with age 55 or older (versus age 35-54 or age 18-34) (AME=20.88% versus 12.80% or 11.82%, p -values <0.05).

Regulatory Approval

Providing information about regulatory approval (versus information absent) had a stronger positive effect on the probability of AI device being accepted for participants with a high (versus low) level of reading health literacy (AME=19.93% versus 5.82%, $p<0.05$), for participants with a high (versus low) level of numeracy (AME=15.79% versus 9.16%, $p<0.05$), for participants whose last routine checkup was within the last year (versus one or more years ago) (AME=18.65% versus 3.96%, $p<0.01$), and for participants who identified as women (versus men) (AME=16.58% versus 9.43%, $p<0.05$).

SPFE Results

Table 4 summarizes the descriptive statistics of the AI label evaluation measures. Most participants reported that the AI labels were easy to read (perceived legibility; 86.7%) and understand (comprehensibility; 84.2%). They also reported that reading the labels was not too mentally demanding (no information overload; 63.9%). Effects of the information factors on perceived legibility, comprehensibility, and information overload were not statistically significant (**Table 5**).

Table 4. Descriptive statistics of the AI label evaluation measures

Measure	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
	N (%)	N (%)	N (%)	N (%)	N (%)
This label is easy to read (Perceived legibility)	15 (1.12)	67 (5.00)	96 (7.16)	474 (35.37)	688 (51.34)
This label is easy to understand (Comprehensibility)	18 (1.34)	73 (5.44)	121 (9.02)	498 (37.14)	631 (47.05)
Reading this label is too mentally demanding for me (Information overload)	524 (39.13)	332 (24.79)	125 (9.34)	202 (15.09)	156 (11.65)
I trust the information on this label (Information credibility)	30 (2.24)	94 (7.01)	220 (16.41)	597 (44.52)	400 (29.83)
Label gives all information needed about the AI device (Perceived label effectiveness)	44 (3.28)	243 (18.13)	177 (13.21)	502 (37.46)	374 (27.91)
Label helps me decide whether the device should be used in my care (Perceived label effectiveness)	15 (1.12)	83 (6.19)	162 (12.09)	600 (44.78)	480 (35.82)
I would trust the results from this AI device	38 (2.83)	138 (10.29)	218 (16.26)	642 (47.87)	305 (22.74)
I would have doubts about this AI device	128 (9.55)	351 (26.17)	274 (20.43)	405 (30.20)	183 (13.65)
I would seek a second opinion	84 (6.26)	162 (12.08)	256 (19.09)	489 (36.47)	350 (26.10)
	Very unlikely	Somewhat unlikely	Undecided	Somewhat likely	Very likely
	N (%)	N (%)	N (%)	N (%)	N (%)
Use the AI device if offered the option	22 (1.64)	125 (9.31)	229 (17.06)	521 (38.82)	445 (33.16)

Table 5. Effects of information element on perceived legibility, comprehensibility, information overload, information credibility, and perceived effectiveness of the label in informing decision-making

	Perceived legibility ^a	Comprehensibility ^a	Information overload ^a	Information credibility ^a	Perceived effectiveness of the label in informing decision-making ^b	
					Label gives all information needed about the AI device	Label helps me decide whether the device should be used in my care
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Added value						
Info present	0.94 (0.70, 1.27)	1.17 (0.88, 1.56)	0.79 (0.58, 1.05)	1.35 (1.05, 1.73)	1.36 (1.06, 1.74)	1.15 (0.89, 1.48)
Info absent	Reference	Reference	Reference	Reference	Reference	Reference
Data privacy						
Opt-in	1.11 (0.83, 1.50)	1.07 (0.80, 1.42)	0.84 (0.63, 1.13)	1.09 (0.85, 1.39)	0.71 (0.56, 0.91)	0.97 (0.75, 1.24)
Opt-out	Reference	Reference	Reference	Reference	Reference	Reference
Expert endorsement						
Info present	1.21 (0.90, 1.63)	0.90 (0.68, 1.21)	0.87 (0.65, 1.16)	1.37 (1.07, 1.76)	1.09 (0.85, 1.39)	1.27 (0.98, 1.63)
Info absent	Reference	Reference	Reference	Reference	Reference	Reference
HCP oversight						
Info present	1.11 (0.83, 1.50)	1.13 (0.85, 1.51)	0.88 (0.66, 1.18)	1.60 (1.25, 2.05)	1.48 (1.16, 1.89)	1.41 (1.09, 1.81)
Info absent	Reference	Reference	Reference	Reference	Reference	Reference
Performance						
High	1.21 (0.90, 1.62)	1.15 (0.87, 1.52)	0.84 (0.63, 1.11)	1.45 (1.14, 1.85)	1 (0.78, 1.27)	1.16 (0.91, 1.48)
Low	Reference	Reference	Reference	Reference	Reference	Reference
Regulatory approval						
Info present	1.16 (0.86, 1.58)	1.29 (0.97, 1.72)	0.99 (0.74, 1.34)	1.55 (1.21, 1.99)	2.05 (1.60, 2.63)	1.11 (0.87, 1.43)
Info absent	Reference	Reference	Reference	Reference	Reference	Reference
Device safety						
Proactive	0.92 (0.68, 1.24)	0.91 (0.69, 1.21)	1 (0.75, 1.33)	1.03 (0.80, 1.31)	0.90 (0.71, 1.15)	1.09 (0.85, 1.39)
Reactive	Reference	Reference	Reference	Reference	Reference	Reference
Validation						
External	0.98 (0.73, 1.32)	1.13 (0.85, 1.51)	1.01 (0.76, 1.35)	1.01 (0.79, 1.29)	1.14 (0.89, 1.45)	1.03 (0.80, 1.32)
Internal	Reference	Reference	Reference	Reference	Reference	Reference

a Model adjusted for patient characteristics that were statistically significantly associated with the outcome.

b Model adjusted for perceived eligibility, comprehensibility, information overload, information credibility, and patient characteristics that were statistically significantly associated with the outcome.

Credibility

Having information about added value (OR=1.35, 95% CI: 1.05, 1.73), expert endorsement (OR=1.37, 95% CI: 1.07, 1.76), HCP oversight (OR=1.60, 95% CI: 1.25, 2.05), high (versus low) performance (OR=1.45, 95% CI: 1.14, 1.85), and regulatory approval (OR=1.55, 95% CI: 1.21, 1.99) were associated with higher levels of perceived credibility of the information (**Table 5**).

Perceived effectiveness of label

Information about added value (OR=1.36, 95% CI: 1.06, 1.74), HCP oversight (OR=1.48, 95% CI: 1.16, 1.89), and regulatory approval (OR=2.05, 95% CI: 1.60, 2.63) were associated with higher likelihood of participants reporting that the label gives them all the information they need about the AI device.

Information about opt-in versus opt-out data privacy protocol (OR=0.71, 95% CI: 0.56, 0.91) were associated with lower likelihood of participants reporting that the label gives them all the information they need about the AI device.

Information about HCP oversight (OR=1.41, 95% CI: 1.09, 1.81) was associated with higher likelihood of participants reporting that the AI label helps them decide whether the AI device should be used in their care (**Table 5**).

Trust and intentions to use AI

Information about expert endorsement (OR=1.30, 95% CI: 1.01, 1.68) and high (versus low) performance (OR=1.48, 95% CI: 1.16, 1.90) were associated with higher levels of trust in the results from the AI device. Information about high (versus low) performance (OR=0.77, 95% CI: 0.61, 0.98) and regulatory approval (OR=0.61, 95% CI: 0.48, 0.78) were associated with lower likelihood of participants reporting having doubts in the AI device. Information about HCP oversight (OR=0.75, 95% CI: 0.57, 0.97) and regulatory approval (OR=0.63, 95% CI: 0.48, 0.83) were associated with lower likelihood of participants reporting needing a second opinion. Information about HCP oversight (OR=1.47, 95% CI: 1.12, 1.94), high (versus low) performance (OR=1.59, 95% CI: 1.22, 2.07), and regulatory approval (OR=1.73, 95% CI: 1.31, 2.30) were associated with higher intention to use the AI device if offered the option (**Table 6**).

Table 6. Effects of information element on patient trust in the AI device and intention to use the AI device

	Trust in the AI device ^a			Intention to use the AI device if offered the option ^a
	I would trust results from this AI device	I would have doubts about this AI device	I would seek a second opinion	
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Added value				
Info present	0.83 (0.64, 1.06)	1.15 (0.90, 1.47)	0.87 (0.67, 1.14)	0.99 (0.75, 1.31)
Info absent	Reference	Reference	Reference	Reference
Data privacy				
Opt-in	1.07 (0.83, 1.37)	1.07 (0.85, 1.36)	1.27 (0.98, 1.65)	1.03 (0.79, 1.34)
Opt-out	Reference	Reference	Reference	Reference
Expert endorsement				
Info present	1.30 (1.01, 1.68)	0.89 (0.70, 1.13)	0.93 (0.71, 1.21)	1.17 (0.89, 1.54)
Info absent	Reference	Reference	Reference	Reference
HCP oversight				
Info present	1.21 (0.94, 1.55)	0.80 (0.63, 1.03)	0.75 (0.57, 0.97)	1.47 (1.12, 1.94)

Info absent	Reference	Reference	Reference	Reference
Performance				
High	1.48 (1.16, 1.90)	0.77 (0.61, 0.98)	0.80 (0.62, 1.04)	1.59 (1.22, 2.07)
Low	Reference	Reference	Reference	Reference
Regulatory approval				
Info present	1.26 (0.98, 1.64)	0.61 (0.48, 0.78)	0.63 (0.48, 0.83)	1.73 (1.31, 2.30)
Info absent	Reference	Reference	Reference	Reference
Device safety				
Proactive	1.08 (0.84, 1.38)	1.13 (0.89, 1.43)	0.98 (0.76, 1.27)	0.93 (0.71, 1.21)
Reactive	Reference	Reference	Reference	Reference
Validation				
External	1.06 (0.83, 1.36)	1.12 (0.88, 1.42)	0.80 (0.61, 1.03)	0.97 (0.74, 1.27)
Internal	Reference	Reference	Reference	Reference

a Model adjusted for perceived eligibility, comprehensibility, information overload, information credibility, perceived effectiveness of the label in informing decision-making, and patient characteristics that were statistically significantly associated with the outcome.

Discussion

To our knowledge, this is the first study to apply two experimental methods, DCE and SPFE, to elicit patient preferences for information about the use of AI devices in their healthcare. Our study provides important evidence and insights for healthcare and AI professionals and policymakers on the relative importance of various information factors to patient decision-making, trust, and acceptance regarding an AI device in cardiology and on subgroup differences in the impact of top preferred information factors. This work is a first step towards developing effective communication strategies about AI in healthcare that ensures transparency and accessibility to facilitate informed decision-making and patient-centered adoption of AI applications.

Results from the two experiments offer complementary insights into the key factors shaping patients' decision-making about use of an AI device in cardiology. Results from the SPFE showed that most participants found the AI label prototypes easy to read and understand, and not mentally demanding. The effects of information factors on perceived legibility, comprehensibility, and information overload were not statistically significant, indicating that patient preferences were not due to a lack of understanding of these labels. Results from the DCE showed that information about provider oversight, regulatory approval, and high device performance were the most influential in increasing patient trust in the AI device and their willingness to use the device in their healthcare. Patients placed less importance on information about external versus internal validation protocol, proactive versus reactive device safety management protocol, and opt-in versus opt-out data privacy protocol. Results from the SPFE reinforced these findings. First, information about the device's added value, expert endorsement, HCP oversight, high performance, and regulatory approval were linked to higher perceived credibility of the AI label. Additionally, information about added value, HCP oversight, and regulatory approval increased the likelihood that participants felt they had all the necessary information about the device. Interestingly, information about HCP oversight was the only factor that significantly influenced participants' perception of the label's usefulness in deciding whether to use the device in their care. Information about expert endorsement, high performance, and regulatory approval were also associated with greater trust in the AI device's results and reduced doubts about the device. Further, information about HCP oversight and regulatory approval lowered the likelihood of participants seeking a second opinion. Lastly, information about HCP oversight, high performance, and regulatory approval contributed to a higher intention to use the AI device if offered.

Implications for Practice

These findings have important implications for the implementation of AI in healthcare. To start, providing transparent and accessible information about HCP oversight, regulatory approval, device performance, and the device's added value is critical for building patient trust and acceptance of AI technology.²⁷ These elements emerged as the most influential in shaping patients' willingness to use an AI device, highlighting the need for

healthcare systems to clearly communicate these aspects when introducing AI devices to patients. Healthcare systems could consider developing decision aids that allow patients to explore different aspects of a specific AI device in collaboration with their providers, helping them weigh the benefits and risks of using the device in their care.²⁶

Our findings also highlight the critical importance of information about endorsement and oversight from regulatory bodies and healthcare providers in boosting patients' confidence and reducing doubts about the AI device's effectiveness and safety, which can lead to higher trust in the device and lower likelihood of needing second opinions. Notably, information about HCP oversight was particularly influential in helping patients decide whether an AI device should be used in their care, consistent with previous research that underscores the value of a "human in the loop" in AI implementation and adoption^{12-14,18} This finding suggests that involving healthcare providers in the AI implementation process and emphasizing their oversight role in patient-facing communications may improve patient engagement and acceptance of AI devices. It is important to acknowledge that oversight occurs at multiple levels – regulatory, provider, and patient – and each level requires access to appropriate information to evaluate the applicability and effectiveness of AI tools in specific contexts.

The relatively lower importance patients placed on data privacy protocols and safety management protocols suggests that, while critical from a policy and ethical standpoint,³⁰ these issues are less prominent in patients' decision-making regarding the use of the AI device. Therefore, while these factors can and should still be addressed in care delivery,⁵⁶ patient-facing communication efforts should prioritize information about HCP oversight, regulatory approval, and device performance to align with patient concerns and information needs to support shared-decision making conversations regarding AI.

Additionally, patient-facing communication and education about AI in healthcare should be tailored to meet the unique needs and preferences of different patient groups.^{57,58} For example, information about a device's added value was found to be particularly influential for patients who were less familiar with AI, had high reading health literacy, had recent routine checkups, and identified as women and people of color. Emphasizing the role of healthcare providers in overseeing AI use was found to be especially important for patients who are less familiar with AI, had recent routine checkups, and were age 55 or older. To effectively address these diverse patient needs and preferences, communication strategies could include tailoring language to patient reading levels and incorporating visual aids, infographics, and videos to help patients better understand complex information about the AI device and make informed decisions about their care. In addition, engaging patients, patient advocacy groups, and community organizations in co-designing patient-facing AI communication and educational materials including AI device labels could help address concerns from underrepresented groups who may have different levels of comfort with and trust in AI technologies and the healthcare system.⁵⁹ Providing AI communication and education materials in various formats and at different time points may improve the accessibility of information about AI and better meet the diverse needs of different patient groups.

Healthcare systems can consider offering in-person consultations during clinical visits, along with printed and/or digital materials for patients to take home, ensuring patients have the opportunity to review and understand the information at their own pace.⁶⁰ Moreover, it is the responsibility of healthcare systems to ensure providers are well-informed about AI devices and are capable of communicating their benefits and risks effectively to patients. Healthcare systems could establish standardized guidelines and best practices for patient-facing communication and education about AI in healthcare. Training providers on how to effectively discuss AI devices with diverse patient populations and address their concerns would be instrumental in ensuring that patients receive clear, transparent information that aids informed decision-making and builds merited trust in AI technologies.

Strength and Limitations

This research has several strengths. First, we utilized innovative experimental methods to elicit patient preferences for information about AI in healthcare, which could be applied to other use cases and medical specialties. Our methods can also be used to develop and evaluate other patient-facing AI communication and education materials. We applied a rigorous process for selection of information factors, including a rapid literature review and three qualitative studies, which strengthened the validity of our findings. However, our findings should be interpreted in the context of the limitations. The study sample, recruited from ResearchMatch.org, is a convenience sample and may not necessarily represent the US adult population. To address this, we intentionally oversampled racial and ethnic minority populations to ensure their representation and enable comparisons between people of color and non-Hispanic white patients in AI information preferences. Additionally, since nearly all participants had health insurance, caution is needed when generalizing the findings to uninsured populations. Third, while selection bias might be a concern, we lack data to compare participants with non-participants. Furthermore, the AI label prototypes used in this study pertained to a hypothetical device in cardiology, which may limit the generalizability of our findings, and the information elements were developed based on the study team's interpretation of the findings from our prior qualitative research, adjusted to complement the study design, and may not fully capture the full range of information that could be included on an AI label. Nonetheless, our methods of evaluating the importance of information elements can be applied to other AI devices, medical conditions, and other types of AI communication and education efforts. Lastly, the study focused solely on the informational content of AI labels. Future research is needed to explore how the format and delivery features of AI communication and education materials impact patient decision-making regarding AI technologies in healthcare.

Conclusions

Through two experimental studies, our research underscores the critical importance of transparent and accessible information about AI devices and their impact on patient trust and acceptance. Information on HCP oversight, regulatory approval, device performance, and its added value emerged as pivotal factors influencing patient

decision-making. Tailoring communication to meet the diverse needs and preferences of patient subgroups is essential for effective and equitable AI adoption in healthcare. Patient-centered communication strategies, coupled with comprehensive education for healthcare providers, would ensure that AI technologies are integrated into healthcare in a way that empowers patients to make informed choices and support overall patient care.

Conflicts of Interest

This research was supported by the Food and Drug Administration (FDA) through the Center of Excellence in Regulatory Science and Innovation (CERSI) grant to Yale University and Mayo Clinic. XZ offers scientific input to research studies through a contracted services agreement between Mayo Clinic and Exact Sciences. MM and JEM received research grants from Arnold Ventures through Yale University. JEM serves on Alexion Pharmaceuticals Ethics Committee and is the founder of Bioethics International. BAB offers scientific input to research studies through a contracted services agreement between Mayo Clinic and Anumana, Inc. AMS, SAM, DWY, and JLR report no additional competing interests.

Funding

This research was funded by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award, Center of Excellence in Regulatory Science and Innovation (CERSI) grant to Yale University and Mayo Clinic [U01FD005938 sub-award CERSI D5].

References

1. Benjamins S, Dhunoo P, Mesko B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med*. 2020;3:118. doi:10.1038/s41746-020-00324-0
2. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. Oct 2018;2(10):719-731. doi:10.1038/s41551-018-0305-z
3. Secinaro S, Calandra D, Secinaro A, Muthurangu V, Biancone P. The role of artificial intelligence in healthcare: a structured literature review. *BMC Medical Informatics and Decision Making*. 2021/04/10 2021;21(1):125. doi:10.1186/s12911-021-01488-9
4. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial Intelligence in Cardiology. *J Am Coll Cardiol*. Jun 12 2018;71(23):2668-2679. doi:10.1016/j.jacc.2018.03.521
5. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial Intelligence in Precision Cardiovascular Medicine. *J Am Coll Cardiol*. May 30 2017;69(21):2657-2664. doi:10.1016/j.jacc.2017.03.571
6. Lopez-Jimenez F, Attia Z, Arruda-Olson AM, et al. Artificial Intelligence in Cardiology: Present and Future. *Mayo Clin Proc*. 2020/05/01/ 2020;95(5):1015-1039. doi:<https://doi.org/10.1016/j.mayocp.2020.01.038>
7. Young AT, Amara D, Bhattacharya A, Wei ML. Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review. *The Lancet Digital Health*. 2021/09/01/ 2021;3(9):e599-e611. doi:[https://doi.org/10.1016/S2589-7500\(21\)00132-1](https://doi.org/10.1016/S2589-7500(21)00132-1)
8. Moy S, Irannejad M, Manning SJ, et al. Patient Perspectives on the Use of Artificial Intelligence in Health Care: A Scoping Review. *J Patient Cent Res Rev*. Spring 2024;11(1):51-62. doi:10.17294/2330-0698.2029
9. Bala S, Keniston A, Burden M. Patient Perception of Plain-Language Medical Notes Generated Using Artificial Intelligence Software: Pilot Mixed-Methods Study. *JMIR Form Res*. Jun 5 2020;4(6):e16670. doi:10.2196/16670
10. Adams SJ, Tang R, Babyn P. Patient Perspectives and Priorities Regarding Artificial Intelligence in Radiology: Opportunities for Patient-Centered Radiology. *J Am Coll Radiol*. Aug 2020;17(8):1034-1036. doi:10.1016/j.jacr.2020.01.007
11. Palmisciano P, Jamjoom AAB, Taylor D, Stoyanov D, Marcus HJ. Attitudes of Patients and Their Relatives Toward Artificial Intelligence in Neurosurgery. *World Neurosurg*. 2020/06/01/ 2020;138:e627-e633. doi:<https://doi.org/10.1016/j.wneu.2020.03.029>
12. 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 Dermatology*. 2020;156(5):501-512. doi:10.1001/jamadermatol.2019.5014
13. Tran VT, Riveros C, Ravaud P. Patients' views of wearable devices and AI in healthcare: findings from the ComPaRe e-cohort. *NPJ Digit Med*. 2019;2:53. doi:10.1038/s41746-019-0132-y
14. Lennartz S, Dratsch T, Zopfs D, et al. Use and Control of Artificial Intelligence in Patients Across the Medical Workflow: Single-Center Questionnaire Study of Patient Perspectives. *J Med Internet Res*. 2021/2/17 2021;23(2):e24221. doi:10.2196/24221

15. McCradden MD, Sarker T, Paprica PA. Conditionally positive: a qualitative study of public perceptions about using health data for artificial intelligence research. *BMJ Open*. 2020;10(10):e039798. doi:10.1136/bmjopen-2020-039798
16. Jutzi TB, Krieghoff-Henning EI, Holland-Letz T, et al. Artificial Intelligence in Skin Cancer Diagnostics: The Patients' Perspective. *Front Med (Lausanne)*. 2020;7:233. doi:10.3389/fmed.2020.00233
17. Aggarwal R, Farag S, Martin G, Ashrafiyan H, Darzi A. Patient Perceptions on Data Sharing and Applying Artificial Intelligence to Health Care Data: Cross-sectional Survey. *J Med Internet Res*. 2021/8/26 2021;23(8):e26162. doi:10.2196/26162
18. Richardson JP, Smith C, Curtis S, et al. Patient apprehensions about the use of artificial intelligence in healthcare. *npj Digital Medicine*. 2021/09/21 2021;4(1):140. doi:10.1038/s41746-021-00509-1
19. Musbahi O, Syed L, Le Feuvre P, Cobb J, Jones G. Public patient views of artificial intelligence in healthcare: A nominal group technique study. *Digit Health*. Jan-Dec 2021;7:20552076211063682. doi:10.1177/20552076211063682
20. 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*. Mar 2021;28(1)doi:10.1136/bmjhci-2020-100293
21. Brereton TA, Malik MM, Lifson M, Greenwood JD, Peterson KJ, Overgaard SM. The Role of Artificial Intelligence Model Documentation in Translational Science: Scoping Review. *Interact J Med Res*. Jul 14 2023;12:e45903. doi:10.2196/45903
22. Mitchell M, Wu S, Zaldivar A, et al. Model Cards for Model Reporting. presented at: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019; Atlanta, GA, USA. <https://doi.org/10.1145/3287560.3287596>
23. Gebru T, Morgenstern J, Vecchione B, et al. Datasheets for datasets. *Commun ACM*. 2021;64(12):86–92. doi:10.1145/3458723
24. Holland S, Hosny A, Newman S, Joseph J, Chmielinski K. The dataset nutrition label. *Data Protection and Privacy*. 2020;12(12):1.
25. Sendak MP, Gao M, Brajer N, Balu S. Presenting machine learning model information to clinical end users with model facts labels. *NPJ Digit Med*. 2020;3:41. doi:10.1038/s41746-020-0253-3
26. Robinson R, Liday C, Lee S, et al. Artificial Intelligence in Health Care—Understanding Patient Information Needs and Designing Comprehensible Transparency: Qualitative Study. Original Paper. *JMIR AI*. 2023;2:e46487. doi:10.2196/46487
27. Zhang Z, Genc Y, Wang D, Ahsen ME, Fan X. Effect of AI Explanations on Human Perceptions of Patient-Facing AI-Powered Healthcare Systems. *J Med Syst*. 2021/05/04 2021;45(6):64. doi:10.1007/s10916-021-01743-6
28. Kostick-Quenet K, Lang BH, Smith J, Hurley M, Blumenthal-Barby J. Trust criteria for artificial intelligence in health: normative and epistemic considerations. *J Med Ethics*. 2024;50(8):544-551. doi:10.1136/jme-2023-109338
29. U.S. Food & Drug Administration. Device Labeling. Updated 10/23/2020. Accessed December 20th, 2024. <https://www.fda.gov/medical-devices/overview-device-regulation/device-labeling>
30. Mooghali M, Stroud AM, Yoo DW, et al. Trustworthy and ethical AI-enabled cardiovascular care: a rapid review. *BMC Medical Informatics and Decision Making*.

2024/09/04 2024;24(1):247. doi:10.1186/s12911-024-02653-6

31. Stroud A, Minter S, Zhu X, Ridgeway J, Miller J, Barry B. Patient information needs for transparent and trustworthy cardiovascular artificial intelligence: A qualitative study. *PLOS Digit Health*. 2025;4(4):e0000826. doi:<https://doi.org/10.1371/journal.pdig.0000826>
32. Merlo G, van Driel M, Hall L. Systematic review and validity assessment of methods used in discrete choice experiments of primary healthcare professionals. *Health Economics Review*. 2020/12/09 2020;10(1):39. doi:10.1186/s13561-020-00295-8
33. Quaife M, Terris-Prestholt F, Di Tanna GL, Vickerman P. How well do discrete choice experiments predict health choices? A systematic review and meta-analysis of external validity. *The European Journal of Health Economics*. 2018/11/01 2018;19(8):1053-1066. doi:10.1007/s10198-018-0954-6
34. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete Choice Experiments in Health Economics: Past, Present and Future. *Pharmacoeconomics*. 2019/02/01 2019;37(2):201-226. doi:10.1007/s40273-018-0734-2
35. de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. <https://doi.org/10.1002/hec.1697>. *Health Econ*. 2012/02/01 2012;21(2):145-172. doi:<https://doi.org/10.1002/hec.1697>
36. Vanniyasingam T, Cunningham CE, Foster G, Thabane L. Simulation study to determine the impact of different design features on design efficiency in discrete choice experiments. *BMJ Open*. 2016;6(7):e011985. doi:10.1136/bmjopen-2016-011985
37. Traets F, Sanchez DG, Vandebroek M. Generating Optimal Designs for Discrete Choice Experiments in R: The idfix Package. *Journal of Statistical Software*. 11/29 2020;96(3):1 - 41. doi:10.18637/jss.v096.i03
38. Carlsson F, Martinsson P. Design techniques for stated preference methods in health economics. <https://doi.org/10.1002/hec.729>. *Health Econ*. 2003/04/01 2003;12(4):281-294. doi:<https://doi.org/10.1002/hec.729>
39. Burgess L, Street DJ. Optimal designs for choice experiments with asymmetric attributes. *Journal of Statistical Planning and Inference*. 2005/09/01/ 2005;134(1):288-301. doi:<https://doi.org/10.1016/j.jspi.2004.03.021>
40. Szinay D, Cameron R, Naughton F, Whitty JA, Brown J, Jones A. Understanding Uptake of Digital Health Products: Methodology Tutorial for a Discrete Choice Experiment Using the Bayesian Efficient Design. *J Med Internet Res*. 2021/10/11 2021;23(10):e32365. doi:10.2196/32365
41. Grömping U. R Package FrF2 for Creating and Analyzing Fractional Factorial 2-Level Designs. *Journal of Statistical Software*. 01/25 2014;56(1):1 - 56. doi:10.18637/jss.v056.i01
42. Cacioppo JT, Petty RE. The need for cognition. *Journal of Personality and Social Psychology*. 1982;42:116-131. doi:10.1037/0022-3514.42.1.116
43. Lins de Holanda Coelho G, H. P. Hanel P, J. Wolf L. The Very Efficient Assessment of Need for Cognition: Developing a Six-Item Version. *Assessment*. 2020/12/01 2018;27(8):1870-1885. doi:10.1177/1073191118793208
44. Thompson HS, Valdimarsdottir HB, Winkel G, Jandorf L, Redd W. The Group-Based Medical Mistrust Scale: psychometric properties and association with breast cancer screening. *Prev Med*. 2004/02/01/ 2004;38(2):209-218. doi:<https://doi.org/10.1016/j.ypmed.2003.09.041>

45. Morris NS, MacLean CD, Chew LD, Littenberg B. The Single Item Literacy Screener: Evaluation of a brief instrument to identify limited reading ability. *BMC Fam Pract.* 2006/03/24 2006;7(1):21. doi:10.1186/1471-2296-7-21
46. McNaughton CD, Cavanaugh KL, Kripalani S, Rothman RL, Wallston KA. Validation of a Short, 3-Item Version of the Subjective Numeracy Scale. *Med Decis Making.* Nov 2015;35(8):932-6. doi:10.1177/0272989x15581800
47. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Med Decis Making.* Sep-Oct 2007;27(5):672-80. doi:10.1177/0272989x07304449
48. Gierisch JM, Earp JA, Brewer NT, Rimer BK. Longitudinal Predictors of Nonadherence to Maintenance of Mammography. *Cancer Epidemiol Biomarkers Prev.* 2010;19(4):1103-1111. doi:10.1158/1055-9965.EPI-09-1120
49. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. *The Patient - Patient-Centered Outcomes Research.* 2015/10/01 2015;8(5):373-384. doi:10.1007/s40271-015-0118-z
50. Harris PA, Scott KW, Lebo L, Hassan N, Lightner C, Pulley J. ResearchMatch: A National Registry to Recruit Volunteers for Clinical Research. *Acad Med.* 2012;87(1):66-73. doi:10.1097/ACM.0b013e31823ab7d2
51. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing; 2023. <https://www.R-project.org/>
52. Wickham H, Averick M BJ, Chang W, et al. Welcome to the tidyverse. *Journal of Open Source Software.* 2019;4(43):1686. doi:10.21105/joss.01686
53. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software.* 10/07 2015;67(1):1 - 48. doi:10.18637/jss.v067.i01
54. *ordinal—Regression Models for Ordinal Data.* Version R package version 2023.12-4. 2023. <https://CRAN.R-project.org/package=ordinal>
55. *marginaleffects: Predictions, Comparisons, Slopes, Marginal Means, and Hypothesis Tests.* Version R package version 0.18.0.9004. 2024. <https://marginaleffects.com/>
56. Williamson SM, Prybutok V. Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare. *Applied Sciences.* 2024;14(2):675.
57. Young A, Tordoff J, Smith A. 'What do patients want?' Tailoring medicines information to meet patients' needs. *Research in Social and Administrative Pharmacy.* 2017/11/01/ 2017;13(6):1186-1190. doi:<https://doi.org/10.1016/j.sapharm.2016.10.006>
58. Smets EMA, Menichetti J, Lie HC, Gerwing J. What do we mean by "tailoring" of medical information during clinical interactions? *Patient Educ Couns.* 2024/02/01/ 2024;119:108092. doi:<https://doi.org/10.1016/j.pec.2023.108092>
59. Adus S, Macklin J, Pinto A. Exploring patient perspectives on how they can and should be engaged in the development of artificial intelligence (AI) applications in health care. *BMC Health Serv Res.* 2023/10/26 2023;23(1):1163. doi:10.1186/s12913-023-10098-2
60. Stroud AM, Minter SA, Zhu X, Ridgeway JL, Miller JE, Barry BA. Patient information needs for transparent and trustworthy artificial intelligence in healthcare.

medRxiv. 2024:2024.07.02.24309850. doi:10.1101/2024.07.02.24309850

Preprint
JMIR Publications

Appendix 1. Choice sets

	Alternative A								Alternative B							
	X1 Data privacy	X2 Perfor mance	X3 Adde d value	X4 Regulato ry approval	X5 Expert endors ement	X6 Validatio n	X7 Device safety	X8 HCP over sight	X1 Data privacy	X2 Perfor mance	X3 Adde d valu e	X4 Regulato ry approval	X5 Expert endors ement	X6 Validati on	X7 Device safety	X8 HCP over sight
1	Opt-out	High	Off	Off	On	Internal	Reactiv e	On	Opt-out	High	Off	Off	On	External	Proacti ve	On
2	Opt-out	High	On	On	On	External	Reactiv e	On	Opt-out	High	Off	On	On	Internal	Proacti ve	Off
3	Opt-in	Low	On	Off	On	Internal	Reactiv e	On	Opt-out	High	On	On	Off	Internal	Reactiv e	Off
4	Opt-in	Low	Off	Off	Off	Internal	Proacti ve	On	Opt-out	Low	Off	On	On	Internal	Reactiv e	On
5	Opt-in	Low	Off	On	On	Internal	Reactiv e	Off	Opt-in	Low	Off	Off	Off	External	Reactiv e	Off
6	Opt-in	High	Off	On	On	Internal	Reactiv e	Off	Opt-out	Low	On	Off	Off	Internal	Proacti ve	On
7	Opt-in	High	Off	Off	Off	External	Reactiv e	On	Opt-in	Low	On	On	On	External	Reactiv e	Off
8	Opt-out	Low	On	On	On	External	Reactiv e	Off	Opt-out	High	Off	Off	On	External	Proacti ve	On
9	Opt-out	Low	Off	Off	Off	External	Proacti ve	On	Opt-out	High	Off	On	On	Internal	Proacti ve	Off
10	Opt-out	High	Off	Off	On	Internal	Reactiv e	On	Opt-out	High	On	On	Off	Internal	Reactiv e	Off
11	Opt-out	High	On	On	On	External	Reactiv e	On	Opt-out	Low	Off	On	On	Internal	Reactiv e	On
12	Opt-in	Low	On	Off	On	Internal	Reactiv e	On	Opt-in	Low	Off	Off	Off	External	Reactiv e	Off
13	Opt-in	Low	Off	Off	Off	Internal	Proacti ve	On	Opt-out	Low	On	Off	Off	Internal	Proacti ve	On
14	Opt-in	Low	Off	On	On	Internal	Reactiv e	Off	Opt-in	Low	On	On	On	External	Reactiv e	Off
15	Opt-in	High	Off	On	On	Internal	Reactiv e	Off	Opt-out	High	Off	Off	On	External	Proacti ve	On
16	Opt-in	High	Off	Off	Off	External	Reactiv e	On	Opt-out	High	Off	On	On	Internal	Proacti ve	Off

Appendix 2. 2IV⁸⁻³ fractional factorial experimental design

Condition	X1. Data privacy and security	X2. Performance	X3. Added value	X4. Regulatory approval	X5. Expert endorsement	X6. Validation	X7. Device safety	X8. HCP oversight
1	Opt-in	High	On	Off	Off	External	Reactive	On
2	Opt-in	High	On	Off	On	External	Proactive	Off
3	Opt-in	Low	On	Off	On	External	Proactive	On
4	Opt-out	Low	On	On	Off	External	Proactive	Off
5	Opt-in	Low	On	Off	Off	External	Reactive	Off
6	Opt-out	Low	Off	Off	On	External	Proactive	Off
7	Opt-out	Low	Off	Off	Off	External	Reactive	On
8	Opt-out	High	On	Off	Off	Internal	Proactive	Off
9	Opt-out	High	Off	On	On	Internal	Proactive	Off
10	Opt-out	Low	Off	On	Off	Internal	Reactive	Off
11	Opt-in	High	Off	On	On	External	Reactive	On
12	Opt-out	Low	Off	On	On	Internal	Proactive	On
13	Opt-in	High	Off	Off	On	Internal	Reactive	Off
14	Opt-in	Low	On	On	Off	Internal	Reactive	On
15	Opt-out	High	Off	On	Off	Internal	Reactive	On
16	Opt-in	Low	Off	Off	On	Internal	Reactive	On
17	Opt-out	Low	On	On	On	External	Reactive	On
18	Opt-in	Low	Off	Off	Off	Internal	Proactive	Off
19	Opt-out	Low	On	Off	On	Internal	Reactive	Off
20	Opt-in	High	On	On	On	Internal	Proactive	On
21	Opt-in	Low	On	On	On	Internal	Proactive	Off
22	Opt-out	High	On	On	On	External	Reactive	Off
23	Opt-out	High	On	On	Off	External	Proactive	On
24	Opt-in	Low	Off	On	Off	External	Proactive	On
25	Opt-out	High	Off	Off	On	External	Proactive	On
26	Opt-in	High	On	On	Off	Internal	Reactive	Off
27	Opt-in	High	Off	On	Off	External	Proactive	Off
28	Opt-in	Low	Off	On	On	External	Reactive	Off
29	Opt-out	High	Off	Off	Off	External	Reactive	Off
30	Opt-out	High	On	Off	On	Internal	Reactive	On
31	Opt-in	High	Off	Off	Off	Internal	Proactive	On
32	Opt-out	Low	On	Off	Off	Internal	Proactive	On

Appendix 3. Subgroup differences in effects of information factors on the probability of the AI device being trusted

Information Factors	Subgroups	AME ^a (95% CI) Percentage Points
Added Value (Info present – info absent)		
	Need for cognition ^b	
	High (Mean + 1SD)	16.21 (13.45, 18.97)
	Moderate (Mean)	14.27 (12.31, 16.22)
	Low (Mean – 1SD)	12.23 (9.47, 14.98)
HCP Oversight (Info present – info absent)		
	Familiarity with AI ^c	
	Somewhat/Slightly/Not at all	19.27 (15.23, 23.32)
	Extremely/Very	11.30 (6.64, 15.95)
	Last routine medical checkup ^d	
	Within last year	18.27 (14.49, 22.04)
	1 or more years ago	11.63 (6.25, 17.02)
	Age group	
	18-34	15.25 (10.43, 20.08)
	35-54	10.34 (5.15, 15.53)
	55 or older	23.44 (17.78, 29.10)
Performance (High – Low)		
	Reading health literacy ^b	
	High	22.02 (18.82, 25.22)
	Low	11.06 (7.85, 14.27)
	Last routine medical checkup ^d	
	Within the last year	20.09 (17.23, 22.95)
	One or more years ago	10.38 (6.32, 14.43)
	Gender	
	Woman	19.38 (16.10, 22.66)
	Man	12.85 (9.62, 16.08)
Regulatory Approval (Info present – info absent)		

	Reading health literacy ^b	
	High	25.07 (20.25, 29.89)
	Low	13.63 (8.72, 18.53)
	Last routine medical checkup ^d	
	Within the last year	23.29 (18.98, 27.61)
	One or more years ago	12.92 (6.72, 19.12)

a AME: average marginal effect. It is the average change in the predicted probabilities (percentage point increase or decrease) of the AI device being trusted/accepted across all participants when moving from one level of the information factor to the other, keeping all other variables in the model constant.

b Model adjusted for education level.

c Model adjusted for health literacy, numeracy, education level, and recency of last medical checkup.

d Model adjusted for health insurance coverage, age group, and perceived household financial status.

Appendix 4. Subgroup differences in effects of information factors on the probability of the AI device being accepted

Information Factors	Subgroups	AME ^a (95% CI) Percentage Points
Added Value (Info present – Info absent)		
	Familiarity with AI ^b	
	Somewhat/Slightly/Not at all	18.06 (15.11, 21.01)
	Extremely/Very	11.49 (7.95, 15.03)
	Reading health literacy ^c	
	High	19.34 (16.03, 22.65)
	Low	11.96 (8.64, 15.28)
	Last routine medical checkup ^d	
	Within the last year	19.19 (16.40, 21.97)
	One or more years ago	8.81 (4.76, 12.87)
	Gender	
	Woman	18.94 (15.54, 22.34)
	Man	11.38 (8.05, 14.71)
	Race/Ethnicity	
	Person of Color	19.64 (16.18, 23.09)
	Non-Hispanic White	12.01 (8.82, 15.19)
HCP Oversight (Info present – Info absent)		
	Familiarity with AI ^b	
	Somewhat/Slightly/Not at all	25.34 (21.88, 28.8)
	Extremely/Very	9.59 (5.29, 13.89)
	Last routine medical checkup ^d	
	Within last year	21.92 (18.69, 25.15)
	1 or more years ago	13.31 (8.33, 18.29)
	Gender	
	Woman	22.40 (18.07, 26.73)
	Man	14.01 (9.87, 18.16)
	Age group	
	18-34	13.77 (8.89, 18.65)
	35-54	13.05 (8.18, 17.92)

		17.93)
	55 or older	30.13 (24.46, 35.79)
Performance (High – Low)		
	Familiarity with AI ^b	
	Somewhat/Slightly/Not at all	18.18 (15.26, 21.09)
	Extremely/Very	10.58 (7.28, 13.88)
	Reading health literacy ^c	
	High	21.41 (18.22, 24.59)
	Low	7.66 (4.47, 10.86)
	Numeracy ^e	
	High (Mean + 1SD)	19.38 (16.22, 22.55)
	Moderate (Mean)	14.59 (12.34, 16.85)
	Low (Mean – 1SD)	9.77 (6.62, 12.93)
	Last routine medical checkup ^d	
	Within the last year	18.91 (16.29, 21.53)
	One or more years ago	7.03 (3.04, 11.02)
	Gender	
	Woman	18.71 (15.43, 21.98)
	Man	9.57 (6.37, 12.77)
	Age group	
	18-34	11.82 (8.14, 15.50)
	35-54	12.80 (9.01, 16.59)
	55 or older	20.88 (16.53, 25.22)
Regulatory Approval (Info present – info absent)		
	Reading health literacy ^c	
	High	19.93 (15.25, 24.62)
	Low	5.82 (0.90, 10.75)
	Numeracy ^c	
	High (Mean + 1SD)	15.79 (11.25, 20.34)
	Moderate (Mean)	12.93 (9.51, 16.35)
	Low (Mean – 1SD)	9.16 (4.18, 14.13)
	Last routine medical checkup ^d	
	Within the last year	18.65 (14.30,

		23.01)
	One or more years ago	3.96 (-2.14, 10.06)
	Gender	
	Woman	16.58 (11.60, 21.57)
	Man	9.43 (4.58, 14.28)

a AME: average marginal effect. It is the average change in the predicted probabilities (percentage point increase or decrease) of the AI device being trusted/accepted across all participants when moving from one level of the information factor to the other, keeping all other variables in the model constant. All analyses adjusted for other participant characteristics.

b Model adjusted for health literacy, numeracy, education level, and recency of last medical checkup.

c Model adjusted for education level.

d Model adjusted for health insurance coverage, age group, and perceived household financial status.