Social Science & Medicine 349 (2024) 116871

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Comparing preferences for skin cancer screening: AI-enabled app vs dermatologist

Susanne Gaube [a,*], Isabell Biebl [b], Magdalena Karin Maria Engelmann [c], Anne-Kathrin Kleine [b], Eva Lermer [b,d]

[a] UCL Global Business School for Health, University College London, UCL East – Marshgate, 7 Sidings St, London, E20 2AE, United Kingdom
[b] Center for Leadership and People Management, Department of Psychology, LMU Munich, Geschwister-Scholl-Platz 1, 80539, Munich, Germany
[c] Department of Experimental Psychology, University of Regensburg, 93040, Regensburg, Germany
[d] Department of Business Psychology, Technical University of Applied Sciences Augsburg, An der Hochschule 1, 86161, Augsburg, Germany

ARTICLE INFO

Handling Editor: Susan J. Elliott

ABSTRACT

Background and aim: Skin cancer is a major public health issue. While self-examinations and professional screenings are recommended, they are rarely performed. Mobile health (mHealth) apps utilising artificial intelligence (AI) for skin cancer screening offer a potential solution to aid self-examinations; however, their uptake is low. Therefore, the aim of this research was to examine provider and user characteristics influencing people's decisions to seek skin cancer screening performed by a mHealth app or a dermatologist.

Methods: Two forced-choice conjoint experiments with $N_{main} = 1591$ and $N_{replication} = 308$ participants from the United States were conducted online to investigate preferences for screening providers. In addition to the provider type (mHealth app vs dermatologist), the following provider attributes were manipulated: costs, expertise, privacy policy, and result details. Subsequently, a questionnaire assessed various user characteristics, including demographics, attitudes toward AI technology and medical mistrust.

Results: Outcomes were consistent across the two studies. The provider type was the most influential factor, with the dermatologist being selected more often than the mHealth app. Cost, expertise, and privacy policy also significantly impacted decisions. Demographic subgroup analyses showed rather consistent preference trends across various age, gender, and ethnicity groups. Individuals with greater medical mistrust were more inclined to choose the mHealth app. Trust, accuracy, and quality ratings were higher for the dermatologist, whether selected or not.

Conclusion: Our results offer valuable insights for technology developers, healthcare providers, and policymakers, contributing to unlocking the potential of skin cancer screening apps in bridging healthcare gaps in underserved communities.

1. Introduction

Skin cancer is the most common type of cancer in the United States (US), with one in five US Americans being affected (American Cancer Society, n.d.). Melanoma of the skin causes the majority of skin cancer deaths (Siegel et al., 2023). However, when detected early and appropriately treated, skin cancer is well-treatable, and survival rates are high (American Academy of Dermatology Association, n.d.; Siegel et al., 2023). Consequently, the American Academy of Dermatology Association (n.d.) recommends regular skin self-examinations to check for signs of skin cancer. People with a higher risk of skin cancer and individuals who found signs of skin cancer during their self-examination should be screened by a dermatologist. However, the rates of self-examinations and professional skin cancer screenings are low (Kasparian et al., 2009). While self-examination can help to detect skin cancer early (Berwick et al., 1996), laypeople's accuracy is insufficient (Hamidi et al., 2010). Even general practitioners (GPs), often the first healthcare specialist consulted by concerned patients, show sub-optimal accuracy in detecting skin cancer (Sangers et al., 2022). Besides that, many regions are underserved with dermatologists, preventing individuals from getting professional skin cancer screenings (e.g., Feng et al., 2018).

Technology advocates argue that mobile health (mHealth)

* Corresponding author. UCL East – Marshgate, 7 Sidings St, London, E20 2AE, United Kingdom.
E-mail address: susanne.gaube@ucl.ac.uk (S. Gaube).

https://doi.org/10.1016/j.socscimed.2024.116871
Received 20 October 2023; Received in revised form 26 March 2024; Accepted 4 April 2024
Available online 15 April 2024

applications can improve the availability of affordable skin cancer screening. Apps analyse smartphone images of skin lesions using AI algorithms (Kong et al., 2021; Udrea et al., 2020). In ideal conditions, these algorithms have shown high sensitivity and reasonable specificity (Udrea et al., 2020). However, in natural settings, their performance is still below specialists'; therefore, using mHealth apps is not a substitute for a professional skin cancer screening (Freeman et al., 2020; Jahn et al., 2022). Nevertheless, since laypeople and GPs show even worse accuracy, mHealth apps might be valuable tools for supporting self-examination and indicating whether to follow up with a dermatologist. However, despite their potential, the uptake of mHealth apps for skin cancer screening is still low (Sangers et al., 2021), and most people prefer an examination by a dermatologist over an mHealth app (Baldauf et al., 2020; Jahn et al., 2022; Longoni et al., 2019).

1.1. The algorithm aversion framework

Algorithm aversion is defined as a preference for human decision-making over algorithms, a higher reliance on human support, and an inclination to evaluate human actions more favourably (Dietvorst et al., 2015; Jussupow et al., 2020). The algorithm aversion model suggests that aversion is influenced by the characteristics of an algorithm that gives advice or performs a task as well as the characteristics of a human that could alternatively do the same (Jussupow et al., 2020). Understanding users' reluctance to use mHealth apps for skin cancer screening as a function of the app's features and the attributes of a dermatologist as an alternative service provider is a better approach than merely focusing on the technology. It reflects the real-world decision-making process, where individuals exploring mHealth apps for skin cancer screening typically also have the alternative option to consult a dermatologist. However, with the exception of the conjoint experiment from Longoni et al. (2019), there is a lack of studies that compared technology and human service provider characteristics simultaneously to understand how these will affect screening provider decisions. Moreover, the algorithm aversion model was developed based on a literature review and is therefore limited to factors that have already been extensively researched. Established characteristics determining algorithm aversion include the algorithm's performance and perceived capabilities, algorithm agency (i.e., performative vs advisory algorithms), and human involvement (Jussupow et al., 2020). Besides these, expertise and social distance were identified as further key human characteristics influencing algorithm aversion. However, other potentially relevant characteristics that have received less consideration in previous works such as costs, data privacy and security, and explainability, should be evaluated. Finally, while considering algorithm and human service provider characteristics is the foundation for understanding algorithm aversion, it has been argued that the algorithm aversion model should be extended to also include user characteristics such as their attitudes towards the providers (Jussupow et al., 2020). Consequently, the present study aims to provide empirical evidence for the impact of both established and novel technology and alternative human service provider characteristics, as well as user characteristics on algorithm aversion in the context of AI-enabled mHealth apps for skin cancer screening.

1.2. Previous research on algorithm aversion for skin cancer screening mHealth apps

Certain aspects of the above-described algorithm and human service provider characteristics have already been investigated concerning mHealth apps for skin cancer screening. In one study focusing on performance and perceived capabilities, patients indicated less confidence and trust in an examination conducted by mHealth apps compared to dermatologists (Jahn et al., 2022). In the same study, less than 25% of all surveyed patients expected reliable results from these apps, and less than 35% reported that using screening apps would reduce their fear of developing skin cancer. In another study, participants were more resistant to AI-enabled skin cancer screenings even when the AI's performance was explicitly stated to be superior to a human provider (Longoni et al., 2019). Considering algorithm agency and human involvement, individuals are more averse to mHealth apps when they are perceived as replacing physicians and making independent decisions instead of assisting physicians' decision-making (Jahn et al., 2022; Jutzi et al., 2020).

Moreover, some user characteristics that might affect aversion to mHealth apps for skin cancer screening have also been examined. One study found that individuals with a stronger sense of uniqueness expressed greater resistance to automated skin exams (Longoni et al., 2019). This is attributed to concerns that AI might be less adept than human physicians in considering people's distinct characteristics and situations. Another study exploring medical mistrust, i.e., a tendency to distrust healthcare providers, institutions, or the system (Thompson et al., 2004), reported that people with low medical mistrust perceive physicians as more trustworthy and fairer (M. K. Lee and Rich, 2021). Conversely, individuals with high medical mistrust evaluated both human and AI providers for skin cancer screening as similarly trustworthy and fair. Furthermore, it has been shown that older patients expressed greater trust in smartphone apps for skin cancer screening than younger patients, while sex did not influence their trustworthiness rating (Jahn et al., 2022).

1.3. The present study

Research examining the effect of technology, alternative human provider, and user characteristics simultaneously when deciding on using a skin cancer screening app is scarce. However, it is crucial to identify the conditions affecting aversion to or appreciation of mHealth apps and individuals who would be willing to use these apps to maximise their potential benefits as aids in self-examining skin health. The current research employs a forced-choice conjoint paradigm in two studies (main and replication) to investigate peoples' preferences for an AI-enabled mHealth app against a dermatologist as the alternative screening provider. Participants' provider preferences and their evaluations are examined by manipulating both providers' characteristics while accounting for user characteristics to assess how these factors might interact. The study's proposed research model and the hypotheses are depicted in Fig. 1.

In accordance with the mHealth acceptance and algorithm aversion literature, we hypothesised:

H1. In general, participants exhibit a preference for the dermatologist over the mHealth app.

Moreover, based on the review of the literature, we decided to manipulate the following provider characteristics to test their influence on participants' provider decisions and determine their relative importance: (1) cost, (2) level of expertise, (3) data privacy policy, (4) and level of result details.

The cost of a product or service is a fundamental criterion for consumer decision (e.g., Erickson and Johansson, 1985) but is understudied in the context of mHealth acceptance for skin cancer screening and algorithm aversion in general. Longoni et al. (2019) found that cheaper skin cancer screening providers were preferred, but cost was less important than the type of provider (robotic vs. human). However, there cost was varied between $20, $40, and $60, but some mHealth apps provide services free of charge or at a lower cost (Google Health, n.d.; SkinVision, n.d.). At the same time the mean price for a professional skin cancer screening is $150 (Matsumoto et al., 2018). Consequently, we let provider cost vary more widely and hypothesised:

H2a. Participants prefer the provider with the lower cost.

H2b. The provider's cost is the most important attribute influencing participants' provider choice.

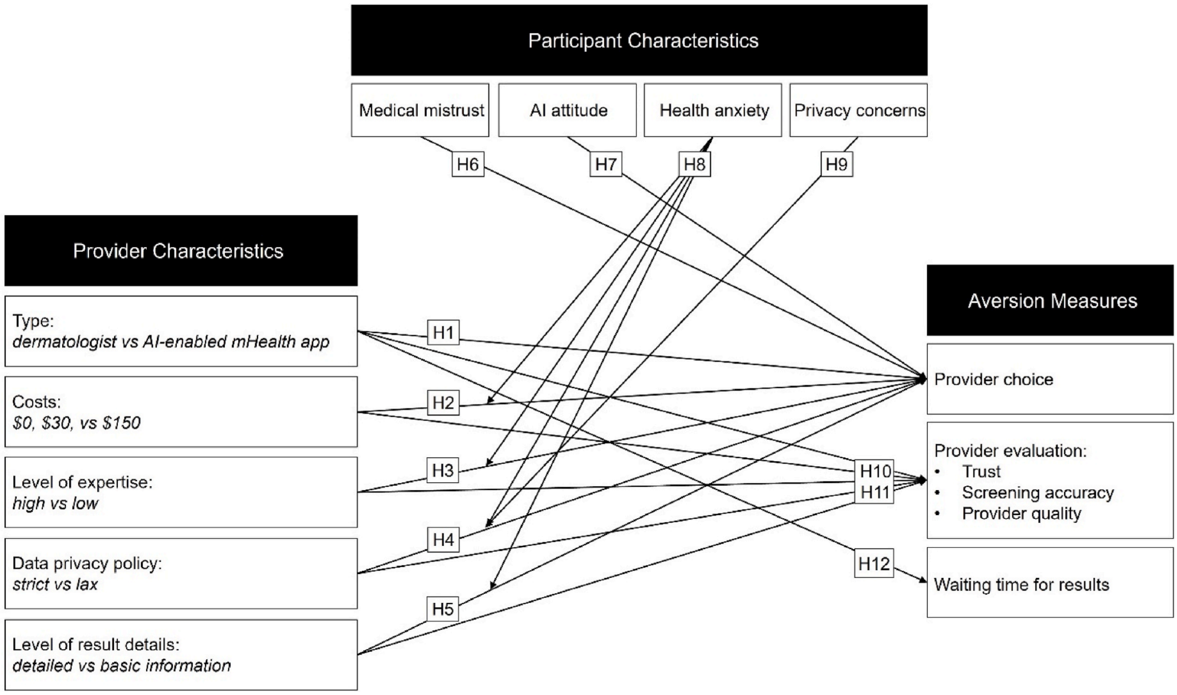

Fig. 1. Research model and proposed hypotheses.

Algorithm aversion depends on both the algorithm's performance and the human service provider's expertise. It has been shown that diagnostic accuracy influences skin cancer provider decisions (Haggenmüller et al., 2021; Longoni et al., 2019). In reality, accuracy is not an available metric when selecting a dermatologist. However, individuals may evaluate expertise based on experience, which can also be used as a proxy for the app's performance and capability. We manipulated providers' expertise and hypothesised:

H3a. Participants prefer the provider with a higher level of expertise.

H3b. The provider's expertise is the second most important attribute influencing provider decisions.

Given the sensitivity of health data, ensuring privacy and data security is crucial. However, the risk of inadequate data security is a significant concern with AI technology in healthcare (Rajpurkar et al., 2022). It has been shown that data safety concerns play a role in rejecting skin cancer diagnosis tools (Haggenmüller et al., 2021). However, this has not been investigated in comparison to a human screening provider, so it is unclear how important this attribute might be. To test this, we manipulated the strictness of the providers' privacy policy and hypothesised:

H4a. Participants prefer the provider with a strict privacy policy.

H4b. The strictness of the provider's privacy policy is less important for the provider decision compared to cost and expertise.

Finally, a limitation of AI technology is its inability to explain the rationale behind predictions or decisions, i.e., the 'black box problem' (Hummelsberger et al., 2023; Rajpurkar et al., 2022). To address this, systems often provide post-hoc explanations to enhance user trust. Many potential users of mHealth apps for skin cancer screening expect some explanation about the algorithm's decision-making process (Haggenmüller et al., 2021). However, it is unclear whether participants demand the same from a human screening provider. Consequently, we manipulated the level of detail given with the screening results and hypothesised:

H5a. Participants prefer the provider that provides more detailed screening results.

H5b. The level of detail concerning the screening result is less important for the provider decision compared to cost and expertise.

To expand the algorithm aversion model, our second aim was to explore how user characteristics influence provider choice. Due to limited prior research, we avoided formulating specific hypotheses regarding participant demographics. Dietvorst et al. (2015) and Jussupow et al. (2020) suggest that participants' attitudes towards support providers may impact algorithm aversion. Therefore, we examined the effects of medical mistrust as a measure of attitudes toward human healthcare providers and general attitudes toward AI technology on participants' provider decisions. Furthermore, health-related anxiety levels can impact health decisions (Salkovskis et al., 2002) and might play a role in their provider choices. Given the sensitivity of health data, peoples' general concerns regarding data security and privacy may be relevant when choosing a screening provider. We hypothesised:

H6. Participants exhibiting higher levels of medical mistrust tend to prefer the AI-enabled mHealth app.

H7. Participants sceptical of AI tend to prefer the dermatologist.

H8. Participants with higher levels of health anxiety prefer a provider with high attribute levels.

H9. Participants with higher levels of data security and privacy concerns prefer a provider with a strict privacy policy.

Algorithm aversion does not only manifest in a preference for human decision-making over algorithms, but also in a tendency to evaluate human actions more favourably. Research has shown that patients perceive an examination conducted via an mHealth app as less trustworthy and accurate compared to an examination conducted by a dermatologist (Jahn et al., 2022). However, the study did not systematically examine the influence of provider attributes on patients' preference for an mHealth app versus an examination by a dermatologist. To the best of our knowledge, no study has yet examined how provider type and attributes relate to the provider evaluation. Consequently, the final aim of our study was to assess how participants evaluated trust in and the quality, including screening accuracy and provider assessment, of the chosen and non-chosen provider. Furthermore, we examined participants' willingness to wait for the screening results and the likelihood of seeking a second opinion as novel additional proxy measures for their provider preference and evaluation. We hypothesised:

H10. Participants rate the trustworthiness, screening result accuracy, and overall provider quality higher for the dermatologists than the mHealth app.

H11. The provider attribute values are associated with the ratings of provider trust, screening accuracy, and quality.

H12. Participants are willing to wait longer for the screening results from a dermatologist than from an AI-enabled mHealth app.

2. Methods

2.1. Participants

In both studies, participants from the US population were recruited via the online research platform Prolific (Prolific, London, UK). We aimed to collect a representative sample by employing quotas for age, gender, ethnicity, and health insurance aligned with US 2020 census data (U.S. Department of Commerce, n.d.). The sampling strategy did not target specific geographical regions. Participants who completed the questionnaire received monetary compensation. For the main study, a minimum sample size of $N_{\min} = 1442$ was estimated for a power (1 - β-error probability) of 0.80 and an estimated effect size of AMCE = 0.09 using the cjpowR package (Schuessler and Freitag, 2020). In the main study, 1702 people started the survey, from which 3 had to be excluded for not giving consent, 39 for not finishing the entire survey, and 69 for failing at least one attention check item. This left us with a sample of $N_{main} = 1591$ for the analysis. Participants who completed the main study were ineligible for inclusion in the replication study. In the replication study, we aimed to gather approximately 20% of the main study's sample. The sample size was planned pragmatically so that we should be able to detect the most important effects from the conjoint experiment in the main study with a power of 0.80 and an estimated effect size of AMCE = 0.20 ($N_{\min} = 283$). 317 people started the replication survey, from which 2 had to be excluded who didn't complete the survey and 7 for failing at least one attention check item. This left us with a sample of $N_{replication} = 308$ for the replication analysis. Participants' main demographics are presented in Table 1. In both samples, more than half of the participants had a university degree (main: 57.5%, replication: 55.8%). Only a minority reported to not have any health insurance coverage (main: 10.0%, replication: 7.5%). Most participants reported no previous history of skin cancer (main: 93.4%, replication: 93.5%); however, a substantial proportion also never had a skin cancer examination (main: 69.5%, replication: 75.6%). Additional sample information can be found in the online supplements.

Table 1
Participant demographics.

	Main Study (N = 1591)	Replication Study (N = 308)
Age (years)		
Mean (SD)	44.3 (16.3)	41.0 (14.9)
Median [Min, Max]	42.0 [18.0, 93.0]	38.0 [18.0, 75.0]
Gender		
Female	783 (49.2%)	153 (49.7%)
Male	767 (48.2%)	149 (48.4%)
Non-binary/third gender	25 (1.6%)	4 (1.3%)
Transgender	6 (0.4%)	2 (0.6%)
Prefer not to answer	10 (0.6%)	0 (0%)
Ethnicity		
White	1216 (76.4%)	215 (69.8%)
Black or African American	138 (8.7%)	39 (12.7%)
Asian	131 (8.2%)	32 (10.4%)
American Indian or Alaska Native	9 (0.6%)	3 (1.0%)
Other	43 (2.7%)	5 (1.6%)
Multiple ethnicities selected	54 (3.4%)	14 (4.5%)

2.2. Research design

Data for the pre-registered studies (main: https://doi.org/10.17605/OSF.IO/P8Q3X; replication https://doi.org/10.17605/OSF.IO/87GVH) were collected between April and May 2022 (main) and between December 2022 and March 2023 (replication) via Qualtrics (Qualtrics, Provo, UT, USA), where the experiment was programmed on. The research plan was reviewed and approved by the local Research Ethics Committee (approval number: 21-2692-101). It encompassed a forced-choice conjoint experiment, succeeded by a survey. Forced-choice conjoint experiments are a widely used and reliable method to study preferences (Green et al., 2001; Hainmueller et al., 2015). In both studies, participants received information on the prevalence of skin cancer and screening recommendations (adopted from Lee and Rich (2021) and Longoni et al. (2019). Subsequently, they were prompted to envision a scenario where they had opted for a skin cancer screening. Within this context, participants were required to choose between two provider options: an examination with an AI-enabled mHealth app called *Skin-AI-D* or by a *dermatologist* at their office. For both providers, participants received information regarding the procedure, which was kept as comparable as possible and at the same length. The exact wording of the introduction and the scenario can be found in the project's online repository. Subsequently, respondents were presented with one vignette consisting of a table displaying the two provider options, along with varying values for four additional attributes associated with each provider. The four attributes were the (a) cost of the screening (independent of health insurance), (b) expertise of the provider, (c) the provider's privacy policy, and (d) the level of detail given with the screening results. Table 2 outlines the attributes and their possible values corresponding to each provider. An example of a possible vignette can be found in the online supplements. In the main study, providers and their respective attributes were presented in the same order for every participant. The primary objective of the replication study was to ensure that the presentation order did not impact the relative importance of the attributes. Therefore, in the replication study, the provider and attribute order was randomised to mitigate potential order effects (Hainmueller et al., 2014). Additional details regarding the operationalisation of attributes are available in the online supplements. By combining attribute values, a total of 48 distinct vignettes could be generated, and one was randomly allocated to each participant. Following the presentation of the vignette, participants were prompted to select a provider for their screening. The conjoint experiments were followed by questions regarding their choice (only main study), psychometrics scales (only main study), and demographic questions (age, gender, ethnicity, level of education, household income, prior

Table 2
Experimental design of the conjoint experiment.

Attributes	AI	Human
Provider	Skin-AI-D	Dermatologist
Costs	$0	$0
	$30	$30
	$150	$150
Expertise	high	high
	low	low
Privacy Policy	strict	strict
	lax	lax
Results	detailed information	detailed information
	basic information	basic information

Note: The exact wording of the attribute values can be found in the online supplements (see Table S2).

experience with skin cancer screenings and history of skin cancer diagnosis).

2.3. Measures

Medical mistrust: We measured participants' mistrust of healthcare institutions using a short version of the Medical Mistrust Index (MMI, LaVeist et al. (2009)) with seven questions (e.g., "You'd better be cautious when dealing with healthcare organisations.") rated on a 4-point scale from 1 (*strongly disagree*) to 4 (*strongly agree*). The average of these items was computed, and the scale showed good internal consistency according to the interpretation of its Cronbach's Alpha value ($\alpha = 0.86$).

Attitudes towards AI: We assessed attitudes towards AI technology with the 21-item General Attitudes towards Artificial Intelligence Scale (GAAIS, Schepman and Rodway (2020), e.g., "Organisations use artificial intelligence unethically.") on a 5-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). The scale consists of two subscales (positive and negative attitudes), which should not be combined according to its developers. Both subscales showed good to excellent internal consistency ($\alpha_{\text{positive}} = 0.92$, $\alpha_{\text{negative}} = 0.86$).

Health anxiety: Health anxiety was assessed using the Short Health Anxiety Inventory (SHAI, Salkovskis et al. (2002)) encompassing 18 items. For each item, participants are presented with four statements (e.g., (a) "I do not worry about my health.", (b) "I occasionally worry about my health.", (c) "I spend much of my time worrying about my health.", and (d) "I spend most of my time worrying about my health.") and are asked to select the statement that best describes their feelings. Each statement has a value assigned ranging from 0 to 3, which is summed across all items to obtain a total score. The scale showed excellent internal consistency ($\alpha = 0.91$).

Data security and privacy concerns: Participant's concerns regarding data security and privacy were measured through three subscales of the Concerns for Information Privacy scale (CFIP, Angst & Agarwal (2009)): collection (4 items, e.g., "It usually bothers me when healthcare entities ask me for personal information."), unauthorised access (3 items, e.g., "Healthcare entities should devote more time and effort to preventing unauthorised access to personal information."), and secondary use (4 items, e.g., "Healthcare entities should not use personal information for any purpose unless it has been authorised by the individuals who provided the information.") on a 5-point scale ranging from 1 (*completely disagree*) to 5 (*completely agree*). The scale showed good internal consistency ($\alpha = 0.86$).

Trust in provider: Participants' trust in both the selected and non-selected providers was assessed using a single item ("How much do you trust [the provider]?") adopted from Gaertig and Simmons (2018) on a 7-point scale from 1 (*not at all*) to 7 (*extremely*).

Perceived accuracy of the screening: The respondents were asked to rate the accuracy of the screening results for both the selected and non-selected providers, on a single item ranging ("How accurate do you think the results of [the provider] will be when it comes to detecting cancerous skin spots?") from 0% to 100%.

Perceived quality of the provider: The quality of the provider was measured using three items (e.g., "How competent is [the provider]?") also adopted from Gaertig and Simmons (2018) on a 7-point scale from 1 (*not at all*) to 7 (*extremely/definitely*). The average of the items was calculated, and the quality scale both for the selected provider and the non-selected showed good internal consistency ($\alpha_{\text{selected}} = 0.86$; $\alpha_{\text{non-selected}} = 0.91$).

Waiting time before choosing the other provider: Participants were queried about how many days they would be willing to wait for the screening results before choosing the other provider or to select the option *"I would never change my chosen screening option"*.

Second opinion: The survey included two questions asking participants about the likelihood that they would seek a second opinion upon receiving a positive or a negative screening result from their chosen provider (e.g., "After receiving the results from [the provider], would you seek additional skin cancer screening to get a second opinion in case of a *positive screening result* (i.e., abnormalities identified)?") on a 7-point Likert ranging from 1 (*not at all*) to 7 (*definitely*).

We measured some additional variables (as preregistered) not reported in the main analysis. Information regarding these measurements can be found in the online supplements.

2.4. Data analysis

Statistical analyses were carried out using R version 4.3.2 (R Core Team, 2024). To test the hypotheses for our conjoint experiments, average marginal component effects (AMCEs) were estimated and plotted using the *cjoint* package (Barari et al., 2023). An AMCE "represents a causal effect of an attribute value against another, averaged over possible interaction effects with the other included attributes, as well as over possible heterogeneous effects across respondents" (Bansak et al., 2021, p. 15). AMCEs can be interpreted as the average change in probability for an attribute value compared to a reference value. We also calculated conditional AMCEs to explore interactions between respondents' characteristics and their decisions based on the vignettes. To test the remaining hypotheses, we utilised logistic and linear mixed-effects regression models, and mean difference tests.

3. Results

3.1. Preference for experienced dermatologists with strict privacy policies and no cost

Initially, we examined the impact of manipulated attributes on participants' decisions regarding their skin cancer screening provider. Overall, the AMCE outcomes were consistent across both studies, as depicted in Fig. 2 (output tables can be found in the online supplements). This indicates that the presentation order exerted minimal influence on decisions. In both studies, the type of provider emerged as the most pivotal determinant of the decision. Across all attributes, the dermatologist was 40.24 % ($SD = 0.02$) and 43.41 % ($SD = 0.05$) more likely to be selected than the mHealth app in the respective studies. A higher price was consistently the most influential attribute besides the type of provider itself. Participants were 22.75 % ($SD = 0.02$) and 23.08 % ($SD = 0.05$) less inclined to opt for a provider with a high price than a free one. Medium price affected the provider decision only in the main study significantly, being 8.70 % ($SD = 0.02$) less likely to be selected as the free version compared to the replication with 1.54 % ($SD = 0.04$). The proposed level of the provider's expertise impacted participants' decisions in both studies. The provider with low expertise was 14.53 % ($SD = 0.02$) and 15.23 % ($SD = 0.05$) less likely to be chosen. Lax data privacy was of comparable, albeit minor, relevance, with the provider being 4.73 % ($SD = 0.02$) and 3.47 % ($SD = 0.04$) less likely selected than one with a strict policy; however, only the effect in the main study was statistically significant. Neither study found that the level of detail with which the results will be presented to the user significantly impacted their decision (2.18 %; $SD = 0.02$ and 4.64 %; $SD = 0.04$). Interpreting the effect size of AMCE values depends on the context of the study. However, most AMCE values in our studies are large in comparison to the median AMCE value of 0.05, i.e., 5 %, found in an analysis of 15 highly cited conjoint experiences (Schuessler and Freitag, 2020).

3.2. Stable preferences across demographics

We calculated conditional AMCEs to explore interactions between respondents' attributes and their provider decisions within the main study. We examined the following characteristics: age, gender, and ethnicity. Regarding age, the sample was divided into younger adults (18–40), middle-aged adults (41–65) and older adults (>65). We compared people who self-identified as female, male, and other.

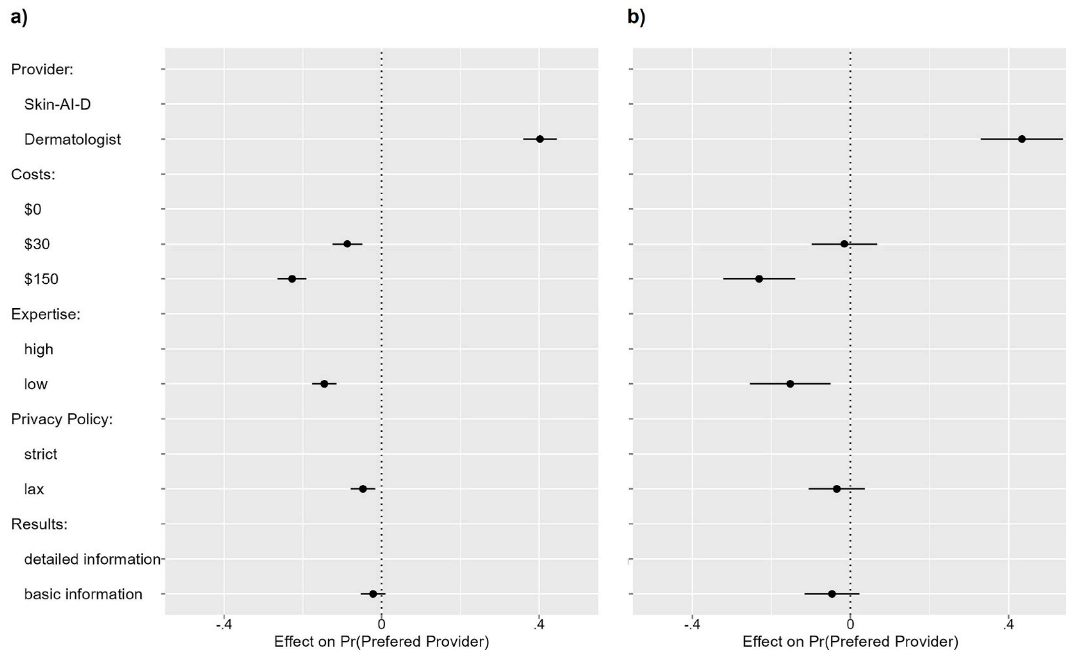

Fig. 2. Effects of provider attributes on the probability of provider choice. The x-axis represents the AMCEs, i.e., the effect on the probability of a provider being preferred. Positive values indicate a higher probability of provider preference, and negative values indicate a lower probability. a) Results from the main study. b) Results from the replication study. Bars represent 95% confidence intervals. The attribute value without a point value and confidence interval represents the reference category for each attribute.

Additionally, participants were grouped into white, black or African American, and other non-white ethnicities. Remarkably, consistent trends emerged across subgroups (see Fig. 3a–c), with only a few noteworthy exceptions. Older adults were less price-sensitive than younger and middle-aged adults. Only for younger adults the level of result detail was a somewhat relevant factor for their decision (Fig. 3a). Female participants were more reluctant to select a provider with a lax privacy policy than male participants (Fig. 3b). The data suggest that white participants were slightly more price-sensitive than the two non-white groups, although the difference in sample sizes impedes the evaluation of these distinctions (Fig. 3c). All output tables can be found in the online supplements.

3.3. *Individuals mistrusting the medical system are more open to mHealth apps*

We also examined further user characteristics influencing the provider decision: medical mistrust, attitudes towards AI technology, health anxiety, and data security and privacy concerns. Table 3 presents findings from the logistic regression, with the provider decision serving as the dependent variable. The equation corresponding to the output can be found in the online supplements. As anticipated, individuals with higher levels of medical mistrust and people with a more favourable attitude towards AI were more inclined to opt for the mHealth app for their skin cancer screening. Conversely, people expressing more negative attitudes towards AI technology were less likely to select the mHealth app. Contrary to our expectations, we did not observe evidence linking health anxiety or data security and privacy concerns to the provider decision. Conditional AMCE results for medical mistrust, health anxiety, and data security and privacy concerns can be found in the online supplements. We reanalysed the regression model, incorporating participants' demographics as control variables to account for potential confounding effects and ensure model reliability. The inclusion of control variables did not alter the main analysis outcomes significantly. The additional results can be found in the online supplements.

3.4. *Dermatologists' results are perceived as more trustworthy, accurate and of higher quality*

In addition to the binary provider choice, we investigated participants' perspectives on (a) trust, (b) perceived accuracy, and (c) overall quality regarding the two providers. Three linear multilevel models were deployed to examine the trust score, perceived accuracy, and screening quality for both selected and non-selected providers as dependent variables, with provider attributes as predictors. All models included fixed effects for the predictors and a random effect for the participants to account for the non-independence of observations. The equations corresponding to the outputs are in the online supplements. Findings indicated that participants exhibited higher levels of trust, accuracy rating, and perceived quality towards dermatologists compared to the mHealth app (see Table 4). The provider's expertise was the sole additional attribute demonstrating a substantial influence on respondents' trust in and perceived accuracy of the providers. High price, low expertise, and a lenient privacy policy were negatively related to anticipated provider quality. Again, we reanalysed all three regression models, incorporating participants' demographics as control variables. The inclusion of control variables did not alter the main analysis outcomes significantly. Consequently, we included only the initial regression results in the manuscript, while the additional results are in the online supplements.

Next, we compared trust, accuracy, and quality ratings by the providers to see if ratings depended on whether the provider was selected or not. The results indicated that even people who opted for the mHealth app trusted the dermatologist more (see Fig. 4a). Moreover, both accuracy (Fig. 4b) and quality (Fig. 4c) ratings for the dermatologist were consistently higher, irrespective of the chosen provider.

3.5. *Anticipating faster results but strong need for second opinions when opting for AI*

Furthermore, participants were asked how long they would be willing to wait for the screening results, encompassing the time until the appointment, before choosing the other provider or to express their

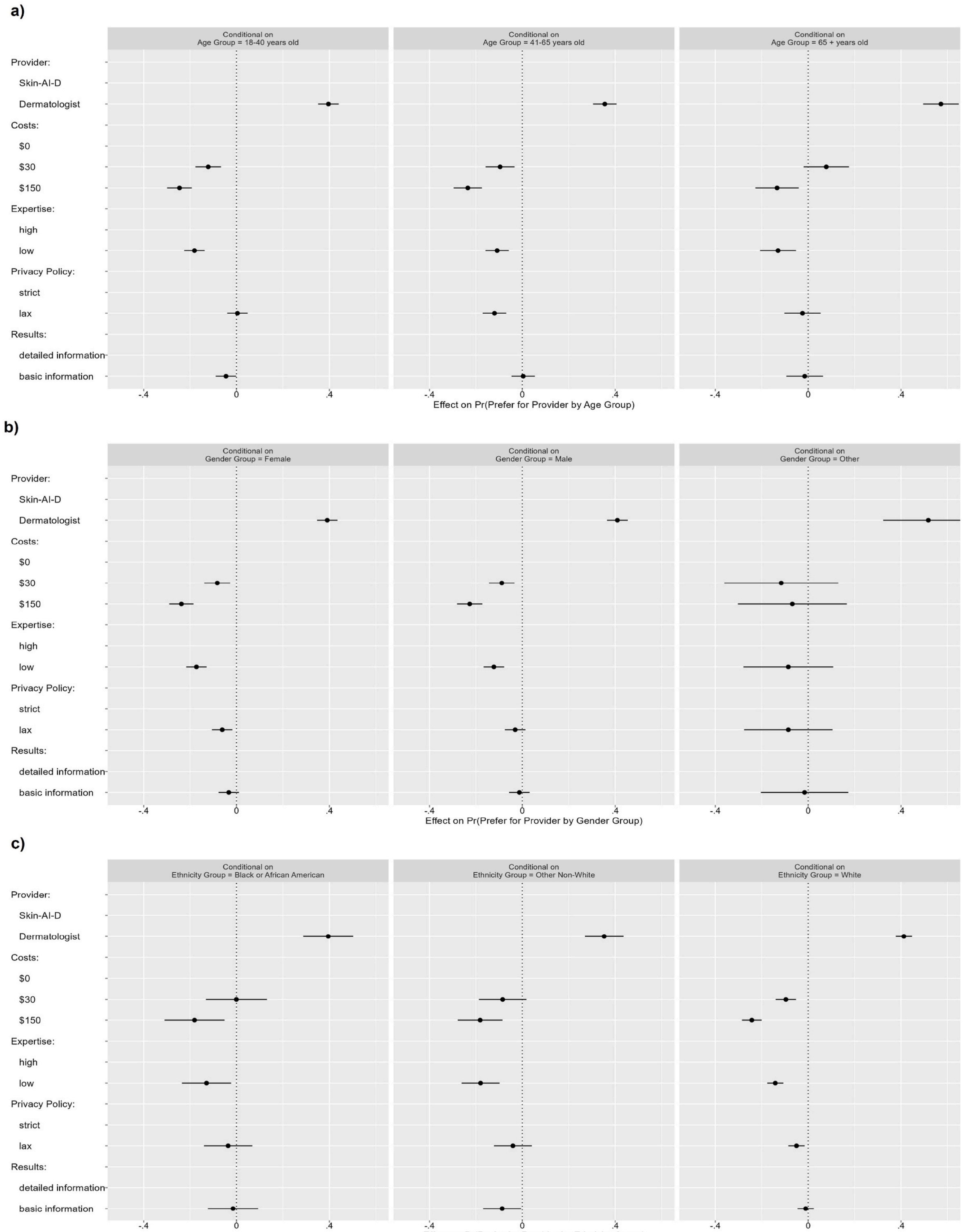

Fig. 3. Effects of provider attributes on the probability of provider choice by a) age, b) gender, and c) ethnicity. The x-axis represents the AMCEs, i.e., the effect on the probability of a provider being preferred. Positive values indicate a higher probability of provider preference, and negative values indicate a lower probability. Bars represent 95% confidence intervals. The attribute value without a point value and confidence interval represents the reference category for each attribute.

unwillingness to change their screening provider. Only 5.34 % of individuals indicated a steadfast commitment to the mHealth app, in contrast to 34.82 % of respondents who expressed they would never switch from the dermatologist. The remaining 59.84 % of participants expressed openness to change providers. Overall, the respondents were willing to wait slightly longer for an appointment and the results from the dermatologist ($M = 13.12$, $SD = 65.29$) compared to the mHealth app ($M = 8.67$, $SD = 9.65$) before contemplating a provider switch ($z = 2.35$, $p = 0.019$, $r = 0.08$).

Subsequently, we analysed participants' responses concerning their intent to pursue a second opinion, contingent on receiving a positive or negative screening result. Unsurprisingly, individuals were more

Table 3
Logistic regression results using provider decision as the criterion.

		95% CI					
Predictor	OR	2.5 %	97.5 %	Est.	SE	z	p
Intercept	54.81	21.52	142.71	4.00	0.48	8.30	**<0.001**
Medical mistrust	0.55	0.44	0.68	−0.60	0.11	−5.55	**<0.001**
Negative attitudes towards AI	1.37	1.15	1.63	0.31	0.09	3.49	**<0.001**
Positive attitudes towards AI	0.44	0.37	0.53	−0.81	0.09	−9.06	**<0.001**
Health anxiety	1.01	0.99	1.02	0.01	0.01	1.00	0.315
Data security and privacy concerns	1.09	0.91	1.30	0.09	0.09	0.96	0.337

Note: OR = odds ratio, Est. = coefficient, SE = standard error, z = Est./SE, p = probability value, $R^2_{Tjur} = 0.108$.

inclined to seek a second opinion after receiving a positive screening result suggesting a potentially cancerous skin abnormality ($t = 28.19$, $p < 00.001$, $d = 0.79$). Regardless of the screening outcome, participants were significantly more likely to pursue a second opinion when the initial result was generated by a mHealth app (see Fig. 5). Table 5 presents a summary of all findings and their assessment against our hypotheses.

4. Discussion

Skin cancer is a prevalent health concern with significant implications for public health and individual well-being. In this work, we examined provider and user characteristics influencing people's preferences for AI-enabled mHealth apps and dermatologists for skin cancer screenings.

We found that the type of provider was the most influential factor in participants' decisions for their skin cancer screening. The dermatologist was chosen significantly more often than the mHealth app (supporting H1). This finding underscores people's enduring preference for seeking human healthcare providers over self-administered mHealth apps, as demonstrated in previous research (Baldauf et al., 2020; Jahn et al., 2022; Longoni et al., 2019). The greater likelihood of opting for the dermatologist while accounting for other provider attributes suggests a strong inclination towards human over algorithmic support, reflecting one of the defining features of algorithmic aversion (Jussupow et al., 2020). Another central aspect of algorithmic aversion is a tendency to evaluate human actions more favourably than algorithms (Jussupow et al., 2020). Our findings support this assumption, as participants exhibited notably higher levels of trust in the dermatologist, along with consistently higher ratings for perceived accuracy and quality, in comparison to the mHealth app (supporting H10). Interestingly, even participants who chose the mHealth app rated the dermatologist on average more favourably across all evaluation criteria. These results suggest that participants perceived human dermatologists as more reliable and capable of delivering accurate and high-quality skin cancer screenings, consistent with previous research (Jahn et al., 2022; M. K. Lee and Rich, 2021). Algorithmic aversion is also affected by people's perception of human involvement in the usage of an algorithm as a "human-algorithm hybrid" (Jussupow et al., 2020). Individuals exhibit a stronger aversion to mHealth apps, when they perceive them as replacements rather than aids to physicians (Jahn et al., 2022; Jutzi et al., 2020), which might also be relevant in the context of the current study. Although the vignette

Table 4
Multilevel regression models with trust, accuracy, and quality as criteria.

	Trust			Accuracy			Quality		
Predictor	Est.	SE	p	Est.	SE	p	Est.	SE	p
Intercept	4.46	0.06	**<0.001**	70.72	0.78	**<0.001**	4.84	0.06	**<0.001**
Provider [Dermatologist]	1.71	0.05	**<0.001**	15.71	0.54	**<0.001**	1.47	0.04	**<0.001**
Costs [$30]	0.02	0.06	0.743	−0.07	0.72	0.927	0.00	0.05	0.964
Costs [$150]	−0.07	0.06	0.253	−0.54	0.72	0.448	−0.10	0.05	**0.040**
Expertise [low]	−0.50	0.05	**<0.001**	−6.43	0.59	**<0.001**	−0.51	0.04	**<0.001**
Privacy Policy [lax]	−0.09	0.05	0.054	−0.97	0.59	0.099	−0.09	0.04	**0.039**
Results [basic]	−0.01	0.05	0.820	0.72	0.59	0.221	−0.01	0.04	0.851

Note: N = 1,591, Observations = 3,182, SE = standard error, t = Est./SE, p = probability value, Marginal $R^2_{Trust} = 0.294$, Conditional $R^2_{Trust} = 0.327$, Marginal $R^2_{Accuracy} = 0.201$, Conditional $R^2_{Accuracy} = 0.343$, Marginal $R^2_{Quality} = 0.299$, Conditional $R^2_{Quality} = 0.314$, outputs with confidence intervals and random effects can be found in the online supplements.

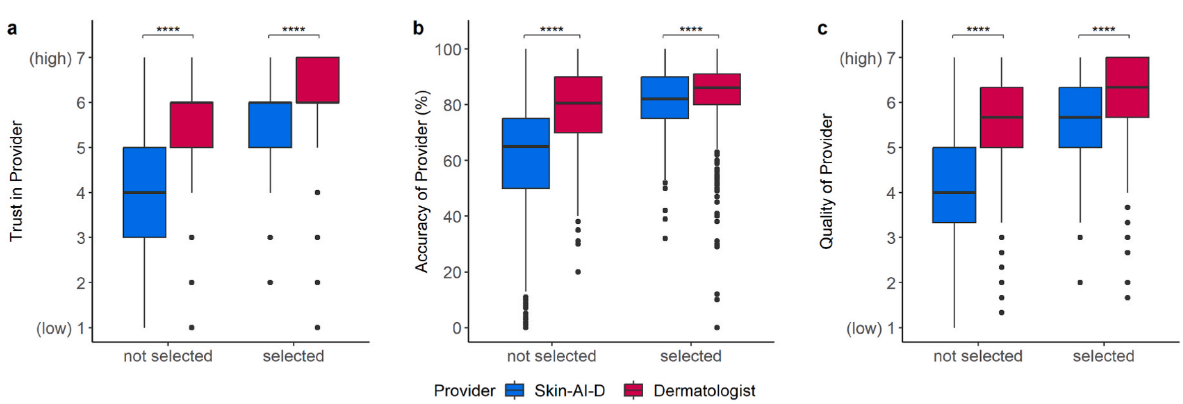

Fig. 4. Ratings of a) trust in, b) perceived accuracy, and c) quality of provider split by selected and non-selected options. The boxplots show the 25th to 75th percentiles (lower and upper hinges) with the median depicted by the central line; the whiskers extend to a maximum of 1.5 × interquartile range (IQR) beyond the boxes.

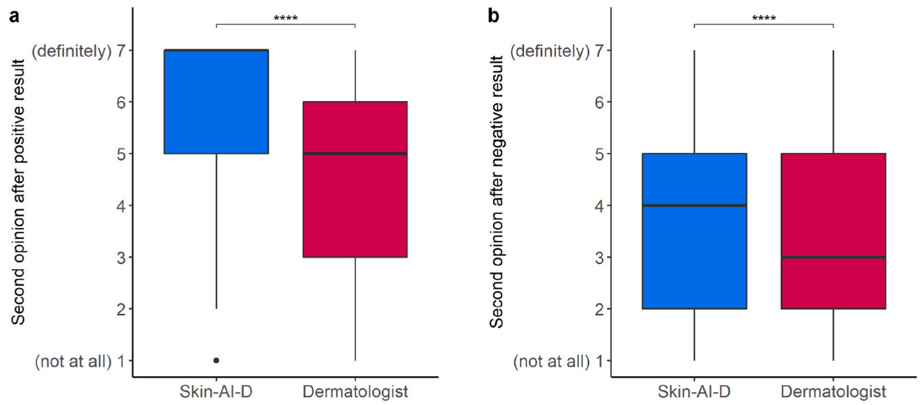

Fig. 5. Likelihood of seeking a second opinion by provider and result type with a) after a positive result and b) after a negative result. The boxplots show the 25th to 75th percentiles (lower and upper hinges) with the median depicted by the central line; the whiskers extend to a maximum of 1.5 × interquartile range (IQR) beyond the boxes.

Table 5
Result summary.

#	Summary of findings	Analysis	Assessment
H1	Preference for the dermatologist over the mHealth app	AMCE	Supported
H2a	Preference for provider with low cost compared to high, and medium cost (only main study)	AMCE	Partially supported
H2b	High price was the most influential attribute besides provider type itself	AMCE	Partially supported
H3a	Preference for provider with high over low expertise	AMCE	Supported
H3b	Expertise was the second most influential attribute besides the provider type itself and cost	AMCE	Supported
H4a	Preference for provider with strict over lax privacy policy (only main study)	AMCE	Partially supported
H4b	Strictness of privacy policy was less important than cost and expertise	AMCE	Supported
H5a	Non-significant preference for provider with detailed over basic screening results	AMCE	Not supported
H5b	Level of result detail was less important than cost and expertise	AMCE	Supported
H6	Individuals with higher medical mistrust were more inclined to opt for the mHealth app	Logistic regression	Supported
H7	Individuals with negative attitudes towards AI were more inclined to opt for the dermatologist	Logistic regression	Supported
H8	Little evidence linking health anxiety to the provider decision	Conditional AMCE	Not supported
H9	Little evidence linking data security and privacy concerns to the provider decision	Conditional AMCE	Not supported
H10	The dermatologist was perceived as more trustworthy, accurate, and of higher quality	Multilevel regression	Supported
H11	Only expertise showed a consistent and substantial influence on trust, accuracy, and quality ratings	Multilevel regression	Partially supported
H12	Respondents were willing to wait slightly longer for the dermatologist compared to the mHealth app	Wilcoxon test	Supported

Note: # = Hypothesis number, AMCE results in Tables S3 and S4; Logistic regression in Table 3, Conditional AMCE results in Figs. S3 and S4; Multilevel regression results in Table 4; Wilcoxon test: $z = 2.35$, $p = 0.019$, $r = 0.08$.

explicitly mentioned that a positive result would lead to a referral to a dermatologist for further examinations, it is plausible that respondents may have interpreted the scenario as implying that the app is intended to substitute professional screenings completely. Consequently, developers of screening apps should strongly emphasise that their product is designed to aid in self-examinations and should not be seen as a substitute for professional advice and may even ask users to confirm that they understand this.

Price sensitivity was another factor strongly affecting people's decisions, with participants showing a preference for free or lower-cost skin cancer screening options (partly supporting H2a). This aligns with previous research (Longoni et al., 2019) and provides further evidence that financial barriers hinder people from seeking optimal preventive cancer services (American Cancer Society, n.d.). Receiving a free screening even outweighed the provider's expertise for the participants (partially supporting H2b). We anticipated that a pricier screening option would be perceived more favourably since a higher price might be perceived as an indicator of higher quality. Surprisingly, costs did not affect trust in and the perceived accuracy of the screening (contrary to H11). Instead, a high price ($150) was even associated with lower provider quality ratings. Hence, mHealth apps should be offered either for free or at a nominal cost to incentivise their adoption and mitigate health disparities. Governments and health insurance providers could facilitate cost reduction by offering subsidies or service coverage.

As anticipated, the provider's expertise significantly influenced the participants' choice in both studies, as individuals were more likely to opt for providers with greater expertise (supporting H3a). This finding is consistent with previous literature demonstrating consumer preferences for highly accurate mHealth apps (Haggenmüller et al., 2021; Longoni et al., 2019). Recognising the challenge of obtaining data on accuracy levels for human providers, we opted to employ expertise as a proxy attribute. Besides the type of screening provider (mHealth app vs. dermatologist), expertise was the only provider attribute that consistently affected their evaluation (supporting H3b). Screening providers with lower expertise were trusted less, and both their accuracy and quality were perceived as inferior compared to providers with higher expertise (supporting H11). The limited adoption of mHealth apps for skin cancer screening may be attributed to their performance still falling short of specialists' standards (Freeman et al., 2020; Jahn et al., 2022). Developers should prioritise enhancing the accuracy of their products in real-world scenarios. Additionally, obtaining CE certification or FDA approval might foster confidence among consumers regarding the safety and effectiveness of the product, which, in turn, might increase user acceptance.

While data protection and model explainability are widely discussed topics in AI research and product regulation (Hummelsberger et al., 2023; Rajpurkar et al., 2022), our findings suggest that these factors might be less relevant for consumer decisions (supporting H4b and H5b). Although a more stringent privacy policy was a significant positive predictor in the main study (supporting H4a), its effect size was small, and we did not find this effect to be significant in the replication study. Moreover, we could not find evidence for a significant association between the level of detail given explaining the screening results and participants' provider decisions (contrary to H5a). Additional research is needed to ascertain whether the lack of statistical significance accurately represents consumer preferences in this domain (i.e., a lack of perceived benefit from a more detailed explanation of how the results were generated) or if the selected research methodology may influence it since previous findings suggested, that users prefer more transparent software (Haggenmüller et al., 2021; Jutzi et al., 2020). That being stated, regulatory bodies must take a more proactive role in enforcing data protection and model transparency guidelines for mHealth apps, as consumer preferences may not provide sufficient motivation for app developers to adhere to these standards.

We also investigated several user characteristics potentially influencing the screening provider decision. In accordance with the literature (M. K. Lee and Rich, 2021), our results indicated that individuals with higher medical mistrust were more inclined to choose the mHealth app (supporting H6). People who distrust the medical system are often more sceptical towards medical professionals, stemming from past negative encounters such as perceived dismissals or discrimination (C. Lee et al., 2009; M. K. Lee and Rich, 2021; Thompson et al., 2004). Particularly among minorities and marginalised groups, higher levels of medical mistrust are often reported, contributing to a decreased propensity to consult medical professionals (e.g., C. Lee et al., 2009; Powell et al., 2019; Thompson et al., 2004). This trend may result in exacerbated health disparities and worsened health outcomes. Consequently, mHealth apps could play a crucial role in enhancing medical services for this demographic, provided they perceive the technology as less biased. However, to achieve better health outcomes for this specific group, mHealth apps should assist users in overcoming their reluctance to seek professional medical assistance when it is required. Predictably, individuals with favourable attitudes towards AI technology were more inclined to opt for the AI-enabled app (supporting H7). According to the classic Technology Acceptance Model (TAM, Davis (1989)), a person's intention to use a technology is determined by their attitude towards it. This attitude, in turn, is heavily influenced by the perceived usefulness and perceived ease of use associated with the technology. Therefore, developers should highlight how mHealth apps for skin cancer screening can enhance the quality of self-examinations and make it as easy as possible for consumers to use the apps.

Our participants were willing to wait only slightly longer for results from dermatologists compared to mHealth apps before considering a provider switch (supporting H12). This finding suggests that people might value the immediacy of technologically generated results while being content with longer waiting times for human providers. Delays in screenings can adversely affect patient outcomes by leading to delayed treatments, a common medical error (Gaube et al., 2021). Consequently, mHealth app developers should highlight the key benefits of their products in providing immediate initial screening results that might expedite treatment planning when needed. Participants' willingness to seek a second opinion varied depending on the screening results and the provider type. Respondents were more likely to seek a second opinion when the initial result was generated by a mHealth app, especially when the result was positive. This finding underscores users' reluctance to rely exclusively on AI algorithms for healthcare decisions, in line with another defining aspect of algorithmic aversion (Jussupow et al., 2020). Establishing connectivity between the mHealth app and medical professionals, with the capability to automatically schedule follow-up examinations with a dermatologist when the app detects something of concern, could accelerate treatment planning and align better with users' expectations. Such an approach has the potential to boost acceptance among both healthcare professionals and consumers.

4.1. Limitations and recommendations for future research

The present study has some limitations. Hypothetical and oversimplified scenarios may not fully capture real-world decision-making, and participants' actual behaviour might differ when faced with real choices. Related to this, making sure that attributes were relevant for both the mHealth app and the dermatologist was challenging and potentially reduced the external validity for the latter. Moreover, participants were recruited from a crowdsourcing platform, which could introduce issues related to the representativeness of the sample (e.g., elderly online platform users might be generally more open to technology than their peers not using such a platform). The uneven sample sizes in certain subgroup analyses, including small sub-samples, somewhat limit the interpretability and generalizability of the affected interaction results. Moreover, the sample size of the replication study was too small to ensure the detection of smaller effects in the conjoint experiment. Finally, the cross-sectional design does not allow for the capture of changes in effects over time, such as shifts in consumer preferences.

Future studies should be conducted among a sufficiently large, representative sample of individuals who are actively seeking skin cancer screening, providing them with more comprehensive descriptions of various service providers to enhance external validity. Additionally, longitudinal studies should assess changes in people's general attitudes towards utilising mHealth tools to support their skin self-examination. Finally, further research on the effects of relevant user characteristics on their reactions to decision aids from algorithms or humans is needed to provide a more comprehensive research framework for the study of algorithm aversion.

5. Conclusion

This research provides a comprehensive understanding of the factors influencing people's preferences and decisions between AI-enabled mHealth apps and dermatologists for skin cancer screenings. Our findings provide valuable insights for technology developers, healthcare providers, and policymakers. First, developers of mHealth apps should emphasise that their technology is intended to assist with self-examinations and is not meant to replace professional advice. This should be underlined by ensuring connectivity with medical professionals for follow-up examinations. Second, mHealth apps should be made available at a low cost, must exhibit excellent performance, and provide speedy results to ensure user acceptance. Third, matters of data protection, security, and model transparency should be subject to oversight by regulatory bodies rather than relying on them to be determined by consumer preferences. Finally, skin cancer screening apps might hold significant potential in addressing healthcare gaps in underserved regions. Therefore, providing safe, reliable, and user-friendly products tailored to these communities might improve acceptance and adoption, which would contribute to enhancing accessibility to medical services and mitigating health disparities.

CRediT authorship contribution statement

Susanne Gaube: Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Isabell Biebl:** Writing – review & editing, Investigation, Formal analysis, Conceptualization. **Magdalena Karin Maria Engelmann:** Writing – review & editing, Investigation, Formal analysis, Conceptualization. **Anne-Kathrin Kleine:** Writing – review & editing. **Eva Lermer:** Writing – review & editing, Supervision, Resources.

Data availability

The dataset, analysis syntax, supplementary information, and all questionnaire items are available online (https://osf.io/xc57g/).

Acknowledgements

We want to thank all participants contributing to this research. This research was funded by a grant from the Volkswagen Foundation (Grant #: 98525).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2024.116871.

References

American Academy of Dermatology Association. (n.d.). Skin cancer. Skin Cancer. Retrieved 5 February 2023, from https://www.aad.org/media/stats-skin-cancer.

American Cancer Society. (n.d.). Skin cancer. Skin Cancer. Retrieved 5 February 2023, from https://www.cancer.org/cancer/skin-cancer.html.

Angst & Agarwal, 2009. Adoption of electronic health records in the presence of privacy concerns: the elaboration likelihood model and individual persuasion. MIS Q. 33 (2), 339–370. https://doi.org/10.2307/20650295.

Baldauf, M., Fröehlich, P., Endl, R., 2020. Trust me, I`m a doctor – user perceptions of AI-driven apps for mobile health diagnosis. 19th International Conference on Mobile and Ubiquitous Multimedia, pp. 167–178. https://doi.org/10.1145/3428361.3428362.

Bansak, K., Hainmueller, J., Hopkins, D.J., Yamamoto, T., 2021. Conjoint survey experiments. In: Druckman, J., Green, D.P. (Eds.), Advances in Experimental Political Science, first ed. Cambridge University Press, pp. 19–41. https://doi.org/10.1017/9781108777919.004.

Barari, S., Berwick, E., Hainmueller, J., Hopkins, D., Liu, S., Strezhnev, A., Yamamoto, T., 2023. *Package 'cjoint'* (2.1.1) [Computer software]. https://cran.r-project.org/web/packages/cjoint/cjoint.pdf.

Berwick, M., Begg, C.B., Fine, J.A., Roush, G.C., Barnhill, R.L., 1996. Screening for cutaneous melanoma by skin self-examination. JNCI Journal of the National Cancer Institute 88 (1), 17–23. https://doi.org/10.1093/jnci/88.1.17.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. 13 (3), 319–340. https://doi.org/10.2307/249008.

Dietvorst, B.J., Simmons, J.P., Massey, C., 2015. Algorithm aversion: people erroneously avoid algorithms after seeing them err. J. Exp. Psychol. Gen. 144 (1), 114–126. https://doi.org/10.1037/xge0000033.

Erickson, G.M., Johansson, J.K., 1985. The role of price in multi-attribute product evaluations. J. Consum. Res. 12 (2), 195–199. https://doi.org/10.1086/208508.

Feng, H., Berk-Krauss, J., Feng, P.W., Stein, J.A., 2018. Comparison of dermatologist density between urban and rural counties in the United States. JAMA Dermatology 154 (11), 1265–1271. https://doi.org/10.1001/jamadermatol.2018.3022.

Freeman, K., Dinnes, J., Chuchu, N., Takwoingi, Y., Bayliss, S.E., Matin, R.N., Jain, A., Walter, F.M., Williams, H.C., Deeks, J.J., 2020. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. BMJ 368, m127. https://doi.org/10.1136/bmj.m127.

Gaertig, C., Simmons, J.P., 2018. Do people inherently dislike uncertain advice? Psychol. Sci. 29 (4), 504–520. https://doi.org/10.1177/0956797617739369.

Gaube, S., Cecil, J., Wagner, S., Schicho, A., 2021. The relationship between health IT characteristics and organizational variables among German healthcare workers. Sci. Rep. 11 (1), 17752 https://doi.org/10.1038/s41598-021-96851-1.

Google Health. (n.d.). Identify Skin Conditions with DermAssist—Google Health. Retrieved 12 January 2024, from https://health.google/consumers/dermassist/.

Green, P.E., Krieger, A.M., Wind, Y., 2001. Thirty years of conjoint analysis: reflections and prospects. Interfaces 31 (3), 56–73. https://doi.org/10.1287/inte.31.3s.56.9676.

Haggenmüller, S., Krieghoff-Henning, E., Jutzi, T., Trapp, N., Kiehl, L., Utikal, J.S., Fabian, S., Brinker, T.J., 2021. Digital natives' preferences on mobile artificial intelligence apps for skin cancer diagnostics: survey study. JMIR mHealth and uHealth 9 (8), e22909. https://doi.org/10.2196/22909.

Hainmueller, J., Hangartner, D., Yamamoto, T., 2015. Validating vignette and conjoint survey experiments against real-world behavior. Proc. Natl. Acad. Sci. USA 112 (8), 2395–2400. https://doi.org/10.1073/pnas.1416587112.

Hainmueller, J., Hopkins, D.J., Yamamoto, T., 2014. Causal inference in conjoint analysis: understanding multidimensional choices via stated preference experiments. Polit. Anal. 22 (1), 1–30. https://doi.org/10.1093/pan/mpt024.

Hamidi, R., Peng, D., Cockburn, M., 2010. Efficacy of skin self-examination for the early detection of melanoma. Int. J. Dermatol. 49 (2), 126–134. https://doi.org/10.1111/j.1365-4632.2009.04268.x.

Hummelsberger, P., Koch, T.K., Rauh, S., Dorn, J., Lermer, E., Raue, M., Hudecek, M.F.C., Schicho, A., Colak, E., Ghassemi, M., Gaube, S., 2023. Insights on the Current State and Future Outlook of AI in Health Care: Expert Interview Study. JMIR AI. https://doi.org/10.2196/47353.

Jahn, A.S., Navarini, A.A., Cerminara, S.E., Kostner, L., Huber, S.M., Kunz, M., Maul, J.-T., Dummer, R., Sommer, S., Neuner, A.D., Levesque, M.P., Cheng, P.F., Maul, L.V., 2022. Over-detection of melanoma-suspect lesions by a CE-certified smartphone app: performance in comparison to dermatologists, 2D and 3D Convolutional Neural Networks in a prospective data set of 1204 pigmented skin lesions involving patients' perception. Cancers 14 (15), 3829. https://doi.org/10.3390/cancers14153829.

Jussupow, E., Benbasat, I., Heinzl, A., 2020. Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion. *Proceedings Of the 28th European Conference on Information Systems (ECIS), an Online AIS Conference.* 28th European Conference on Information Systems (ECIS2020) – A Virtual AIS Conference. Online. https://aisel.aisnet.org/ecis2020_rp/168.

Jutzi, T.B., Krieghoff-Henning, E.I., Holland-Letz, T., Utikal, J.S., Hauschild, A., Schadendorf, D., Sondermann, W., Fröhling, S., Hekler, A., Schmitt, M., Maron, R.C., Brinker, T.J., 2020. Artificial intelligence in skin cancer diagnostics: the patients' perspective. Front. Med. 7, 233. https://doi.org/10.3389/fmed.2020.00233.

Kasparian, N.A., McLoone, J.K., Meiser, B., 2009. Skin cancer-related prevention and screening behaviors: a review of the literature. J. Behav. Med. 32 (5), 406–428. https://doi.org/10.1007/s10865-009-9219-2.

Kong, F.W., Horsham, C., Ngoo, A., Soyer, H.P., Janda, M., 2021. Review of smartphone mobile applications for skin cancer detection: what are the changes in availability, functionality, and costs to users over time? Int. J. Dermatol. 60 (3), 289–308. https://doi.org/10.1111/ijd.15132.

LaVeist, T.A., Isaac, L.A., Williams, K.P., 2009. Mistrust of health care organizations is associated with underutilization of health services. Health Serv. Res. 44 (6), 2093–2105. https://doi.org/10.1111/j.1475-6773.2009.01017.x.

Lee, C., Ayers, S.L., Kronenfeld, J.J., 2009. The association between perceived provider discrimination, healthcare utilization and health status in racial and ethnic minorities. Ethn. Dis. 19 (3), 330–337. https://doi.org/10.13016/h7o5-zl06.

Lee, M.K., Rich, K., 2021. Who is included in human perceptions of AI?: trust and perceived fairness around healthcare AI and cultural mistrust. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems 1–14. https://doi.org/10.1145/3411764.3445570.

Longoni, C., Bonezzi, A., Morewedge, C.K., 2019. Resistance to medical artificial intelligence. J. Consum. Res. 46 (4), 629–650. https://doi.org/10.1093/jcr/ucz013.

Matsumoto, M., Secrest, A., Anderson, A., Saul, M.I., Ho, J., Kirkwood, J.M., Ferris, L.K., 2018. Estimating the cost of skin cancer detection by dermatology providers in a large health care system. J. Am. Acad. Dermatol. 78 (4), 701–709. https://doi.org/10.1016/j.jaad.2017.11.033.

Powell, W., Richmond, J., Mohottige, D., Yen, I., Joslyn, A., Corbie-Smith, G., 2019. Medical mistrust, racism, and delays in preventive health screening among African-American men. Behav. Med. 45 (2), 102–117. https://doi.org/10.1080/08964289.2019.1585327.

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing [Computer software]. https://www.R-project.org/.

Rajpurkar, P., Chen, E., Banerjee, O., Topol, E.J., 2022. AI in health and medicine. Nat. Med. 28 (1), 31–38. https://doi.org/10.1038/s41591-021-01614-0.

Salkovskis, P.M., Rimes, K.A., Warwick, H.M.C., Clark, D.M., 2002. The Health Anxiety Inventory: development and validation of scales for the measurement of health anxiety and hypochondriasis. Psychol. Med. 32 (5), 843–853. https://doi.org/10.1017/S0033291702005822.

Sangers, T., Nijsten, T., Wakkee, M., 2021. Mobile health skin cancer risk assessment campaign using artificial intelligence on a population-wide scale: a retrospective cohort analysis. J. Eur. Acad. Dermatol. Venereol. 35 (11), e772–e774. https://doi.org/10.1111/jdv.17442.

Sangers, T., Reeder, S., van der Vet, S., Jhingoer, S., Mooyaart, A., Siegel, D.M., Nijsten, T., Wakkee, M., 2022. Validation of a market-approved artificial intelligence mobile health app for skin cancer screening: a prospective multicenter diagnostic accuracy study. Dermatology 238 (4), 649–656. https://doi.org/10.1159/000520474.

Schepman, A., Rodway, P., 2020. Initial validation of the general attitudes towards artificial intelligence scale. Computers in Human Behavior Reports 1, 100014. https://doi.org/10.1016/j.chbr.2020.100014.

Schuessler, J., Freitag, M., 2020. Power analysis for conjoint experiments [Preprint] SocArXiv. https://doi.org/10.31235/osf.io/9yuhp.

Siegel, R.L., Miller, K.D., Wagle, N.S., Jemal, A., 2023. Cancer statistics, 2023. CA A Cancer J. Clin. 73 (1), 17–48. https://doi.org/10.3322/caac.21763.

SkinVision. (n.d.). SkinVision—Skin Cancer Melanoma Detection App. Retrieved 12 January 2024, from https://www.skinvision.com/.

Thompson, H.S., Valdimarsdottir, H.B., Winkel, G., Jandorf, L., Redd, W., 2004. The Group-Based Medical Mistrust Scale: psychometric properties and association with breast cancer screening. Prev. Med. 38 (2), 209–218. https://doi.org/10.1016/j.ypmed.2003.09.041.

Udrea, A., Mitra, G.D., Costea, D., Noels, E.C., Wakkee, M., Siegel, D.M., Carvalho, T.M., Nijsten, T.E.C., 2020. Accuracy of a smartphone application for triage of skin lesions based on machine learning algorithms. J. Eur. Acad. Dermatol. Venereol. 34 (3), 648–655. https://doi.org/10.1111/jdv.15935.

U.S. Department of Commerce. (n.d.). Explore Census Data. Retrieved 5 May 2023, from https://data.census.gov/.