Fine for others but not for me: The role of perspective in patients' perception of artificial intelligence in online medical platforms

Matthias F.C. Hudecek [a,b,*,1], Eva Lermer [c,d,1], Susanne Gaube [c,e], Julia Cecil [c], Silke F. Heiss [f], Falk Batz [g]

[a] Department of Experimental Psychology, University of Regensburg, Regensburg, Germany
[b] Department of Psychology, New York University, New York, United States
[c] Center for Leadership and People Management, Department of Psychology, LMU Munich, Munich, Germany
[d] Department of Business Psychology, Technical University of Applied Sciences Augsburg, Augsburg, Germany
[e] Global Business School for Health, University College London, London, UK
[f] FOM University of Applied Sciences Munich, Munich, Germany
[g] Department of Obstetrics and Gynecology and Center for Gynecological Endocrinology and Reproductive Medicine, University Hospital, LMU Munich, Munich, Germany

ARTICLE INFO

ABSTRACT

In the near future, online medical platforms enabled by artificial intelligence (AI) technology will become increasingly more prevalent, allowing patients to use them directly without having to consult a human doctor. However, there is still little research from the patient's perspective on such AI-enabled tools. We, therefore, conducted a preregistered 2x3 between-subjects experiment ($N = 266$) to examine the influence of *perspective* (oneself vs. average person) and *source of advice* (AI vs. male physician vs. female physician) on the perception of a medical diagnosis and corresponding treatment recommendations. Results of robust ANOVAs showed a statistically significant interaction between the source of advice and perspective for all three dependent variables (i. e., evaluation of the diagnosis, evaluation of the treatment recommendation, and risk perception). People prefer the advice of human doctors to an AI when it comes to their own situation. In contrast, the participants made no differences between the sources of medical advice when it comes to assessing the situation of an average person. Our study contributes to a better understanding of the patient's perspective of modern digital health technology. As our findings suggest the perception of AI-enabled diagnostic tools is more critical when it comes to oneself, future research should examine the relevant factors that influence this perception.

1. Theoretical background

1.1. Introduction

The rapid development of digital technologies is leading to enormous changes in healthcare (Hummelsberger et al., 2023). These include, for example, the growing emergence of online medical platforms that use artificial intelligence (AI) to provide medical advice and consultations directly to patients (Haupt, 2019). These platforms have become increasingly prevalent in recent years as they offer several benefits, such as convenience, accessibility, and cost-effectiveness (Bharti et al., 2020). However, despite the growing popularity of these platforms, research on how people accept and interact with them is still limited (Nadarzynski et al., 2020). The existing literature suggests that there is positive attitude toward AI in healthcare (e.g., Maassen, 2021; Robertson et al., 2023), but also a preference for human physicians over AI technology (e. g., Lennartz et al., 2021; Yakar et al., 2022). So far, it was not investigated whether the generally positive attitude towards AI technologies in healthcare varies in situations where patients themselves are affected by the technology or when it affects another person. The aim of this study was to close this gap. Based on Construal Level Theory (CLT; Trope & Liberman, 2010), we investigated how diagnoses and treatment recommendations of online medical platforms are perceived compared to human doctors if the situation affects a person directly vs. an average person. The current research is important because although AI-enabled digital health technology has become increasingly prevalent in

healthcare, comparatively little is known about how patients view the usage of these tools. Understanding patients' preferences and concerns is crucial for designing AI-enabled tools that meet their needs and expectations.

1.2. Patients' perception of online medical platforms and AI in healthcare

The term online medical platform generally refers to all online health services providing consultation and/or treatment of medical conditions through information technology (Jiang et al., 2021). Online medical platforms, sometimes also referred to as telemedicine, offer a variety of services including online medical consultation, counseling, and health management (e.g., El-Sherif et al., 2022; Liang et al., 2021). With the increasing prevalence of AI technology in healthcare, it is essential to anticipate patients' reactions toward these services as they become a significant aspect of medical care (Richardson et al., 2022). This is especially critical because patients' appreciation regarding AI may hinder the widespread adoption and utilization of these technologies. Research on non-medical uses of AI has demonstrated that the general public's perception of AI can vary widely from worries of loss of control of AI or ethical concerns for AI to hopes for AI in healthcare and education (Fast & Horvitz, 2017). Negative attitudes toward AI are reflected by the so-called algorithm aversion described by Dietvorst et al. (2015). Algorithm aversion as a construct means that humans often favor human advice or forecasters over a statistical algorithm (i.e., AI) in making their decision. As such, people usually even prefer human advice when they have experienced that the algorithm outperforms the human. This tendency may be costly, especially for decision-making situations in which algorithms are better forecasters than humans (e.g., Grove et al., 2000; Kaufmann & Wittmann, 2016; Ægisdóttir et al., 2006). Jussupow et al. (2020) differentiate in their review between three occurrences of aversion: First, people can choose between a human and an algorithm that provides advice or performs a task. Second, people can use the assessment of the human or the algorithm to form their own decision. And third, the human or algorithm providing advice can be evaluated differently. All three types of algorithm aversion seem to be relevant in the context of online medical platforms. For example, a systematic review found that many individuals showed reservations toward the use of AI in healthcare and stated instead a preference for human care (Young et al., 2021). A recent study showed how people respond to risk management recommendations from an AI compared to human experts. Participants were asked to make decisions regarding medical risks, among other things, and it was found that participants generally preferred human experts over AI. In the follow-up study, participants were asked to make a decision before receiving recommendations from human experts or AI. The results showed that participants changed their decisions more readily when receiving dissenting recommendations from human experts compared to the AI (Larkin et al., 2022). These findings underscore the significance of patient involvement to ensure that AI technology is integrated into healthcare in a way that promotes public trust and alleviates potential fears (Richardson et al., 2021). So far, in the medical field, many studies have shown that respondents generally have a positive attitude toward AI. For instance, in one German study, more than half of the participants ($N = 462$ patients) rated the utilization of AI in medicine as positive or very positive, while only a small percentage (4.77%) reported negative or very negative perceptions. Respondents did not express significant concerns regarding AI but strongly agreed that physicians should retain control over AI technology (Fritsch et al., 2022). This ties in with research on different roles of AI. Typically, researchers differentiate between AI as a support system for physicians or health care professionals vs. independently performing diagnoses (Holzinger et al., 2022). Nelson et al. (2020) also found similar results using a qualitative approach ($N = 48$ patients from the UK) focusing on skin cancer screening. While 75% of patients said they would recommend AI to their friends and family, 94% emphasized the significance of a symbiotic relationship between humans and AI

(Nelson et al., 2020), thus advocating AI as a support system instead of an independent actor. Furthermore, a recent review article indicates that patients and the public generally supported the use of AI in healthcare (Young et al., 2021). Contrary to the aforementioned algorithm aversion, these findings might reflect the so-called algorithm appreciation (Logg et al., 2019). Under certain conditions it was shown in a series of experimental studies that participants prefer advice of an algorithm over human advice (Logg et al., 2019; You et al., 2022). A more recent study that built on the results of Logg and colleagues, found that algorithmic aversion and appreciation could be created manipulating the framing of the algorithmic vs. human agent (Hou & Jung, 2021). It is thus interesting to further investigate which factors can explain why people behave in an algorithmic appreciative way in one situation and in an aversive way in another.

In summary, there seems to be a positive attitude toward AI in healthcare, but also a preference for human physicians over AI technology. A question that arises in this context is whether male and female physicians are equally preferred to AI. Numerous studies suggest that female and male physicians are perceived differently (Hall et al., 2011; Mast, 2007). For example, patients talk differently to male vs. female physicians (Hall & Roter, 2002), and rate both the nature of their interactions (Mast & Kadji, 2018) and their competence (Hall et al., 2015) as well as the satisfaction with the physicians (Duberstein et al., 2007) differently. Similar results were also found in a study that asked participants to imagine a virtual medical visit (Mast et al., 2007). Hall et al. (2015) investigated whether patients rated male and female physicians differently based on their patient-centered skills. The results showed that male physicians were judged more positively when they displayed higher patient-centeredness, while female physicians did not receive the same credit for this behavior. The study suggests that female physicians may not be seen as competent when displaying patient-centeredness, as it is expected of them. Therefore, it is interesting not only to examine the difference between human and machine, but also to distinguish between the gender in the human condition.

1.3. Perception as a matter of perspective?

In the context of online medical platforms and AI technology in healthcare, it was not yet investigated whether the generally positive attitude towards AI technologies in healthcare varies in situations where patients themselves are affected by the technology or when it affects another person. Construal Level Theory (CLT) offers a useful theoretical framework for this. CLT posits that the way individuals mentally construct and perceive events and objects is significantly influenced by their perceived psychological distance (PD) from these entities (Trope & Liberman, 2010). PD describes the extent to which a circumstance is "not part of one's direct experience" (Trope et al., 2007, p. 2). Objects that are psychologically distant from the subject are represented abstractly and theoretically. Peng et al. (2013) figuratively compare a high construal level (i.e., high PD) to the perspective of a bird viewing a sprawling forest from above. In contrast, they analogize a low construal level to an animal that closely sees single trees at the bottom of the forest. CLT proposes that the subjective PD to an event or object influences whether one thinks about it more concretely or abstractly (Lermer et al., 2015; 2016b). CLT has been applied in various contexts (e.g., climate change: Brügger, 2020, Wang et al., 2019; smartphone usage while driving: Lim et al., 2021; autonomous shuttle busses: Schandl et al., 2023; attitude toward physical exercise: Wang et al., 2022) and has a wide reach in evaluating risks in different domains (Lermer et al., 2016a). For example, studies on risk assessment have shown, that risk perceptions vary depending on who is at risk. Lermer et al. (2013) found that risks that affect oneself are perceived lower than when they affect other people (e.g., an average citizen). One assumption is that the phenomenon of unrealistic optimism (Harris & Hahn, 2011) is decisive for this effect. According to the motto "it won't happen to me", most people consider the probability that a negative event will occur

lower for oneself than if it affects someone else (Wills, 1981). However, according to the Construal Level Theory (CLT; Trope & Liberman, 2010), risks that affect oneself should be rated higher than risks that affect others, because here the psychological distance (PD) is lower. Consequently, studies on risk perception have shown that more concrete thinking leads to higher risk assessments than abstract thinking (Lermer et al., 2016a). It is unclear whether these research findings can be applied to the context of online medical platforms. Unrealistic optimism would suggest that risk assessment is lower for oneself than for others. Therefore, if the use of AI is expected to involve risk, then this risk should be estimated higher for others than for oneself. However, the opposite assumption would be made according to the CLT. Other relevant variables to capture the attitude toward AI in healthcare represent the evaluation of diagnosis and treatment recommendation. Trust in a diagnosis and the belief that it is correct is crucial for the relationship between the patient and the treating contact and, not least, for the course of treatment (Lu et al., 2018; Nguyen et al., 2009). To the best of the authors' knowledge, there are no studies that investigated the effect of perspective on evaluation of diagnosis and treatment recommendation.

Thus, the current state of research and the gaps identified here lead to the following two research questions (RQ) explored in this study. What influence do (RQ-1) perspective (target person: self vs. average person) and (RQ-2) source of advice (AI vs. male physician vs. female physician) have on the evaluation of the A) diagnosis, B) treatment recommendation, C) risk perception?

Overall, this study aims to contribute to a better understanding of patients' perspectives on modern digital health technology by identifying relevant factors that influence their perceptions.

2. Materials and method

2.1. Design

The experimental design and the research questions were preregistered on Open Science Framework (https://osf.io/6kzxf; this is a temporary and anonymous link for the peer-review process. The repository will be made public upon acceptance). We conducted a 2 (perspective) x 3 (source of advice) between-subjects online experiment. Participants were randomly assigned to one of the six conditions. Perspective as the first independent variable had two levels. Accordingly, participants were asked to imagine

[condition self]: *that you have recently been vaccinated and now have severe discomfort that also causes concern, and you are wondering if this is still normal. Therefore, you go to an online medical platform and seek medical advice. There you describe all your symptoms and get then assigned to a medical contact.*

[condition average person]: *that an average German citizen has recently been vaccinated and now has severe discomfort that also causes concern, and wonders if this is still normal. Therefore, the person goes to an online medical platform and seeks medical advice. There the person describes all his or her symptoms and gets then assigned to a medical contact.*

Source of advice, as the second independent variable had three levels: AI, male physician, and female physician. As the scenario continued, participants had to imagine the following:

[condition self]: *On the online medical platform, you will be assigned to a medical contact.*

[condition AI]: *You have been assigned to the Symptom-Checkbot. Symptom-Checkbot is an artificial intelligence that has been trained to make medical diagnoses that exactly match your health problem.*

[condition male/female physician]: *You have been assigned to Dr. Andreas Weber/Dr. Andrea Weber who is a specialist in medical diagnoses that exactly match your health problem.*

[condition average person]:

On the online medical platform, the average person will be assigned to a medical contact.

[condition AI]: *An average person has been assigned to the Symptom-Checkbot. Symptom-Checkbot is an artificial intelligence that has been trained to make medical diagnoses that exactly match the health problem of the average person.*

[condition male/female physician]: *The average person has been assigned to Dr. Andreas Weber/Dr. Andrea Weber who is a specialist in medical diagnoses that exactly match the health problem of the average person.*

At the end of the scenario, all participants received the same diagnosis and treatment recommendation:

[AI vs. female vs. male physician] *gives the following diagnosis: These are normal side effects of vaccination.*

[AI vs. female vs. male physician] *gives the following treatment recommendation: Do nothing and do not worry.*

In addition to the text of the scenario, source of advice was accompanied by graphic illustrations. For the male/female physician condition, participants were shown a figure with a stethoscope that looked either male or female. In the AI condition, an abstract image of a human head with computer neurons was shown. We decided to use the vaccination scenario as a cover story since it was a very present and common topic at the time of data collection due to vaccinations against COVID-19. In addition, we were looking for a description of symptoms that sounded severe enough (i.e., "severe discomfort") to make people believe it would make sense to seek medical advice without at the same time stressing the participants in an unethical way or giving them the impression that they needed to see an emergency doctor immediately. All study materials can be found in the online repository (https://osf.io/vwxcj/). Following the scenario, participants were asked to rate the dependent variables (i.e., evaluation of the diagnosis, treatment recommendation, and risk perception), and complete a survey assessing psychometric scales and demographics.

2.2. Measures

To examine the three dependent variables (i.e., evaluation of the diagnosis, treatment recommendation, and risk perception), we developed eight items. Four items were used to measure the evaluation of the diagnosis. This variable captured both participants' trust in the diagnosis (e.g., "I trust the diagnosis") as well as their trust in the medical expertise of the source of advice (e.g., "I trust the medical expertise of [condition]"; condition refers to female/male doctor or Symptom-Checkbot) on a 7-point Likert scale from "strongly disagree" to "strongly agree". These items were averaged to an index of evaluation of diagnosis ($\alpha = 0.89$). Two items examined the evaluation of treatment recommendation. This variable measured the extent to which individuals felt comforted by the recommendation and whether they would adhere to it (e.g., "I feel comforted by the treatment recommendation"). Both items were measured on a 7-point Likert scale from "strongly disagree" to "strongly agree" and were averaged and resulted in an index of evaluation of treatment recommendation ($\alpha = 0.80$). Finally, two items were used to measure risk perception. This variable measured the perceived risk that the diagnosis and the treatment recommendation could be wrong (e.g., "How high do you estimate the risk that the diagnosis could be wrong?") on a scale from 0% to 100%. The average of these two items showed excellent internal consistency α of 0.92.

In addition, the following control variables were considered: attitude towards artificial intelligence, health anxiety, technology commitment, and socio-demographic variables (gender, age, social status, and health insurance status).

Attitudes toward AI was measured with the General Attitudes towards Artificial Intelligence Scale (GAAIS; Schepman & Rodway, 2020). Participants had to answer 20 items (e.g., "Artificially intelligent systems can help people feel happier") on a 5-point Likert scale from "strongly disagree" to "strongly agree". Items were averaged to two subscales (positive attitude: $\alpha = 0.67$ and negative attitude: $\alpha = 0.43$). Also, an attention check item was part of the scale ("I would be grateful if you could select agree"). Participants who failed to answer the attention check correctly were excluded from the analysis.

Health anxiety was examined using the health anxiety inventory (MK-HAI) developed by Bailer and Witthöft (2006). This scale assesses a person's tendency for health-related concerns with 14 items (e.g., "I spend a lot of time worrying about my health") on a five-point Likert scale from "strongly disagree" to "strongly agree". All items were averaged to an index of health anxiety ($\alpha = 0.92$).

Technology commitment was measured with the Technology Commitment Short Scale (Neyer et al., 2016). This construct captures the personal attitude toward and handling of modern technology. Participants had to answer 12 items (e.g., "I am very curious regarding new technical developments") on a 5-point Likert scale from "not true at all" to "completely true". Items were averaged to calculate technology commitment ($\alpha = 0.82$).

The of MacArthur Scale of Subjective Social Status (Adler et al., 2000; Hudecek et al., 2022) was used to assess subjective social status. Participants placed themselves on a given drawing of a ladder with ten rungs according to the following description: "Think of this ladder as representing where people stand in our society. At the top of the ladder are the people who are the best off, those who have the most money, most education, and best jobs. At the bottom are the people who are the worst off, those who have the least money, least education, and worst jobs or no job".

We measured a set of socio-demographic variables. Participants were asked to report their age, gender, formal education in terms of German educational achievement levels (ranging from no educational attainment at all to a university degree), and their average working hours per month. We assessed participants' current health condition and their attitude toward vaccinations as well as their vaccination status against COVID-19. We also asked participants to provide their height and weight to calculate their Body mass index (BMI). Moreover, we assessed participants' health insurance status (public vs. private health insurance).

A manipulation check was included at the end of the survey. For this, participants had to indicate from whom they had previously received the diagnosis and treatment recommendation on the online medical platform as part of the scenario. Individuals who failed to answer the manipulation check correctly were excluded from the analysis.

2.3. Sample

The data were collected using pools from three universities across Germany. Participants were informed about the purpose of the study and gave consent prior to participation. The consent also included that the data would be used for publication and stored as open data in the Open Science Framework. Participants received course credit for participation. The online questionnaire was set to "require" answers from participants apart from the demographic variables also including the COVID-19 vaccination status. Thus, there are no missing values on the psychological measures in the dataset.

In total, 350 participants completed the survey. Before running the analyses, we excluded certain participants from the dataset according to predefined criteria that were part of the preregistration. The questionnaire was programmed that only those participants who answered the attention check item correctly could complete the questionnaire. However, all subjects who failed to answer the manipulation check correctly were excluded ($N = 69$). Moreover, as stated in our preregistration, participants who had negative attitudes toward vaccination and were not vaccinated against COVID-19 were excluded from the analyses ($N = 15$). The reason for doing so is that the experimental scenarios required the participants to imagine that they had recently been vaccinated. The final sample comprised $N = 266$ persons. Participants were 74% female students with an average age of $M = 24.83$ years ($SD_{age} = 5.81$, $Range_{age} = 18–51$). Regarding the educational level, 9% had completed vocational training, 86% had a university of applied sciences entrance qualification (i.e., German Fachhochschulreife, $N = 40$) or a university entrance qualification, i.e., German Abitur $N = 190$), and 5% of the respondents already had an academic degree. The majority of the participants worked at least part-time with an average of over 25 h per week ($SD = 14.66$), which clearly distinguishes these participants (63% of the total sample) from classic student samples.

2.4. Statistical analyses

Data were analyzed using R Studio (version 2022.07.1 on macOS, R version 4.2.1). We used two-way ANOVAs to analyze the effect of perspective and source of advice on the dependent variables. We checked the statistical assumptions using the Shapiro-Wilk test (normality assumption) and the Levene test (homoscedasticity). As the assumption of normal distribution was violated for all dependent variables and across all conditions and homoscedasticity could only be fully established for treatment recommendation, we calculated robust ANOVAs using the WRS2 package by Mair and Wilcox (2020). The calculation of robust ANOVAs is based on 20% trimmed means. As Mair and Wilcox (2020) describe, the "appeal of a 20% trimmed mean is that it achieves nearly the same amount of power as the mean when sampling from a normal distribution" (p. 1). It is important to note that the robust two-way ANOVA does not report any degrees of freedom since an adjusted critical value is used to determine significance, instead of the F-value, a Q statistic is reported (Mair & Wilcox, 2020). After running the ANOVA, we calculated robust two-sided pairwise comparisons to analyze the simple main and interaction effects.

To analyze the effect of the control variables, we computed multiple linear regressions for each dependent variable with the independent and the control variables as predictors. As Cook's barplots indicated several outliers for the dependent variables, we calculated robust regression analyses using the robustbase package (Maechler, 2022).

Table 1
Means and standard deviations of the dependent variables by condition.

Perspective	Source of diagnosis	N	EOD	EOT	Risk perception
Self	AI	45	3.61 (1.44)	3.84 (1.51)	50.53 (19.99)
	Female physician	45	4.80 (1.42)	5.17 (1.43)	28.66 (19.05)
	Male physician	43	4.92 (1.51)	5.34 (1.56)	32.38 (22.23)
Average person	AI	47	4.10 (1.17)	4.43 (1.31)	43.28 (19.87)
	Female physician	47	4.27 (1.36)	4.61 (1.55)	42.60 (21.37)
	Male physician	39	4.28 (1.23)	4.56 (1.48)	41.44 (26.80)

Note. EOD = evaluation of diagnosis; EOT = evaluation of treatment recommendation; Standard deviations are shown in brackets.

3. Results

Table 1 displays means and standard deviations for the dependent variables for each condition (AI vs. female physician vs. male physician, self vs. average person).

3.1. Evaluation of diagnosis

To answer research questions RQ-1-A and RQ-2-A, exploring the influence of perspective (self vs. average person) and source of advice (AI vs. male physician vs. female physician) on the evaluation of the diagnosis, we performed a robust two-way ANOVA. Results show a significant main effect of source of diagnosis on the evaluation of the diagnosis ($Q = 15.54$, $p = .001$). Pairwise comparisons regarding the main effect of source of diagnosis show that the evaluation of diagnoses was significantly lower for diagnoses coming from AI compared to male ($\widehat{\psi} = -1.74$, $p = .001$) or female physicians ($\widehat{\psi} = -1.46$, $p = .001$). The evaluation of diagnosis between male and female physicians was not significantly different ($\widehat{\psi} = -0.28$, $p = .584$). No significant main effect was found for perspective ($Q = 2.11$, $p = .149$). In addition, there was a statistically significant interaction between source of diagnosis and perspective ($Q = 10.03$, $p = .009$). Post hoc tests revealed that the effect of source of advice was dependent on perspective (see Fig. 1a). Specifically, participants' evaluation of the diagnosis from AI was significantly lower compared to diagnoses from male ($\widehat{\psi} = -1.32$, $p = .012$) or female physicians ($\widehat{\psi} = -1.24$, $p = .006$) when it came to their own perspective (condition self). In contrast, no differences were found between the different sources of advice for the perspective of an average person ($\widehat{\psi} = -0.08$, $p = .880$).

3.2. Evaluation of treatment recommendation

Turning to the analysis of research questions RQ-1-B and RQ-2-B regarding the evaluation of treatment recommendation, the robust two-way ANOVA again showed a significant main effect of source of advice on evaluation of treatment recommendation ($Q = 19.30$, $p = .001$) but not for perspective ($Q = 3.16$, $p = .078$). As before, the evaluation of treatment recommendation was significantly lower for recommendations coming from AI compared to male ($\widehat{\psi} = -2.08$, $p < .001$) or female physicians ($\widehat{\psi} = -1.81$, $p < .001$). The overall evaluation of the treatment recommendation between male and female physicians was not significantly different ($\widehat{\psi} = -0.28$, $p = .625$). As for the evaluation of the diagnosis, the results showed that a significant interaction
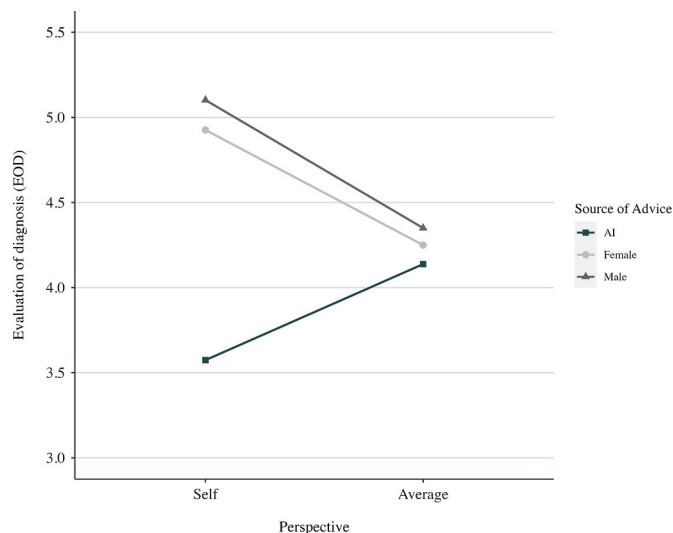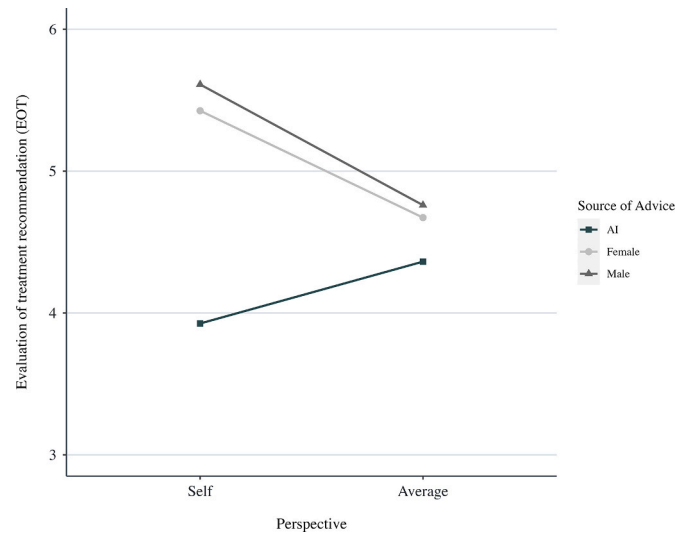

Fig. 1b. Interaction plot for evaluation of treatment recommendation (EOT). *Note.* Figure shows trimmed means (20%).

effect exists between source of advice and perspective in predicting the evaluation of treatment recommendation ($Q = 7.93$, $p = .023$). Post hoc tests revealed that the effect of source of advice was again dependent on perspective (see Fig. 1b). Specifically, participants' evaluation of treatment recommendation from AI was significantly lower compared to diagnoses from male ($\widehat{\psi} = -1.29$, $p = .029$) or female physicians ($\widehat{\psi} = -1.19$, $p = .012$) when it came to their own perspective (condition self). In contrast, no differences were found between the different sources of advice for the perspective of an average person ($\widehat{\psi} = -0.10$, $p = .861$).

3.3. Risk perception

Regarding research questions RQ-1-C and RQ-2-C with risk perception as the dependent variable, the robust two-way ANOVA revealed again a significant main effect for source of advice ($Q = 14.44$, $p = .002$) and a significant interaction effect ($Q = 11.82$, $p = .004$). As before, no significant main effect was found for perspective ($Q = 3.43$, $p = .067$). Again, we found a similar pattern as for the previous research questions. Risk perception was significantly higher for AI compared to male ($\widehat{\psi} = 23.78$, $p = .007$) or female physicians ($\widehat{\psi} = 25.30$, $p = .001$), while the

Fig. 1a. Interaction plot for evaluation of diagnosis (EOD). *Note. Figure* shows trimmed means (20%).
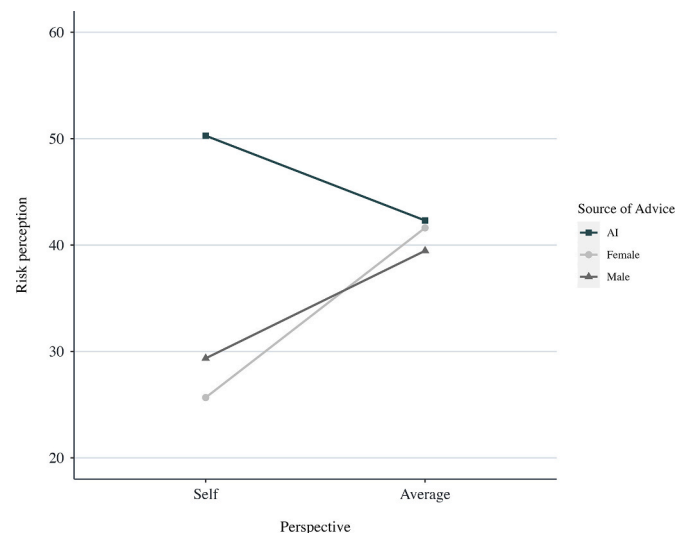

Fig. 1c. Interaction plot for risk perception. *Note.* Figure shows trimmed means (20%).

Table 2
Means, standard deviations, and correlations with confidence intervals for the control and dependent variables.

Variable	M	SD	1	2	3	4	5	6	7	8
1. EOD	4.32	1.42								
2. EOT	4.65	1.54	.77** [.71, .81]							
3. Risk perception	39.88	22.57	−.67** [−.73, −.59]	−.57** [−.64, −.48]						
4. GAAIS positive	3.61	0.54	.03 [−.09, .15]	−.02 [−.14, .10]	−.03 [−.15, .09]					
5. GAAIS negative	3.23	0.68	.01 [−.11, .13]	−.03 [−.15, .09]	−.01 [−.13, .11]	.52** [.43, .60]				
6. Technology Commitment	3.82	0.52	.04 [−.08, .16]	−.02 [−.14, .10]	−.02 [−.14, .10]	.52** [.43, .61]	.46** [.36, .55]			
7. Health anxiety	2.39	0.75	.06 [−.06, .18]	.13* [.01, .25]	−.11 [−.23, .01]	−.10 [−.22, .02]	−.24** [−.35, −.13]	−.25** [−.36, −.13]		
8. Social status	6.21	1.18	.00 [−.12, .12]	.04 [−.08, .16]	.08 [−.04, .20]	.10 [−.02, .22]	.20** [.08, .31]	.27** [.15, .38]	−.11 [−.23, .01]	
9. Age	24.83	5.41	.01 [−.11, .13]	−.02 [−.14, .10]	−.02 [−.14, .10]	−.06 [−.18, .06]	−.05 [−.17, .07]	.12* [.00, .24]	−.09 [−.21, .03]	.19** [.07, .30]

Note. EOD = evaluation of diagnosis, EOT = evaluation of treatment recommendation; *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates $p < .05$. ** indicates $p < .01$.

Table 3a
Robust multiple regression analysis on evaluation of diagnosis (EOD).

Variable	B	SE	t	p	95% CI
(Intercept)	3.90	1.35	2.88	.004	[1.2, 6.6]
Distance[a]	−0.57	0.31	−1.86	.063	[-1.2, 0]
Perspective[b]	0.19	0.28	0.67	.503	[-0.4, 0.7]
Perspective[c]	0.10	0.28	0.37	.714	[-0.4, 0.7]
GAAIS positive	−0.01	0.22	−0.05	.958	[-0.4, 0.4]
GAAIS negative	0.15	0.20	0.74	.460	[-0.2, 0.5]
Technology commitment	0.09	0.23	0.40	.689	[-0.4, 0.5]
Health anxiety	0.14	0.12	1.12	.262	[-0.1, 0.4]
BMI	−0.04	0.03	−1.66	.098	[-0.1, 0]
Gender	0.10	0.28	0.35	.727	[-0.4, 0.6]
Social Status	−0.01	0.09	−0.11	.916	[-0.2, 0.2]
Age	0.01	0.01	0.91	.361	[0, 0]
Health insurance[d]	−0.29	0.33	−0.87	.384	[-0.9, 0.4]
Distance x Perspective[e]	1.10	0.43	2.52	.012	[0.2, 1.9]
Distance x Perspective[f]	1.37	0.47	2.92	.004	[0.5, 2.3]

Note. CI = confidence interval.
[a] Condition self = 1, condition average person = 0.
[b] Condition female = 1, AI = 0.
[c] Condition male = 1, condition AI = 0.
[d] Public health insurance = 1, private health insurance = 0.
[e] Condition self = 1, condition average person = 0, Condition female = 1, AI = 0.
[f] Condition self = 1, condition average person = 0, Condition male = 1, AI = 0; GAAIS = general attitude towards artificial intelligence scale.

risk perception between male and female physicians was not significantly different ($\widehat{\psi} = -1.53$, $p = .853$). In addition, post hoc tests regarding the significant interaction revealed that the effect of source of advice was dependent on perspective (see Fig. 1c). Specifically, participants' risk perception was significantly higher for AI compared to male ($\widehat{\psi} = 18.08$, $p = .037$) or female physicians ($\widehat{\psi} = 23.92$, $p = .001$) when it came to their own perspective (condition self). In contrast, no differences were found between the different sources of advice for the perspective of an average person ($\widehat{\psi} = -5.84$, $p = .481$).

3.4. Control variables

Table 2 displays the correlations between the control variables with the dependent variables. Robust multiple regression analyses for the dependent variables revealed effects for some of the control variables (see Table 3a–c). Regarding the evaluation of the diagnosis, no effects were found for any of the control variables. Turning to the evaluation of the treatment recommendation, we found a significant positive effect for health anxiety ($b = 0.35$, $SE = 0.15$, $p = .020$). All other control variables had no significant effect. Regarding risk perception as the third dependent variable, health anxiety ($b = -4.52$, $SE = 2.16$, $p = .037$) again turned out to be a significant control variable. Also, health insurance status ($b = 10.36$, $SE = 4.75$, $p = .026$) had a significant impact on risk perception, suggesting that participants with public health insurance perceived higher risk compared to participants with private health insurance.

4. Discussion

Digital health technology using AI to interact with patients will become increasingly important in the near future (Madhav & Tyagi, 2022). However, there is limited research on patients' perspectives, especially when it comes to experimental study designs. Consequently, we conducted a scenario-based experiment to examine the influence of perspective (self vs. average person) and source of advice (AI vs. male physician vs. female physician) on the perception of a medical diagnosis, corresponding treatment recommendations and risk perception in the context of online medical platforms. Overall, our study provides three key findings:

Table 3b
Robust multiple regression analysis on evaluation of treatment recommendation (EOT).

Variable	B	SE	t	p	95% CI
(Intercept)	4.36	1.38	3.17	.002	[1.7, 7.1]
Distance[a]	−0.67	0.36	−1.88	.062	[-1.4, 0]
Perspective[b]	0.28	0.34	0.83	.407	[-0.4, 0.9]
Perspective[c]	0.19	0.35	0.56	.579	[-0.5, 0.9]
GAAIS positive	−0.23	0.23	−0.99	.322	[-0.7, 0.2]
GAAIS negative	0.24	0.23	1.06	.289	[-0.2, 0.7]
Technology commitment	−0.02	0.26	−0.07	.942	[-0.5, 0.5]
Health anxiety	0.35	0.15	2.34	.020	[0.1, 0.6]
BMI	−0.04	0.02	−1.51	.133	[-0.1, 0]
Gender	0.31	0.25	1.24	.216	[-0.2, 0.8]
Social Status	0.07	0.10	0.71	.476	[-0.1, 0.3]
Age	0.00	0.02	0.02	.987	[0, 0]
Health insurance[d]	−0.65	0.36	−1.81	.072	[-1.4, 0.1]
Distance x Perspective[e]	1.23	0.46	2.65	.009	[0.3, 2.1]
Distance x Perspective[f]	1.63	0.50	3.24	.001	[0.6, 2.6]

Note. CI = confidence interval.
[a] Condition self = 1, condition average person = 0.
[b] Condition female = 1, AI = 0.
[c] Condition male = 1, condition AI = 0.
[d] Public health insurance = 1, private health insurance = 0.
[e] Condition self = 1, condition average person = 0, Condition female = 1, AI = 0.
[f] Condition self = 1, condition average person = 0, Condition male = 1, AI = 0;
GAAIS = general attitude towards artificial intelligence scale.

Table 3c
Robust multiple regression analysis on risk perception.

Variable	B	SE	t	p	95% CI
(Intercept)	53.77	19.89	2.70	.007	[14.8, 92.8]
Distance[a]	7.39	4.61	1.60	.110	[-1.6, 16.4]
Perspective[b]	−0.77	4.95	−0.16	.877	[-10.5, 8.9]
Perspective[c]	−3.50	6.52	−0.54	.592	[-16.3, 9.3]
GAAIS positive	1.18	3.72	0.32	.751	[-6.1, 8.5]
GAAIS negative	−2.23	2.83	−0.79	.430	[-7.8, 3.3]
Technology commitment	−4.24	4.56	−0.93	.353	[-13.2, 4.7]
Health anxiety	−4.52	2.16	−2.10	.037	[-8.8, −0.3]
BMI	0.26	0.44	0.58	.560	[-0.6, 1.1]
Gender	0.98	4.48	0.22	.826	[-7.8, 9.8]
Social Status	2.31	1.39	1.66	.098	[-1.2, 0.3]
Age	−0.47	0.39	−1.19	.234	[-1.2, 0.3]
Health insurance[d]	10.63	4.75	2.24	.026	[1.3, 19.9]
Distance x Perspective[e]	−22.21	6.73	−3.30	.001	[-35.4, −9.0]
Distance x Perspective[f]	−16.17	7.78	−2.08	.039	[-31.4, −0.9]

Note. CI = confidence interval.
[a] Condition self = 1, condition average person = 0.
[b] Condition female = 1, AI = 0.
[c] Condition male = 1, condition AI = 0.
[d] Public health insurance = 1, private health insurance = 0.
[e] Condition self = 1, condition average person = 0, Condition female = 1, AI = 0.
[f] Condition self = 1, condition average person = 0, Condition male = 1, AI = 0;
GAAIS = general attitude towards artificial intelligence scale.

First, the results show factors that influence the patient's perception of AI in the context of online medical platforms. In general, diagnoses, treatment recommendations, and risk perception of AI compared to human physicians are perceived more negatively. This reflects previous research that has shown that algorithm aversion (Dietvorst et al., 2015) not only applies to medical professionals (Gaube et al., 2021) but is also common among patients (Wu et al., 2021). Specifically, the significant interaction effects indicate that respondents rate all three dependent variables worse for the AI if they are affected by its diagnosis and treatment recommendation themselves. However, when it comes to judging the situation for other people, AI is not perceived more negatively than human physicians, i.e., there was no difference between the different perspectives. The results thus tie in with existing research on

Construal Level Theory (CLT; Trope & Liberman, 2010). CLT research has shown, that the evaluation of risk in various domains (Lermer et al., 2016a), smartphone usage while driving (Lim et al., 2021), or attitude toward physical exercise (Wang et al., 2022) depend on whether a particular event or situation is more distal or proximal. In the case of more concrete thinking, risks are perceived higher (Lermer et al., 2015), and the likelihood of engaging in, for example, pro-environmental behaviors increase (Wang et al., 2020). In addition, the present study results also expand existing research. Previous studies have demonstrated both positive (Nelson et al., 2020) and negative (Promberger & Baron, 2006) perceptions of AI from a patient's perspective. Here, our study provides a possible explanation for these different perceptions based on the CLT. Accordingly, AI is perceived more negatively when a person is directly affected, while it is perceived more positively when others are affected. This potential mechanism also fits the explanation offered by Longoni and colleagues (Longoni et al., 2019). They suggest that patients' negative attitudes toward AI in healthcare are due to what they call uniqueness neglect. This describes the concern that one's unique characteristics, circumstances, and symptoms will be neglected when being cared for by AI-based healthcare providers.

Second, we found that patients did not evaluate male and female physicians differently. According to this, the results of the current study oppose gender bias from the patient's perspective, and they indicate that it does not matter to patients whether they are advised by a male or female doctor on an online medical platform. Previous research identifying gender biases on the patient side has mostly referred to the face-to-face interaction between patient and physician (e.g., Mast & Kadji, 2018). However, due to the setting of an online medical platform, this kind of interaction was not given in our study; therefore, it seems plausible that no differences between female and male physicians were identified. One potential technical advancement of online medical platforms is the use of chatbots (Nadarzynski et al., 2020). However, in this case, the design and response behavior of the chatbot should ensure that it communicates empathically. Previous studies have indicated that chatbots' ability to show empathy is important for patients accepting such technologies (Liu & Sundar, 2018). Future studies could therefore investigate what type of patient-AI interaction and communication on online medical platforms improves patient acceptance and trust in diagnoses and treatment recommendations.

Third, there was little influence of possible moderator variables. Of the control variables that were included in this study, only health anxiety showed a significant statistical effect on treatment recommendation ratings and risk perception. This finding ties into existing research, as health anxiety is linked to safety-seeking behavior (Helbig-Lang & Petermann, 2010; Lermer et al., 2021) and higher levels of risk perception (Lindner et al., 2022; Mohd Salleh Sahimi et al., 2021). Accordingly, people with higher health anxiety appear to have slightly higher confidence in treatment recommendations per se. However, their overall risk perception is lower. Here, subsequent analyses comparing participants in the human vs. AI conditions show that this association can only be sustained regarding the risk perception of diagnoses and treatment recommendation coming from AI. This should be considered when implementing appropriate services. Otherwise, individuals with high health anxiety might not follow the recommendations, resulting in actual health risks. In addition, a significant association between insurance status and risk perception was found. In our study sample, the proportion of individuals with private health insurance was small (11%). Still, these individuals had higher trust in the recommendations of physicians and AI on the online medical platform in general. Future research should further investigate whether this tendency can be replicated with a larger sample of participants with private health insurance. If such an effect should actually exists, this would result in important consequences for the design of corresponding services. In particular, people with public health insurance should then be addressed in such a way that they are also convinced by the diagnoses they receive on an online platform and can follow the recommendations. Remarkably,

there were no significant associations with other potential moderator variables (e.g., general attitude toward artificial intelligence). This suggests that the degree to which a person perceives and evaluates diagnoses and recommendations on an online medical platform is primarily related to whether they come from AI or a human and whether or not the person is personally affected. Further research could follow up on this with a comprehensive study to determine whether and if so, which person-related characteristics influence the perception and acceptance of AI in a medical context from the patient's perspective. For example, a study on the identification of different user groups through the application of latent profile analyses (Spurk et al., 2020) could be a promising starting point.

4.1. Limitations

Some limitations of the study need to be mentioned. First, we used a scenario-based design which might affect the validity of the results. Although it is known from different contexts of psychological research that the results of scenario-based experiments are comparable to real in-field experiments and show similar validity (Chen et al., 2019; Weyrich et al., 2020; Zhang et al., 2023), future studies should replicate the design in a more realistic setting. Second, the psychometric quality of one moderator variable capturing participants' general attitude towards AI was not adequate, as the value of Cronbach's alpha for negative GAAIS was significantly below the acceptable threshold of .70 (Scher-melleh-Engel & Werner, 2012). We have no plausible explanation for this low reliability since this instrument is a validated measure that has already shown good reliability values in other studies (e.g., Carolus et al., 2023). In the present case, the validity of the results might therefore be limited with regard to this variable. Future studies should therefore replicate the results on the basis of the current instruments and additionally use alternative measures to capture participants' general attitude toward AI (e.g., the ATAI scale of Sindermann et al., 2021) to check for potential differences. Third, as in many psychological studies, the sample is biased towards female participants and has a low mean age of 24.83 years. However, controlling for gender and age had no effects on the dependent variables. In addition, it must be noted that the sample was skewed towards a higher education level, which might have impacted the participants' perception of artificial intelligence. Future studies should replicate the current findings and aim for a sample with a more heterogeneous educational background. In general, a higher education level is associated with higher acceptance rates of new technologies (Czaja et al., 2006; Rice et al., 2019). However, a recent study with a sample from Israel found that education level was not a significant predictor for patients' acceptance of AI-based technology in primary care (Chalutz Ben-Gal, 2023). Another aspect that future studies should address refers to the design of the scenario. The symptoms in the current study were framed as "severe discomfort". It seems likely that the results can be generalized to cases or diseases where a severe discomfort is typical of the course of the illness (e.g., having a cold, flu). However, for other diseases or cases (e.g., emergency situations), the study should be replicated to test the generalizability of the results. In addition, the current design could also be used to test the effect for existing AI-assisted applications and symptom checkers (e.g., screen cancer screening).

4.2. Practical implications

The results of our study suggest some practical implications. First, considering that diagnoses, treatment recommendations, and risk perception of AI compared to human physicians are perceived more negatively in general, the main point of contact for patients should remain a real human. Therefore, online medical platforms providing the possibility for patients to check symptoms should not only be developed in a way that they are used by the patients directly. Instead, such symptom checkers could be developed for the use of physicians or healthcare organizations. Thus, humans will not be replaced by AI, but AI will be able to support physicians or healthcare organizations while patients are still in contact with real humans. When AI applications in healthcare become more widespread and accepted in the future, symptom checkers can also be increasingly developed for direct use by patients. Second, our results stress that when it comes to judging the situation for other people, AI is no longer perceived more negatively than human physicians. This effect could provide some opportunities for the design of online medical platforms. When patients use these services and interact directly with them, the symptom checker could refer to others. For example, when recommending actions or providing explanations, it could not only address the user directly, but also emphasize that "other people" with the same symptoms have benefited from certain measures or behaviors. In this way, the psychological distance could be increased, which in turn – according to the results of the current study – should have a positive effect on the perception of the diagnoses and recommendations. However, it must be noted that these considerations still need to be tested in future studies. Third, we found that patients with higher health anxiety tend to perceive greater risks of incorrect diagnoses and treatment recommendations. Therefore, we suggest in line with precision medicine and personalized health care (e.g., Johnson et al., 2021), developing AI-based symptom checkers that can recognize signs of health anxiety in patient interactions (e.g., through natural language processing of patient queries) to respond appropriately. For example, once the symptom checker has identified patients with increased health anxiety, it could provide additional information to alleviate the health anxiety or involve a human doctor or healthcare professional in the conversation.

4.3. Conclusions

In our preregistered scenario-based 2x3 experiment, we examined the influence of perspective (self vs. average person) and source of advice (AI vs. male physician vs. female physician) on the perception of a medical diagnosis and corresponding treatment recommendations and risk perception. In addition, we assessed the importance of several control variables such as health anxiety, general attitude towards artificial intelligence, and socio-demographic variables (e.g., social status). Results show that people prefer the advice of human doctors compared to an AI when it comes to their own situation. In contrast, there are no differences in terms of the source of advice for judging the situation of an average person. Our study contributes to a better understanding of the patient's perspective of modern digital health tools. As results indicate that the perception of AI-enabled tools is more critical when it comes to yourself, future research should examine the relevant factors that influence this perception.

Data availability statement

The data described in this article are openly available in the Open Science Framework at https://osf.io/vwxcj/. Please note that this is a provisional and anonymous link for the peer review process. In case of acceptance, we will include the link to the public repository.

Ethical approval and informed consent statement

Our study was conducted in full compliance with the ethical guidelines of the German Psychological Society (DGPs) and the American Psychological Association (APA). At the time of data collection, it was not common practice at most German universities to obtain ethical approval for studies that did not contain sensitive personal data, did not

involve vulnerable groups, and did not pose risks to participants. Therefore, in accordance with the ethical guidelines of the Department of Psychology at LMU Munich, ethical approval was not required for this study. Only anonymous questionnaires were used for the study. No identifying information was obtained from participants. All participants were fully informed about the study and gave informed consent by agreeing to participate. In doing so, they were explicitly informed that all data would be kept confidential and that they could withdraw from the study at any time without giving a reason.

CRediT authorship contribution statement

Matthias F.C. Hudecek: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Eva Lermer:** Conceptualization, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing, Supervision. **Susanne Gaube:** Conceptualization, Methodology, Writing – review & editing. **Julia Cecil:** Writing – review & editing. **Silke F. Heiss:** Methodology, Writing – review & editing. **Falk Batz:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

There are no acknowledgements to be made.

References

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology, 19*(6), 586–592. https://doi.org/10.1037/0278-6133.19.6.586

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34*(3), 341–382. https://doi.org/10.1177/0011000005285875

Bailer, J., & Witthöft, M. (2006). *Deutsches modifiziertes health anxiety inventory (MK-HAI). Zusammenstellung sozialwissenschaftlicher Items und skalen (ZIS)*. https://doi.org/10.6102/ZIS71

Bharti, U., Bajaj, D., Batra, H., Lalit, S., Lalit, S., & Gangwani, A. (2020). *Medbot: Conversational artificial intelligence powered chatbot for delivering tele-health after COVID-19. 2020 5th international Conference on Communication and electronics systems (ICCES)*, 870–875. https://doi.org/10.1109/ICCES48766.2020.9137944

Brügger, A. (2020). Understanding the psychological distance of climate change: The limitations of construal level theory and suggestions for alternative theoretical perspectives. *Global Environmental Change, 60*, Article 102023. https://doi.org/10.1016/j.gloenvcha.2019.102023

Carolus, A., Koch, M., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). *MAILS – meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies*. https://doi.org/10.48550/ARXIV.2302.09319

Chalutz Ben-Gal, H. (2023). Artificial intelligence (AI) acceptance in primary care during the coronavirus pandemic: What is the role of patients' gender, age and health awareness? A two-phase pilot study. *Frontiers in Public Health, 10*, Article 931225. https://doi.org/10.3389/fpubh.2022.931225

Chen, G., Smith, T. A., Kirkman, B. L., Zhang, P., Lemoine, G. J., & Farh, J.-L. (2019). Multiple team membership and empowerment spillover effects: Can empowerment processes cross team boundaries? *Journal of Applied Psychology, 104*(3), 321–340. https://doi.org/10.1037/apl0000336

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement (create). *Psychology and Aging, 21*(2), 333–352. https://doi.org/10.1037/0882-7974.21.2.333

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033

Duberstein, P., Meldrum, S., Fiscella, K., Shields, C. G., & Epstein, R. M. (2007). Influences on patients' ratings of physicians: Physicians demographics and

personality. *Patient Education and Counseling, 65*(2), 270–274. https://doi.org/10.1016/j.pec.2006.09.007

El-Sherif, D. M., Abouzid, M., Elzarif, M. T., Ahmed, A. A., Albakri, A., & Alshehri, M. M. (2022). Telehealth and artificial intelligence insights into healthcare during the COVID-19 pandemic. *Healthcare, 10*(2), 385. https://doi.org/10.3390/healthcare10020385

Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence, 31*(1). https://doi.org/10.1609/aaai.v31i1.10635

Fritsch, S. J., Blankenheim, A., Wahl, A., Hetfeld, P., Maassen, O., Deffge, S., Kunze, J., Rossaint, R., Riedel, M., Marx, G., & Bickenbach, J. (2022). Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. *Digital Health, 8*. https://doi.org/10.1177/20552076221116772

Gaube, S., Suresh, H., Raue, M., Merritt, A., Berkowitz, S. J., Lermer, E., Coughlin, J. F., Guttag, J. V., Colak, E., & Ghassemi, M. (2021). Do as AI say: Susceptibility in deployment of clinical decision-aids. *NPJ Digital Medicine, 4*(1), 31. https://doi.org/10.1038/s41746-021-00385-9

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19–30. https://doi.org/10.1037/1040-3590.12.1.19

Hall, J. A., Blanch-Hartigan, D., & Roter, D. L. (2011). Patients' satisfaction with male versus female physicians: A meta-analysis. *Medical Care, 49*(7), 611–617.

Hall, J. A., & Roter, D. L. (2002). Do patients talk differently to male and female physicians? *Patient Education and Counseling, 48*(3), 217–224. https://doi.org/10.1016/S0738-3991(02)00174-X

Hall, J. A., Roter, D. L., Blanch-Hartigan, D., Mast, M. S., & Pitegoff, C. A. (2015). How patient-centered do female physicians need to Be? Analogue patients' satisfaction with male and female physicians' identical behaviors. *Health Communication, 30*(9), 894–900. https://doi.org/10.1080/10410236.2014.900892

Harris, A. J. L., & Hahn, U. (2011). Unrealistic optimism about future life events: A cautionary note. *Psychological Review, 118*(1), 135–154. https://doi.org/10.1037/a0020997

Haupt, C. (2019). Artificial professional advice. *Yale Journal of Health Policy, Law, and Ethics, 18*(3), 55–77. https://ssrn.com/abstract=3400898.

Helbig-Lang, S., & Petermann, F. (2010). Tolerate or eliminate? A systematic review on the effects of safety behavior across anxiety disorders. *Clinical Psychology: Science and Practice, 17*(3), 218–233. https://doi.org/10.1111/j.1468-2850.2010.01213.x

Holzinger, A., Dehmer, M., Emmert-Streib, F., Cucchiara, R., Augenstein, I., Ser, J. D., Samek, W., Jurisica, I., & Díaz-Rodríguez, N. (2022). Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence. *Information Fusion, 79*, 263–278. https://doi.org/10.1016/j.inffus.2021.10.007

Hou, Y. T.-Y., & Jung, M. F. (2021). Who is the expert? Reconciling algorithm aversion and algorithm appreciation in AI-supported decision making. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW2), 1–25. https://doi.org/10.1145/3479864

Hudecek, M., Fischer, P., Gaube, S., & Lermer, E. (2022). Who thinks COVID-19 is a hoax? Psychological correlates of beliefs in conspiracy theories and attitudes towards anti-coronavirus measures at the end of the first lockdown in Germany. *Journal of Open Psychology Data, 10*, 9. https://doi.org/10.5334/jopd.64

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, 2*, Article e47353. https://doi.org/10.2196/47353

Jiang, X., Xie, H., Tang, R., Du, Y., Li, T., Gao, J., Xu, X., Jiang, S., Zhao, T., Zhao, W., Sun, X., Hu, G., Wu, D., & Xie, G. (2021). Characteristics of online health care services from China's largest online medical platform: Cross-sectional survey study. *Journal of Medical Internet Research, 23*(4), Article e25817. https://doi.org/10.2196/25817

Johnson, K. B., Wei, W., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Zhao, J., & Snowdon, J. L. (2021). Precision medicine, AI, and the future of personalized health care. *Clinical and Translational Science, 14*(1), 86–93. https://doi.org/10.1111/cts.12884

Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why are we averse to algorithms? A comprehensive literature review on algorithm aversion. In *Proceedings of the 28th European conference on information systems (ECIS)* (Vol. 168). https://aisel.aisnet.org/ecis2020_rp/168.

Kaufmann, E., & Wittmann, W. W. (2016). The success of linear bootstrapping models: Decision domain-, expertise-, and criterion-specific meta-analysis. *PLoS One, 11*(6), Article e0157914. https://doi.org/10.1371/journal.pone.0157914

Larkin, C., Drummond Otten, C., & Árvai, J. (2022). Paging Dr. JARVIS! Will people accept advice from artificial intelligence for consequential risk management decisions? *Journal of Risk Research, 25*(4), 407–422. https://doi.org/10.1080/13669877.2021.1958047

Lennartz, S., Dratsch, T., Zopfs, D., Persigehl, T., Maintz, D., Große Hokamp, N., & Pinto Dos Santos, D. (2021). Use and control of artificial intelligence in patients across the medical workflow: Single-center questionnaire study of patient perspectives. *Journal of Medical Internet Research, 23*(2), Article e24221. https://doi.org/10.2196/24221

Lermer, E., Hudecek, M. F. C., Gaube, S., Raue, M., & Batz, F. (2021). Early and later perceptions and reactions to the COVID-19 pandemic in Germany: On predictors of behavioral responses and guideline adherence during the restrictions. *Frontiers in Psychology, 12*, Article 769206. https://doi.org/10.3389/fpsyg.2021.769206

Lermer, E., Streicher, B., Sachs, R., & Frey, D. (2013). How risky? The impact of target person and answer format on risk assessment. *Journal of Risk Research, 16*(7), 903–919. https://doi.org/10.1080/13669877.2012.761267

Lermer, E., Streicher, B., Sachs, R., Raue, M., & Frey, D. (2015). The effect of construal level on risk-taking: Construal level and risk. *European Journal of Social Psychology, 45*(1), 99–109. https://doi.org/10.1002/ejsp.2067

Lermer, E., Streicher, B., Sachs, R., Raue, M., & Frey, D. (2016a). Thinking concretely increases the perceived likelihood of risks: The effect of construal level on risk estimation: Effect of construal level on risk estimates. *Risk Analysis, 36*(3), 623–637. https://doi.org/10.1111/risa.12445

Lermer, E., Streicher, B., Sachs, R., Raue, M., & Frey, D. (2016b). The effect of abstract and concrete thinking on risk-taking behavior in women and men. *Sage Open, 6*(3), Article 215824401666612. https://doi.org/10.1177/2158244016666127

Liang, X., Kong, Y., & Peng, G. (2021). Online medical platform oriented identification of user needs related to alzheimer's disease and ontology construction. In N. Streitz, & S. Konomi (Eds.), *Distributed, Ambient and pervasive interactions (S. 316–330)*. Springer International Publishing. https://doi.org/10.1007/978-3-030-77015-0_23

Lim, D., Jay), Lin, J.-S., Chung, U. C., & Ko, Y. (2021). The role of construal fit in threat appeal to persuade young drivers not to text while driving. *Journal of Social Marketing, 11*(4), 406–423. https://doi.org/10.1108/JSOCM-06-2020-0109

Lindner, C., Kotta, I., Marschalko, E. E., Szabo, K., Kalcza-Janosi, K., & Retelsdorf, J. (2022). Increased risk perception, distress intolerance and health anxiety in stricter lockdowns: Self-control as a key protective factor in early response to the COVID-19 pandemic. *International Journal of Environmental Research and Public Health, 19*(9), 5098. https://doi.org/10.3390/ijerph19095098

Liu, B., & Sundar, S. S. (2018). Should machines express sympathy and empathy? Experiments with a health advice chatbot. *Cyberpsychology, Behavior, and Social Networking, 21*(10), 625–636. https://doi.org/10.1089/cyber.2018.0110

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes, 151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research, 46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Lu, X., Zhang, R., Wu, W., Shang, X., & Liu, M. (2018). Relationship between internet health information and patient compliance based on trust: Empirical study. *Journal of Medical Internet Research, 20*(8), e253. https://doi.org/10.2196/jmir.9364

Maassen, O., Fritsch, S., Palm, J., Deffge, S., Kunze, J., Marx, G., Riedel, M., Schuppert, A., & Bickenbach, J. (2021). Future medical artificial intelligence application requirements and expectations of physicians in German university hospitals: Web-based survey. *Journal of Medical Internet Research, 23*(3), Article e26646. https://doi.org/10.2196/26646

Madhav, A. V. S., & Tyagi, A. K. (2022). The world with future technologies (Post-COVID-19): Open issues, challenges, and the road ahead. In A. K. Tyagi, A. Abraham, A. Kaklauskas, & Hrsg (Eds.), *Intelligent interactive multimedia Systems for e-healthcare applications (S. 411–452)*. Springer Singapore. https://doi.org/10.1007/978-981-16-6542-4_22

Maechler, M. (2022). Oktober 14). *Package „robustbase"*. https://cran.r-project.org/web/packages/robustbase/robustbase.pdf.

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods, 52*(2), 464–488. https://doi.org/10.3758/s13428-019-01246-w

Mast, M. S. (2007). On the importance of nonverbal communication in the physician–patient interaction. *Patient Education and Counseling, 67*(3), 315–318. https://doi.org/10.1016/j.pec.2007.03.005

Mast, M. S., Hall, J. A., & Roter, D. L. (2007). Disentangling physician sex and physician communication style: Their effects on patient satisfaction in a virtual medical visit. *Patient Education and Counseling, 68*(1), 16–22. https://doi.org/10.1016/j.pec.2007.03.020

Mast, M. S., & Kadji, K. K. (2018). How female and male physicians' communication is perceived differently. *Patient Education and Counseling, 101*(9), 1697–1701. https://doi.org/10.1016/j.pec.2018.06.003

Mohd Salleh Sahimi, H., Azman, N., Nik Jaafar, N. R., Mohd Daud, T. I., Baharudin, A., Ismail, A. K., Abdul Malek, A. Z., Hassan, M. R., & Mohammed Nawi, A. (2021). Health anxiety and its correlations with self-perceived risk and attitude on COVID-19 among Malaysian healthcare workers during the pandemic. *International Journal of Environmental Research and Public Health, 18*(9), 4879. https://doi.org/10.3390/ijerph18094879

Nadarzynski, T., Bayley, J., Llewellyn, C., Kidsley, S., & Graham, C. A. (2020). Acceptability of artificial intelligence (AI)-enabled chatbots, video consultations and live webchats as online platforms for sexual health advice. *BMJ Sexual & Reproductive Health, 46*(3), 210–217. https://doi.org/10.1136/bmjsrh-2018-200271

Nelson, C. A., Pérez-Chada, L. M., Creadore, A., Li, S. J., Lo, K., Manjaly, P., Pournamdari, A. B., Tkachenko, E., Barbieri, J. S., Ko, J. M., Menon, A. V., Hartman, R. I., & Mostaghimi, A. (2020). Patient perspectives on the use of artificial intelligence for skin cancer screening: A qualitative study. *JAMA Dermatology, 156*(5), 501. https://doi.org/10.1001/jamadermatol.2019.5014

Neyer, F. J., Felber, J., & Gebhardt, C. (2016). Kurzskala Technikbereitschaft (TB, technology commitment). In *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. https://doi.org/10.6102/ZIS244

Nguyen, G. C., LaVeist, T. A., Harris, M. L., Datta, L. W., Bayless, T. M., & Brant, S. R. (2009). Patient trust-in-physician and race are predictors of adherence to medical management in inflammatory bowel disease. *Inflammatory Bowel Diseases, 15*(8), 1233–1239. https://doi.org/10.1002/ibd.20883

Peng, J., He, F., Zhang, Y., Liu, Q., Miao, D., & Xiao, W. (2013). Differences in simulated doctor and patient medical decision making: A construal level perspective. *PLoS One, 8*(11), Article e79181. https://doi.org/10.1371/journal.pone.0079181

Promberger, M., & Baron, J. (2006). Do patients trust computers? *Journal of Behavioral Decision Making, 19*(5), 455–468. https://doi.org/10.1002/bdm.542

Rice, S., Winter, S. R., Mehta, R., & Ragbir, N. K. (2019). What factors predict the type of person who is willing to fly in an autonomous commercial airplane? *Journal of Air Transport Management, 75*, 131–138. https://doi.org/10.1016/j.jairtraman.2018.12.008.

Richardson, J. P., Curtis, S., Smith, C., Pacyna, J., Zhu, X., Barry, B., & Sharp, R. R. (2022). A framework for examining patient attitudes regarding applications of artificial intelligence in healthcare. *Digital Health, 8*. https://doi.org/10.1177/20552076221089084

Richardson, J. P., Smith, C., Curtis, S., Watson, S., Zhu, X., Barry, B., & Sharp, R. R. (2021). Patient apprehensions about the use of artificial intelligence in healthcare. *NPJ Digital Medicine, 4*(1), 140. https://doi.org/10.1038/s41746-021-00509-1

Robertson, C., Woods, A., Bergstrand, K., Findley, J., Balser, C., & Slepian, M. J. (2023). Diverse patients' attitudes towards Artificial Intelligence (AI) in diagnosis. *PLOS Digital Health, 2*(5), Article e0000237. https://doi.org/10.1371/journal.pdig.0000237

Schandl, F., Lermer, E., & Hudecek, M. F. C. (2023). If it concerns me: An experimental investigation of psychological distance and the acceptance of autonomous shuttle buses [Preprint]. *Open science framework*. https://doi.org/10.31219/osf.io/ac9np

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

Schermelleh-Engel, K., & Werner, C. S. (2012). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger, A. Kelava, & Hrsg (Eds.), *Testtheorie und fragebogenkonstruktion (S. 119–141)*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20072-4_6.

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., Sariyska, R., Stavrou, M., Becker, B., & Montag, C. (2021). Assessing the attitude towards artificial intelligence: Introduction of a Short measure in German, Chinese, and English language. *KI - Künstliche Intelligenz, 35*(1), 109–118. https://doi.org/10.1007/s13218-020-00689-0

Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and "how to" guide of its application within vocational behavior research. *Journal of Vocational Behavior, 120*, Article 103445. https://doi.org/10.1016/j.jvb.2020.103445

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440–463. https://doi.org/10.1037/a0018963

Trope, Y., Liberman, N., & Wakslak, C. (2007). Construal levels and psychological distance: Effects on representation, prediction, evaluation, and behavior. *Journal of Consumer Psychology, 17*(2), 83–95. https://doi.org/10.1016/S1057-7408(07)70013-X

Wang, X., Duan, X., Li, S., & Bu, T. (2022). Effects of message framing, psychological distance, and risk perception on exercise attitude in Chinese adolescents. *Frontiers in Pediatrics, 10*, Article 991419. https://doi.org/10.3389/fped.2022.991419

Wang, S., Hurlstone, M. J., Leviston, Z., Walker, I., & Lawrence, C. (2019). Climate change from a distance: An analysis of construal level and psychological distance from climate change. *Frontiers in Psychology, 10*, 230. https://doi.org/10.3389/fpsyg.2019.00230

Weyrich, P., Scolobig, A., Walther, F., & Patt, A. (2020). Do intentions indicate actual behaviour? A comparison between scenario-based experiments and real-time observations of warning response. *Journal of Contingencies and Crisis Management, 28*(3), 240–250. https://doi.org/10.1111/1468-5973.12318

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90*(2), 245–271. https://doi.org/10.1037/0033-2909.90.2.245

Wu, J., Xu, L., Yu, F., & Peng, K. (2021). Acceptance of medical treatment regimens provided by AI vs. Human. *Applied Sciences, 12*(1), 110. https://doi.org/10.3390/app12010110

Yakar, D., Ongena, Y. P., Kwee, T. C., & Haan, M. (2022). Do people favor artificial intelligence over physicians? A survey among the general population and their view on artificial intelligence in medicine. *Value in Health, 25*(3), 374–381. https://doi.org/10.1016/j.jval.2021.09.004

You, S., Yang, C. L., & Li, X. (2022). Algorithmic versus human advice: Does presenting prediction performance matter for algorithm appreciation? *Journal of Management Information Systems, 39*(2), 336–365. https://doi.org/10.1080/07421222.2022.2063553

Young, A. T., Amara, D., Bhattacharya, A., & Wei, M. L. (2021). Patient and general public attitudes towards clinical artificial intelligence: A mixed methods systematic review. *The Lancet Digital Health, 3*(9), e599–e611. https://doi.org/10.1016/S2589-7500(21)00132-1

Zhang, R., Kang, H., Jiang, Z., & Niu, X. (2023). How does workplace ostracism hurt employee creativity? Thriving at work as a mediator and organization-based self-esteem as a moderator. *Applied Psychology, 72*(1), 211–230. https://doi.org/10.1111/apps.12374