The Impact of AI Perceived Transparency on Trust in AI Recommendations in Healthcare applications

Shabankareh, M., Khamoushi Sahne, SS., Nazarian, A. and Foroudi, P.

The final, published version in Asia-Pacific Journal of Business Administration is available at:

https://doi.org/10.1108/APJBA-12-2024-0690

The Impact of AI Perceived Transparency on Trust in AI Recommendations in Healthcare applications

Abstract

Purpose- The integration of artificial intelligence (AI) in healthcare has transformed the way users interact with health applications, offering personalized recommendations and decision-making support. However, building trust in AI-driven systems remains a significant challenge, particularly in high stakes environments like healthcare, where user concerns about fairness, control, and privacy are paramount. This study aims to investigate how AI transparency influences trust in healthcare applications, focusing on the mediating roles of perceived fairness and control, and the moderating role of privacy concerns.

Design/methodology/approach- A quantitative research design was employed, utilizing survey data collected from healthcare application users. Structural Equation Modeling (SEM) and moderation analysis were used to test the proposed conceptual framework, exploring the interrelationships among the variables.

Findings- The results revealed that AI transparency significantly influences trust in healthcare applications indirectly through perceived fairness, while perceived control had a limited mediating effect. Privacy concerns were found to amplify the relationship between fairness and trust but did not significantly moderate the effects of transparency or control on trust. These findings emphasize the central role of fairness and privacy in building trust, highlighting the nuanced interplay between ethical perceptions and user concerns in high-stakes contexts.

Originality- This study contributes to the literature by integrating fairness, control, and privacy concerns into a unified framework for understanding trust in AI healthcare applications. By demonstrating how transparency operates indirectly and how privacy concerns shape user perceptions, this research provides novel insights for designing ethically robust and user-centric AI systems tailored to sensitive domains like healthcare.

Keywords: AI transparency, Recommender systems, Healthcare applications, AI fairness, Privacy concerns

1. Introduction

Integration of AI in healthcare demonstrates incredible growth in respect of taking care of health for professionals and individuals (Chen and Decary, 2020). AI Health Applications have become necessary and vital in that they give personalized real-time recommendations based on user-specific information. In such situations, patients depend on these applications even to monitor their well-being and chronic conditions (Samal et al., 2021). For instance, applications remind patients to take drugs or prescribe symptom checkers driven by AI that would help patients make decisions on whether a doctor's visit is required (Tan et al., 2024). The applications are not constrained to patients only. Even healthy persons, athletes, and fitness-conscious persons get different benefits from these AI-driven applications for tracking their fitness and managing their stress levels. Some of them offer customized mindfulness exercises which cater to mental well-being and relaxation. In essence, AI health applications analyze diversified real-time data sources such as wearables, sensors, and manual inputs to provide personalized recommendations. More importantly, the application does predict health risks for proactive measures, besides providing healthcare providers with insights derived from data to optimize diagnosis and treatment (Karatas et al., 2022).

There should be no compromise on transparency regarding AI-based health applications. When users are confused as to how AI gets recommendations, this creates suspicions about the system's reliability-particularly for high-stakes domains like health, in which decisions relate to the well-being of oneself and others. It goes ahead to state that AI systems do not provide clarity on how it works to ensure comfort in their advice for reliance. The World Health Organization and organizations like the Bipartisan Policy Center highlight the fact that transparency in AI-driven healthcare will instill trust and ensure ethics, as AI becomes more and more integral in diagnosis, treatment, and efficiency (WHO, 2021, Adams, 2024). According to experts, one of the major reasons for that is that both providers and patients need to trust this technology. Transparency means that AI systems' decisions should be understandable and explainable to users, further building trust in AI-driven care recommendations Castelo et al. (2019) found that while users tend to believe algorithms more in objective tasks such as financial advice, they are a bit skeptical when it comes to more subjective decisions like medical recommendations-even when the algorithms outperform humans. Studies have shown that people would rather take medical advice from human doctors rather than an algorithm. Studies by Promberger and Baron (2006) and Longoni et al. (2019) suggest this skepticism stems from concerns that algorithms might not account for individual circumstances, fueling mistrust despite AI's proven accuracy. This evidence-supported belief contributes to mistrust in AI-powered application processes despite proven accuracy, which makes users hold back from following recommended advice put forward by applications. Although these users may question the ability of AI in personalising care, several studies have proven that most of the time, AI systems show human performance levels or even outperform humans in medical activities. Such was a conclusion by Xie et al. (2019) that AI models diagnosed skin cancer with amazing accuracy matching the performance of experienced dermatologists. In the meantime, other works have underlined that AI is able to integrate complex datasets and analyze them-such as by Schork (2019) and Ahmed et al. (2020) provide recommendations peculiar to each health profile for every patient-specific condition, proving that AI can offer highly personalized care.

Although extensive research demonstrates that AI systems can deliver highly accurate and personalized medical recommendations, the lack of transparency in their decision-making processes remains a critical barrier to user trust, even when AI performance is proven to exceed that of human experts. One major issue that hasn't been addressed despite the growing usage of AI in healthcare is how user trust in AI suggestions is influenced by perceived control, transparency, and fairness. Previous research has examined these elements separately, but little is known about how they interact, especially in delicate settings like healthcare where privacy issues make building trust even more difficult. Designing AI systems that people are willing to use and rely on requires an understanding of these interactions. Fehr et al. (2024) found that many AI medical products in Europe lack transparency, particularly in fairness, bias, and data validation. This highlights the need for stricter legal transparency requirements. Similarly, Shick et al. (2024) point out that patients and caregivers often feel uneasy about AI devices due to limited awareness and concerns over their impact on care, stressing the need for educational resources and addressing issues such as data security, costs, and technical requirements. Bernal and Mazo (2022) emphasize that while AI has transformative potential, gaining trust from healthcare professionals and patients necessitates enhanced transparency and the establishment of new regulations for AI's design, validation, and deployment. Moreover, Galiana et al. (2024) stress that the ethical implications of AI in medicine are complex and crucial, highlighting the need for AI technologies to be safe, fair, and respectful of patient privacy, with a strong emphasis on transparency and ongoing training for professionals. Kiseleva et al. (2022) propose viewing AI transparency as a multilayered system of measures, suggesting that interpretability and transparency must be context-specific and shaped by legal frameworks and expectations. These studies collectively emphasize the need to balance AI's integration into healthcare with transparency, fairness, and ethics, highlighting how critical these factors are to ensuring trust and successful implementation.

As AI is becoming more integrated into everyday life, healthcare applications are increasingly widespread, providing users with 24/7 access to health services and information. Prior research has often overlooked the evaluation of AI applications in healthcare from the user's perspective, particularly concerning transparency, perceived fairness, and perceived control and their impact on user trust in AI-driven recommendations and applications/platforms remains underexplored. This study seeks to address this gap by empirically investigating the interplay of AI transparency, fairness, and perceived control in influencing user trust in healthcare applications, with privacy concerns serving as a moderating variable. To achieve this, the study is guided by the following research questions: The main research question is: How does AI transparency influence user trust in healthcare applications? Additionally, the study explores two sub-questions: (1) What are the mediating roles of perceived fairness and perceived control in this relationship? and (2) How do privacy concerns moderate the impact of AI transparency, fairness, and control on user trust? This research enhances the knowledge of user trust in AI-driven healthcare and provides valuable insights for designing systems that meet user expectations and ethical norms. Besides, given the great importance of privacy in health contexts, this research investigates the moderating role played by privacy concerns. Various authors have underlined that in the case of sensitive health information, privacy concerns become paramount, and users become highly sensitive to data security issues (Awotunde et al., 2021, Chenthara et al., 2019, Abouelmehdi et al., 2018). Understanding how privacy concerns interact with transparency, fairness, and user control is crucial for fostering trust in AI systems within healthcare. By empirically testing these relationships, this study not only aims to enhance understanding of how user trust is shaped regarding AI-generated health recommendations but also responds to prior calls for research by Bach et al. (2024) to evaluate trust concepts in specific AI-enabled contexts. This research bridges the knowledge gap, enhancing theoretical understanding of trust in AI while offering practical insights for developers and policymakers. The results can inform the development of AI-based healthcare apps that are regarded as more transparent, equitable, and privacy-aware, hence enhancing user confidence and adoption.

2. Literature review

2.1. Theoretical foundations

This research formulates the study's hypotheses by using theoretical insights from multiple established frameworks that elucidate user trust in AI-driven decision-making systems. Specifically, trust theory (Mayer et al., 1995) serves as the foundation for understanding how transparency, fairness, and control influence trust in AI recommendations. This theory asserts that trust is established when individuals regard a system as competent, benevolent, and ethical attributes that are closely linked to transparency and fairness. Additionally, procedural justice theory (Thibaut, 1975) underscores the significance of fairness and openness in the establishment of trust, indicating that individuals are more inclined to trust institutions that adhere to transparent, impartial, and equitable decision-making procedures. Furthermore, theory of planned behavior (Ajzen, 1991) establishes a basis for comprehending how users' capacity to understand and impact AI-driven decisions influences their trust. Lastly, privacy calculus theory (Culnan and Armstrong, 1999) elucidates how privacy apprehensions serve as a moderating element in the establishment of trust, suggesting that users evaluate possible dangers in relation to the advantages of AI-generated suggestions prior to developing trust. By integrating these theoretical perspectives, this study offers a comprehensive framework for examining the complex interplay between AI transparency, fairness, control, privacy concerns, and user trust in healthcare applications.

To complement the foundational frameworks discussed above, recent scholarship has increasingly emphasized the importance of ethical principles, explainability, and adaptive trust-building mechanisms in AI systems. Research in AI ethics has highlighted the need for systems to be aligned with values such as accountability, transparency, and user autonomy, especially in sensitive domains like healthcare (Siala and Wang, 2022, Singhal et al., 2024). Explainable AI (XAI) frameworks further expand this perspective by focusing on the interpretability and accessibility of algorithmic processes for end-users (Chaddad et al., 2023, Loh et al., 2022). These frameworks argue that providing users with meaningful, understandable insights into how AI systems operate fosters a dynamic and iterative process of trust formation. By integrating these emerging perspectives, this study contributes to a more nuanced and context-sensitive understanding of user trust in healthcare AI applications.

2.2. Perceived transparency of AI and trust in AI recommendations

AI transparency has been defined in various ways across the literature, reflecting its complexity and importance in different contexts. Liao and Vaughan (2023) describes transparency as the ability to provide relevant stakeholders with the necessary understanding of an AI model's capabilities, limitations, functionality, and how to control or utilize its outputs, thereby fostering informed usage. Bernal and Mazo (2022) and Bhatt et al. (2020) offers a more detailed view, defining AI transparency as the design of algorithms that are inherently intelligible to humans, whether independently or with external tools. This involves sharing comprehensive information about the algorithm's processes, including documentation, validation procedures, dataset descriptions, data analysis, and model outputs. Similarly, Shin (2021) highlights three critical components of AI transparency: Understandability, which requires that algorithmic evaluations be accessible and comprehensible to the public; Explainability, ensuring that AI-generated outputs can be easily understood by affected individuals; and Observability, which focuses on enabling users to grasp the relationship between an algorithm's internal processes and its external results.

Based on the definitions provided, user Perceived Transparency of AI (PTAI) in healthcare refers to how well users can understand and engage with the AI systems used in health applications. This includes several dimensions: first, understandability, where the inner workings and criteria of the algorithms are presented in a way that is accessible and easily comprehended by non-expert users. Second, explainability, which ensures that the outputs, such as recommendations or diagnoses, are clearly articulated and justifiable to those affected. Finally, observability allows users to trace and comprehend how the AI's internal processes correspond to its external results. Together, these elements ensure that healthcare AI systems are transparent in their operations, allowing users to grasp how decisions are made and how these systems function.

Trust in AI Recommendations (TAIR) refers to the degree of confidence users have in the outputs or suggestions provided by AI systems, based on their belief that these recommendations are reliable, trustworthy, and the vendor's commitment to fulfilling obligations in their exchange relationship (Shin, 2021, Shin and Park, 2019). This view highlights that in the healthcare contexts, where the need for reliable information is paramount, users are particularly attentive to the level of detail provided, leading to greater trust when there is openness about the system's processes. The concept of user trust, as elaborated by Gefen and Straub (2003), extends to AI-driven systems by asserting that trust forms when users can understand and make sense of the outputs and processes of the technology they are interacting with in e-services. Transparency is key here which will enable users to get an inside view of how AI works and thus give them the perception that this is a more reliable system, rather than an opaque one. Empirical studies will underpin or bring into focus the fact that AI transparency can have significant effects on users' perceptions (Felzmann et al., 2020, Shin and Park, 2019, Shin, 2021). This is to say, users appeared to be better informed when systems made their outputs and decision-making processes known and could understand the logic that lay behind recommendations made by AI. This understanding, in turn, influences user satisfaction and perceived control, making them more willing to engage with and rely on AI systems. In healthcare applications, where personal health data and critical decision-making are involved, transparency ensures that users feel more at ease with AI-driven diagnoses and recommendations, ultimately leading to greater adherence to these systems. While previous studies have identified transparency as a vital element in the adoption of AI. However, much of the research has concentrated on general AI applications, neglecting high-stakes contexts like healthcare, where the mechanisms for building trust may vary considerably. The current body of literature primarily examines transparency through a usability lens, highlighting its significance in enhancing system understanding and mitigating uncertainty. Limited research has focused on the interaction between transparency, fairness perceptions, and user control in influencing trust, especially in the context of sensitive health data. This study addresses the identified gap by building on trust theory (Mayer et al., 1995) , which asserts that system integrity is essential for the formation of trust. Transparency improves perceptions of integrity by minimizing opacity in AI decision-making, thus promoting confidence in AI recommendations. We propose the following hypothesis:

H1. PTAI has a significant effect on TAIR in the healthcare applications

2.3. PTAI, Perceived fairness of AI, Perceived control over AI, and TAIR

Li and Zheng (2024) defines AI fairness as the capacity of AI systems to make decisions that are unbiased, non-discriminatory, and fair, focusing on the outcome of AI decisions being free from discrimination or bias. Other scholars further extend this understanding of AI fairness by emphasizing three core dimensions (Shin, 2021, Shin and Park, 2019). The first is nondiscrimination, which states that the AI system is unbiased toward any group. Then there is accuracy, which places great emphasis on the algorithm's data sources in identifying, logging, and benchmarking for fairness. Lastly, due process involves a belief that the AI system has impartial processes with no prejudice in the decision-making process. Therefore, user Perceived Fairness of AI (PFAI) in healthcare applications can be defined as the user's belief that the AI system operates without bias, ensures accuracy by transparently managing data sources, and follows impartial procedures in making decisions, leading to outcomes that are equitable and non-discriminatory.

Procedural justice theory (Thibaut, 1975), which emphasizes that people consider procedures to be fair when decision-making procedures are more transparent and provide enough information on how the results are produced. Based on this theory, transparency in decision-making would lead to increased perception of fairness in procedures, because one feels more informed and part of the process, and thus one trusts the system more. Building on this, empirical work in algorithmic decision-making has shown that indeed, transparency is one of the important ways through which concerns over bias and unfairness are mitigated. For example, Binns et al. (2018) show that when algorithms provide clear and understandable explanations for their outputs, users are more likely to perceive these systems as fair, even in complex environments such as healthcare. This finding suggests that transparency helps bridge the gap between the technical workings of AI and users' concerns about potential biases, leading to more favorable perceptions of fairness (Felzmann et al., 2020, Memarian and Doleck, 2023). When users believe that an AI system makes decisions impartially, without bias or favoritism, and follows transparent processes, they are more likely to view the system as trustworthy. This perception of fairness, encompassing nondiscrimination, data accuracy, and due process, is expected to foster greater confidence in AI-driven healthcare recommendations. Doshi-Velez et al. (2017) concluded that that transparency and fairness in AI systems improve users' understanding of how AI reaches its decisions, fostering trust in these decisions. Moreover, Topol (2019) and Binns et al. (2018) findings indicated that that when AI systems in healthcare provide transparent and equitable treatment (thus appearing fair), users are more inclined to accept and rely on the system's advice.

The role of transparency in shaping fairness perceptions becomes particularly critical in AI-driven healthcare, where decision-making algorithms must navigate concerns of bias, accountability, and ethical integrity (Binns et al., 2018, Felzmann et al., 2020). Users assessing AI-driven suggestions consider not just the fairness of outcomes but also the system's transparency and compliance with procedural justice standards. This corresponds with results indicating that when individuals sense openness in decision-making processes, they are more inclined to deduce justice and cultivate faith in the system's ethical integrity. Although fairness is extensively examined in AI ethics, the primary emphasis has been on alleviating algorithmic bias, with insufficient consideration of how perceptions of fairness develop in relation to transparency in user-centric applications. This study broadens this viewpoint by investigating fairness not just as an ethical limitation, but as a perceptual process whereby transparency cultivates trust.

Therefore, it can be hypothesized that:

H2. PTAI has a significant effect on PFAI in the healthcare applications

H3. PFAI has a significant effect on TAIR in the healthcare applications

Perceived control can be defined as the extent to which individuals believe they possess the necessary resources (such as time, skills, or financial capacity), opportunities, and abilities to perform a particular behavior (Tucker et al., 2020). Previous studies showed that the more a user feels that they have these capabilities and resources, the more confident they are in using these services in the context of online services (Tucker et al., 2020, Belanche et al., 2022). This sense of control boosts their likelihood to engage with online services, as they feel empowered to navigate and influence outcomes in the digital environment. In this study, perceived control over AI (PCAI) refers to a user's sense of being in charge of their actions when using AI systems, feeling that the use of the system is clear, manageable, and under their control. It reflects the absence of confusion and the user's confidence in navigating the platform. When users feel that they can make informed decisions while interacting with the AI, it boosts their perceived control, leading to greater engagement with the healthcare system. In their study, Rohden and Espartel (2024) highlight that risk-averse users may view decisions assisted by recommendation algorithms as more susceptible to negative outcomes, particularly due to the fear generated by a lack of understanding of how the technology functions. This perception can reduce the consumer's sense of control over the choices they are making, as they may associate autonomous technologies, like recommendation systems, with diminished agency (De Bellis and Johar, 2020). In AI-driven healthcare applications, perceived control is especially pertinent when consumers interact with automated systems that affect their medical decision-making. The degree to which users perceive their ability to supervise and control AI interactions influences their confidence in these technologies. Although previous studies indicate that more transparency improves user understanding, there has been little focus on how this openness correlates with an enhanced perception of control over AI-generated suggestions. Transparency enhances the comprehensibility of AI outputs and cultivates a feeling of agency, enabling users to perceive themselves as active participants in decision-making rather than just consumers of automated recommendations. This corresponds with theory of planned behavior (Ajzen, 1991), which posits that individuals are more likely to trust and embrace technology when they believe they have control over their interactions. Therefore, we propose the following hypotheses:

H4. PTAI has a significant effect on PCAI in the healthcare applications

H5. PCAI has a significant effect on TAIR in the healthcare applications

2.4. The moderating role of Privacy concerns of AI

In the literature, Smith et al. (1996) define privacy concerns as "the degree to which individuals are concerned about how their personal information is collected, stored, and used by organizations". This definition emphasizes the worry individuals have about the handling of their data by organizations, whether through collection, storage, or usage, especially as personal data becomes more integral to digital services. Similarly, Dinev and Hart (2006) elaborate on privacy concerns, describing them as "an individual’s general tendency to worry about the collection and use of personal information by third parties". This perspective highlights a broader apprehension users have toward external parties, who may collect and potentially misuse their personal data without full transparency or control. Building on these foundational definitions, Privacy Concerns of AI (PC) in healthcare applications refers to users' anxieties regarding how AI systems collect, store, and utilize their sensitive health-related data. In such contexts, users are particularly concerned about how AI algorithms handle personal medical information, how secure these systems are, and whether third-party entities could gain unauthorized access. As AI becomes more integrated into healthcare apps, users' control over their data, the transparency of AI's decision-making processes, and potential misuse of this information become central aspects of privacy concerns.

The rise of digital platforms, apps, and online services has amplified users' anxieties about the collection, use, and security of their personal information (Kim et al., 2023). As Suh and Han (2003) suggest, privacy concerns are increasingly central to user trust in digital services as they directly impact users' perceptions of data security and control, especially when it comes to sensitive information. This highlights the growing importance of privacy across all sectors, including finance, healthcare, and e-commerce (Abbas and Khan, 2015, Maseeh et al., 2021). Additionally, Bansal et al. (2015) found that privacy concerns can significantly influence users' willingness to adopt new digital platforms. In their study, they observed that unresolved privacy issues can lead to a "lack of trust and hesitance in adopting or continuing to use online services, especially those that handle sensitive data." Furthermore, Martin et al. (2017) emphasized that privacy concerns are a major determinant of user trust in online platforms. They argue that if users feel their personal information is mishandled or exposed to potential risks, they are unlikely to trust or engage with the system fully. Abbas and Khan (2015) argue that e-health clouds can only gain worldwide acceptance if they earn the confidence and trust of healthcare organizations and patients by providing robust mechanisms to protect sensitive electronic health information. Similarly, Schomakers et al. (2019) highlight that privacy concerns are a significant barrier to the adoption of e-health technologies, influenced by perceived data sensitivity, trust in data protection, and individual privacy values.

Privacy issues not only hinder AI adoption but also influence consumers' assessments of transparency, fairness, and control in AI-driven healthcare services. Individuals with significant privacy apprehensions may exhibit skepticism towards transparent AI systems, interrogating whether enhanced openness correlates with improved data security or instead amplifies exposure to possible privacy threats. Likewise, apprehensions over equity may intensify in privacy-sensitive contexts, as users can view AI-generated conclusions as prejudiced or exploitative of personal information. Moreover, although perceived control often bolsters confidence, persons with significant privacy apprehensions may continue to feel insecure, despite believing they possess a degree of control over the system. This aligns with privacy calculus theory (Culnan and Armstrong, 1999), which posits that users make a cost-benefit analysis when engaging with technology, weighing the advantages of AI-driven recommendations against potential privacy risks. Given that privacy concerns can alter how users interpret transparency, fairness, and control, it is expected that privacy concerns will moderate these relationships. Specifically, we argue that privacy concerns negatively moderate the relationships between PTAI, PFAI, and PCAI with TAIR. Individuals with significant privacy apprehensions may continue to doubt AI-generated suggestions, notwithstanding the openness, equity, and user autonomy exhibited by AI systems. This concern emerges when privacy-conscious consumers frequently view transparency as heightened vulnerability to data threats instead of a guarantee of ethical AI techniques. Based on these, we hypothesize:

H6. PC moderate the relationship between PTAI and TAIR in the healthcare applications

H7. PC moderate the relationship between PFAI and TAIR in the healthcare applications

H8. PC moderate the relationship between PCAI and TAIR in the healthcare applications.

Figure 1 illustrates the conceptual model of the study.

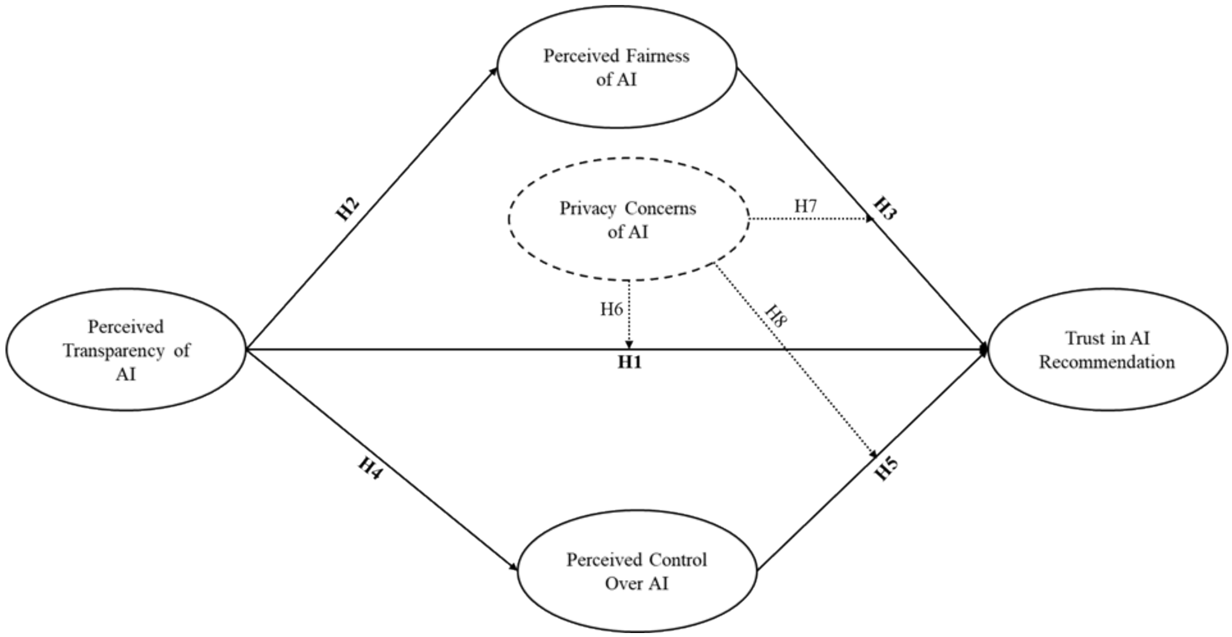

Figure 1- Conceptual model of the research

3. Methodology

3.1. Sample, data collection, and survey

The data for this study were collected in Iran between September 10, 2024, and October 5, 2024. As a developing country in the Middle East, Iran provides a unique context for the aim of this study. Data collection involved a self-administered questionnaire specifically designed to test the hypotheses. The questionnaire was set up using Google Docs to facilitate easy access and distribution among respondents. One of the authors was responsible for distributing and collecting the data, targeting individuals through various social media platforms, particularly those engaged with medical, health and fitness content. Additionally, data collection efforts extended to physical locations such as gyms, wellness clubs, and hospitals, where individuals are more likely to use AI-powered health/medical application. Participants were invited to participate voluntarily, and the responsible author provided a direct link to the questionnaire. The research utilized a convenience sample technique, focusing on individuals who often engage with AI-driven health and medical applications. Alongside internet distribution, participants were recruited from physical venues such as gyms, wellness centers, and hospitals, where they were contacted in person and offered to participate willingly. The recruiting technique guaranteed diverse replies from people participating in different health and fitness activities. Throughout the data collection period, the author was available to address any questions or concerns from participants, thereby enhancing the clarity and validity of the collected responses. The study sample included individuals from Alborz and Tehran provinces who actively used AI-powered health/medical applications. In total, 406 complete questionnaires were collected that were suitable for analysis.

The sample size of 406 respondents was determined based on prior recommendations for conducting SEM, which suggest a minimum of 10 responses per observed variable (Hair Jr et al., 2010). Additionally, a priori power analysis using GPower software indicated that a minimum of 384 responses would be required to detect medium effect sizes with a power of 0.80 and an alpha level of 0.05 (Faul et al., 2009). The final sample size exceeded this threshold, ensuring robust statistical analysis and providing a solid foundation for exploring the research questions posed in this study.

The questionnaire was carefully structured into three distinct sections to ensure clarity and ease of understanding for respondents. It began with an introduction expressing gratitude for their participation, outlining the study's objectives, and assuring them of confidentiality in handling their information. The first part included a key question about their use of AI-powered health/medical applications, with those who answered "yes" directed to the following section. The second part focused on collecting demographic information, such as age, usage frequency, duration, and primary purpose, which was essential for contextualizing the findings. The final section contained the main questions, organized into five sub-sections based on the study's variables, with each variable's name clearly labeled above its related questions. A 7-point Likert scale (Strongly disagree- Strongly agree) was used to assess attitudes. To maintain response integrity, a strategically placed deviation question was included to verify thoughtful engagement, prompting participants to revisit questions if inconsistencies were detected. This structured approach aimed to gather valuable data while enhancing engagement and reliability through thoughtful design elements. Table 1 provides a detailed summary of the demographic characteristics of the respondents. The dataset includes variables such as gender, marital status, age, education level, app usage duration, and frequency of daily app use.

'Table 1 Here'

3.2. Item measurements

The measurement items for this study were adapted from prior, well-established research in the field of artificial intelligence and carefully aligned with the study's objectives through minor modifications. Specifically, the items for the variables perceived transparency of AI was measured using three items capturing understandability, explainability, and observability, perceived fairness of AI was assessed through three dimensions: nondiscrimination, accuracy, and due process. Trust in AI recommendation was measured based on perceived reliability and trustworthiness of AI recommendations were sourced from (Shin, 2021). The items for perceived control over AI, which reflect a user's sense of control, were derived from Belanche et al. (2022) and comprised 3 items. Finally, privacy concern of AI, which capture awareness and control over personal information handling, were adapted from Chellappa (2008) and included 6 items. Initially, the items were extracted in English, and after minor adjustments to ensure alignment with the research goals, they were validated by two academic experts based in the UK. Following this, the items were translated into Persian (the official language of Iran) by a professional translator and were again reviewed and approved by the same experts.

A pilot study was then conducted to ensure the clarity, validity, and reliability of the questionnaire, aiming to identify any potential issues or ambiguities. The questionnaire was distributed to 25 business administration students, who were asked to flag any unclear or confusing statements. No significant issues were reported, and the instrument demonstrated strong reliability, with a Cronbach's alpha of 0.93, indicating high internal consistency. Once the pilot study confirmed the robustness of the questionnaire, it was administered to the main study population. This rigorous process ensured that the measurement items were both reliable and valid, supporting the credibility of the study's findings. To evaluate non-response bias, a Mann-Whitney U test was performed to compare early and late respondents. The results showed no significant differences, indicating that non-response bias was not present (Lambert and Harrington, 1990).

4. Results

4.1. Construct Validity

Initial tests for factor analysis, validity, and reliability were conducted using the sample data to refine the research measures. A two-phase approach, as recommended by Anderson and Gerbing (1988), was employed. During the first phase, exploratory factor analyses were carried out to evaluate the construct groupings, ensuring alignment with the expected theoretical factor structures. A detailed overview of the research constructs is provided in Table 2. To ensure reliability, the internal consistency of the indicators was assessed to confirm their adequacy in representing the associated latent constructs. Following the guidelines of Bagozzi and Yi (1988), all measures were required to achieve a reliability coefficient (rho) greater than 0.70 (Hair, 2009, Nunnally, 1978). Appendix 1 presents the CR and AVE values for all constructs. Composite Reliability (CR) measures the internal consistency of the items, with a threshold of ≥0.70 indicating good reliability (Hair et al., 2010). Average Variance Extracted (AVE) assesses the amount of variance captured by a construct's items relative to measurement error, with a recommended threshold of ≥0.50 (Fornell and Larcker, 1981, Foroudi and Dennis, 2023). The results confirm that all constructs meet these criteria, supporting the reliability and convergent validity of the measurement model.

'Table 2 Here'

We performed Confirmatory Factor Analysis (CFA) to thoroughly assess the unidimensionality of the research constructs. This process involved testing subsets of items and validating the constructs through measurement models by examining their internal consistency, as suggested by Anderson and Gerbing (1988). During these tests, we applied a constraint that fixed the relationship between each pair of latent variables to 1. This constraint consistently resulted in a significant reduction in model fit ($\Delta\chi^2 = 2.10$; df = 1–4; $p < 0.01$), in line with Anderson and Gerbing (1988) findings. Additionally, we compared the variance extracted from each construct against the squared off-diagonal values in the Phi matrix (Fornell and Larcker, 1981). This analysis confirmed that the items for each construct measured distinct underlying concepts. Skewness and kurtosis values for all variables were within the ±2 range, confirming univariate normality. Linearity was evaluated via bivariate scatterplots and Pearson correlations, showing linear trends among variables. Multicollinearity diagnostics using Variance Inflation Factor (VIF) values were all below 3, indicating no multicollinearity concerns. To further evaluate the constructs, Pearson's correlation matrix (two-tailed) was employed at a 0.01

significance level to examine linearity and multi-collinearity. The results revealed that most independent variables had significant positive correlations with the dependent variables, and the majority demonstrated linear relationships.

A common recommendation is that, in addition to χ^2 results, researchers should report at least one absolute fit index and one incremental fit index (Abbas and Khan, 2015), including their values and degrees of freedom. In line with this guideline, several indices were used to evaluate the model's fit: the Comparative Fit Index (CFI) achieved a value of 0.984, exceeding the recommended threshold of 0.90; the Tucker-Lewis Index (TLI) scored 0.979, also above 0.90; the Incremental Fit Index (IFI) reached 0.984, surpassing 0.90; and the Root Mean Square Error of Approximation (RMSEA) was 0.033, well below the 0.08 threshold (Garver and Mentzer, 1999, Hair et al., 2006, Tabachnick, 2007). The chi-square value was 136.254 with 94 degrees of freedom. These indices collectively indicate a strong and comprehensive model fit. Steenkamp and Van Trijp (1991) confirmed the nomological validity of the measurement model for these three factors. Additional fit indices, including the Relative Fit Index (RFI) at 0.936 and the Normed Fit Index (NFI) at 0.950, also exceeded the 0.90 threshold, further supporting the robustness of the measurement model (Hair et al., 2006). Achieving such a strong fit can be challenging, but these results provide robust support for the model. To address potential common method bias, Harman's one-factor test was first applied. Additionally, latent factors were analyzed by comparing the chi-square difference between fully constrained and original models, following the recommendations of Malhotra et al. (2006) and Podsakoff et al. (2003). The shared variance between these models revealed statistically significant differences. Furthermore, a latent common method factor was introduced in a comparative model using the unmeasured latent method construct approach. The chi-square difference between the original and method-factor models was significant, indicating that CMV is unlikely to inflate the observed relationships. These multiple diagnostics confirm that common method variance does not pose a substantial threat to the validity of this study's findings.

4.1. Hypothesis Examination

We used Hayes' PROCESS macro to perform the mediation analyses and test the hypothesized relationships. The results revealed both direct and indirect effects among the key constructs, offering new insights into the drivers of trust in AI recommendations.

First, Perceived Transparency of AI (PTAI) had a significant positive effect on both Perceived Fairness of AI (PFAI) (Coefficient = 0.3249, $P < 0.001$) and Perceived Control over AI (PCAI) (Coefficient = 0.3306, $P < 0.001$). However, the direct effect of PTAI on Trust in AI Recommendations (TAIR) was not significant (Coefficient = 0.0281, $P = 0.596$). This suggests that transparency builds trust indirectly, primarily through fairness and control mechanisms.

Among the mediators, PFAI significantly enhanced TAIR (Coefficient = 0.2590, $P < 0.001$), while PCAI did not (Coefficient = -0.0632, $P = 0.222$). These results indicate that fairness plays a more important role in shaping trust than control.

We also examined the indirect pathways. The mediation through PFAI (PTAI → PFAI → TAIR) was significant (Effect = 0.1157, BootLLCI = 0.0520, BootULCI = 0.1901), reinforcing the central role of fairness in fostering trust. In contrast, the mediation through PCAI (PTAI → PCAI → TAIR) was not significant (Effect = -0.0203, BootLLCI = -0.0942, BootULCI = 0.0408), suggesting that control perceptions do not strongly mediate the transparency–trust link.

The moderation analysis further revealed that Privacy Concerns of AI influenced how fairness affects trust. Specifically, the interaction between PFAI and Privacy Concerns was significant (Coefficient = 0.2215, $P = 0.0395$). This indicates that users with higher privacy concerns are more sensitive to fairness when forming trust. However, privacy concerns did not significantly moderate the effects of either PTAI (P = 0.112) or PCAI (P = 0.975) on TAIR.

Lastly, the overall models showed moderate explanatory power, with R^2 values of 0.1448 for PFAI, 0.1675 for PCAI, and 0.2098 for TAIR. These findings highlight the importance of fairness and privacy concerns in building trust in AI healthcare applications. Transparency plays a key role, but its effect is mostly indirect, working through users' perceptions of fairness.

'Table 3 Here'

5. Discussion and conclusion

AI-powered applications and platforms are rapidly being used in medical and healthcare settings, providing users with unique solutions for monitoring and improving their health. These applications have transformed healthcare service delivery by offering individualized, data-driven suggestions and improving access to medical information. Despite their increasing prominence, trust remains a significant factor influencing user adoption, particularly in high-risk healthcare settings. While previous research has recognized the importance of transparency, fairness, and user control in building trust in AI systems, their interplay and combined influence on trust development remains unexplored, particularly in healthcare applications. Furthermore, privacy concerns have received little attention as a moderating element in this association. Therefore, this study aims to examine how AI transparency influences trust in AI recommendations in healthcare, with fairness and control as mediators, and privacy concerns as a moderating factor, providing deeper insights into the mechanisms shaping trust in AI-driven health applications.

The findings reveal key insights into the role of AI transparency in shaping user trust and perceptions in healthcare applications. First, while PTAI did not have a direct effect on TAIR, it significantly influenced PFAI and PCAI. This indicates that transparency alone may not suffice to build trust but works indirectly by enhancing users' perceptions of fairness and control. These results suggest that in highly sensitive contexts such as healthcare, where users deal with personal health data and critical decisions, trust is shaped by a combination of transparency and users' confidence in the fairness and manageability of the system.

The lack of a direct relationship between PTAI and TAIR might also stem from the healthcare context itself, where users tend to scrutinize recommendations more carefully due to the high stakes involved. Prior studies, such as those by Shin (2021) and Shin and Park (2019), have highlighted transparency as a critical driver of trust, but our findings suggest that in healthcare, users may demand additional reassurances, such as fairness and control, to develop trust in AI systems.

Furthermore, the significant influence of PTAI on both PFAI and PCAI aligns with procedural justice theory (Thibaut, 1975), which highlights the importance of transparent processes in cultivating perceptions of fairness and empowerment. This study extends previous work by demonstrating that the indirect effect of transparency—via fairness and control—is particularly important in healthcare, unlike in domains such as e-commerce or entertainment where trust may be more easily gained. These findings underscore the importance of context-sensitive approaches when designing AI systems for healthcare. Ethical considerations, user autonomy, and perceived procedural fairness all play critical roles in shaping how transparency affects trust in this domain.

The findings highlight the distinct roles of perceived fairness (PFAI) and perceived control (PCAI) in shaping trust in AI recommendations (TAIR). Specifically, PFAI demonstrated a significant positive effect on TAIR, underscoring the importance of fairness perceptions in building trust. When users perceive AI systems as fair, unbiased, and equitable in their decision-making processes, they are more likely to trust the recommendations provided. In contrast, PCAI did not show a significant effect on TAIR. This suggests that users' sense of control over AI systems may not directly contribute to trust in the same way that fairness does. One plausible explanation lies in the nature of the healthcare context. Given the sensitivity of medical decisions, users may prioritize fairness an ethical dimension over control, which is more functional or interactive in nature. Despite its insignificance, the PCAI → TAIR path was retained in the model to preserve theoretical alignment with the Theory of Planned Behavior (Ajzen, 1991), which highlights perceived behavioral control as a driver of intention and trust. The non-significant effect in our context may reflect domain-specific dynamics where ethical and fairness considerations outweigh perceived autonomy. Future research could explore this pathway further in less sensitive or more interactive domains.

Prior studies, such as those by Binns et al. (2018) and Felzmann et al. (2020), have similarly emphasized that fairness is a critical factor in fostering trust in AI systems, particularly in domains where decisions carry significant personal implications. On the other hand, the lack of significance for PCAI could stem from the assumption that users in healthcare settings often rely on the expertise of systems rather than seeking full autonomy or control, reducing the direct impact of control perceptions on trust. Moreover, users may not fully understand how to exert control over advanced AI functionalities, especially when technical literacy is limited. This may lead to a situation where control is either undervalued or even perceived as burdensome.

These findings further suggest that the relative importance of fairness and control may vary across different contexts. For instance, in less sensitive applications, such as e-commerce, control perceptions might play a more prominent role, as users value flexibility and decision-making autonomy. However, in healthcare, where ethical considerations and accuracy are paramount, fairness perceptions become more influential. This distinction highlights the need for context-specific strategies in designing AI systems, ensuring that they address the factors most critical to fostering trust in their intended use cases.

The study also tested the indirect effects of PTAI on TAIR through PFAI and PCAI, providing further insights into how transparency influences trust. Results showed that the pathway through PFAI was significant, reinforcing the critical role of fairness in building trust by addressing users' concerns about equity and ethics. However, the pathway through PCAI was not significant, suggesting that control perceptions are less influential in healthcare, where users prioritize fairness and transparency over personal agency. These findings highlight the context-specific nature of trust-building factors, with fairness taking precedence in sensitive environments like healthcare.

This study examined the moderating role of Privacy Concerns (PC) in the relationship between PTAI, PFAI, PCAI, and TAIR, yielding both expected and unexpected findings. The results demonstrated that PC significantly moderated the relationship between PFAI and TAIR, highlighting that privacy concerns amplify the impact of fairness perceptions on trust. However, the moderation effects of PC on PTAI → TAIR and PCAI → TAIR pathways were not significant, suggesting that the influence of transparency and control on trust is less sensitive to users' privacy concerns in the context of healthcare applications.

The significant moderation in the PFAI → TAIR pathway aligns with prior research, such as Bansal et al. (2015) and Schomakers et al. (2019), which emphasize that privacy concerns heighten users' scrutiny of fairness in data-driven systems. When fairness is perceived as strong, privacy-conscious users are more likely to trust the system, viewing it as a safeguard against potential misuse of their sensitive health data. This finding also resonates with Martin et al.

(2017), who argue that trust is bolstered when systems demonstrate both ethical behavior and respect for data privacy. In healthcare, fairness becomes a critical reassurance, particularly for users who are already anxious about the handling of their personal medical information.

In contrast, the lack of significant moderation in the PTAI → TAIR and PCAI → TAIR pathways suggests that privacy concerns do not meaningfully amplify the effects of transparency or control on trust. This divergence from findings like those of Suh and Han (2003), which emphasize privacy concerns as a core component of trust in digital platforms, could be attributed to the inherent complexity of healthcare applications. Unlike in contexts like e-commerce, where transparency and control are more directly linked to user satisfaction, healthcare users may expect fairness and security as baseline attributes, diminishing the relative importance of transparency and control when privacy concerns are high.

Additionally, these results may reflect the nuanced relationship between user perceptions and AI systems. Users with heightened privacy concerns may view control and transparency as insufficient to address their anxieties about data misuse. As suggested by Dinev and Hart (2006), privacy concerns often extend beyond visible system features, reflecting broader apprehensions about third-party access and systemic vulnerabilities. This could explain why fairness, which is more closely associated with ethical data handling, plays a stronger role in building trust under these conditions.

Comparatively, the findings align with Galiana et al. (2024) and Fehr et al. (2024) in highlighting the critical role of privacy and fairness in fostering trust in healthcare technologies. However, they contrast with studies which suggest that transparency universally mitigates privacy concerns by enhancing user understanding of data processes. This discrepancy may point to the context-specific nature of trust dynamics, where the sensitive nature of healthcare data shifts the balance of priorities toward fairness. These insights suggest that privacy concerns act as a lens through which users evaluate the ethical and procedural aspects of AI systems, with fairness playing a more prominent role than transparency or control. Future research should investigate whether similar patterns emerge in other high-stakes domains, such as finance, or whether privacy concerns exhibit different moderating effects in less sensitive contexts.

5.1. Theoretical implications

The results of this study contribute to the existing literature on trust in AI systems within the healthcare domain by providing a nuanced understanding of how fairness, control, and privacy concerns interact with transparency to shape user trust (Abbas and Khan, 2015, Bernal and Mazo, 2022, Castelo et al., 2019, Fehr et al., 2024, Kiseleva et al., 2022, Bach et al., 2024). Specifically, the findings provide new theoretical insights into the role of fairness, control, and privacy concerns in shaping trust in AI recommendations, particularly in the sensitive context of healthcare. While transparency is widely regarded as a fundamental driver of trust in AI, this research highlights that its influence is largely mediated through fairness perceptions, rather than exerting a direct effect. This nuanced finding challenges prior assumptions in the literature that transparency alone suffices to build trust and enriches theoretical understanding by showing that fairness acts as a critical intermediary mechanism. Furthermore, the insignificant role of control as a mediator suggests that in high-stakes environments, such as healthcare, users may deprioritize their sense of agency in favor of ethical assurances provided by fairness and privacy protections. These findings compel researchers to reexamine the relative importance of fairness and control in trust-building frameworks, particularly in contexts involving sensitive decision-making and personal data.

The study further contributes to the theoretical discourse by uncovering the moderating role of privacy concerns in trust dynamics. Specifically, the findings reveal that privacy concerns amplify the impact of fairness on trust, underscoring the interconnected nature of ethical perceptions and user anxieties about data security. This expands the scope of existing theories, such as procedural justice theory (Thibaut, 1975) , by integrating the role of privacy concerns as a contextual factor that shapes user evaluations of fairness and its subsequent influence on trust. In contrast, the absence of significant moderation effects in the relationships involving control and transparency suggests that privacy concerns selectively influence certain pathways, indicating that user priorities in trust formation are not uniform but highly dependent on the perceived ethical stakes of the context. These findings challenge existing trust models by introducing a layered perspective that emphasizes the situational interplay between fairness, control, and privacy concerns, offering a more robust framework for understanding trust in high-stakes AI applications.

This research also addresses a critical gap in the literature by demonstrating the limitations of perceived control in mediating trust in AI-driven healthcare applications. While previous studies have positioned control as a central factor in user engagement with digital systems, this study reveals that its role diminishes in contexts where ethical considerations and fairness dominate user priorities. This finding challenges general trust-building models that assign equal weight to control across all domains and calls for a more context-sensitive approach in theoretical frameworks. By showing that perceived control has a limited impact in healthcare, the study prompts researchers to rethink the applicability of traditional constructs in high stakes, ethically charged environments, where users may place greater emphasis on fairness and privacy protections over their own agency.

5.2. Practical implications

The findings of this study offer valuable practical insights for developers, designers, and policymakers working with AI systems in healthcare. First, the results underscore the critical role of fairness in building trust, suggesting that developers should prioritize designing AI systems that not only provide transparent outputs but also ensure that their decision-making processes are perceived as fair and unbiased. This could involve implementing mechanisms that clearly explain how decisions are made and demonstrating that these decisions are equitable across diverse user groups. For instance, fairness audits and bias detection tools could be integrated into healthcare AI platforms to proactively address user concerns and enhance trust. Additionally, embedding transparent AI interfaces that communicate decision logic in simple and user-friendly language can further improve users' understanding and confidence in the system.

Second, the limited role of perceived control in trust formation highlights the importance of focusing efforts on ethical assurances rather than excessive user autonomy in high-stakes contexts. Developers should ensure that healthcare AI systems are designed to reduce complexity and emphasize reliability, as users in these environments may prioritize confidence in the system’s ethical standards over their own control over the system. Simplified interfaces that guide users through decision-making processes while maintaining transparency and fairness may prove more effective than overly customizable features. It is also important that design choices are aligned with users’ cognitive and emotional comfort, particularly for populations with lower digital literacy or limited exposure to AI technologies.

Finally, the significant moderation effect of privacy concerns on fairness and trust emphasizes the need for robust data protection mechanisms to address users' anxieties about data security. Policymakers and organizations should establish and communicate clear privacy standards, ensuring that users are confident about how their sensitive health data is collected, stored, and used. In addition to strong technical safeguards, clear and proactive communication strategies such as in-app notifications, privacy dashboards, or consent summaries can help inform users about data practices and build a sense of control and transparency. Educating users about these protections and maintaining transparency in privacy practices can further mitigate concerns and foster greater adoption of healthcare AI systems. Furthermore, developers should adopt best practices in data protection, including end-to-end encryption, data anonymization, and secure storage protocols. Aligning healthcare AI applications with internationally recognized data protection frameworks such as the Health Insurance Portability and Accountability Act (HIPAA), the General Data Protection Regulation (GDPR), and the ISO/IEC 27799 standard for health informatics privacy can provide a solid foundation for ethical and legal compliance. Ensuring meaningful informed consent and empowering users with control over their personal data will be essential in fostering long-term trust and responsible adoption of AI in healthcare.

These practical implications highlight the importance of balancing fairness, ethical assurances, and privacy protections to build trust in healthcare AI systems, offering actionable insights for industry stakeholders aiming to optimize user engagement and satisfaction. By translating these findings into actionable strategies such as transparency-centered design, fairness audits, and user-centric privacy communication developers, healthcare professionals, and policymakers can more effectively implement responsible and trustworthy AI-driven health technologies.

5.3. Limitations and further research

A number of limitations arise from the present study. First, the proposed model was tested with a specific sample in a specific context, and it therefore could be limiting to generalize into other domains or populations. Second, the present study focused on the perceptual factors-transparency, fairness, and control-leaving aside other important variables such as user experience or emotional responses that may contribute considerably to the formation of trust in AI. Additionally, this study relied solely on self-reported data, which introduces potential sources of bias such as social desirability bias, recall bias, and acquiescence bias. Although a deviation question was included to improve response attentiveness, we acknowledge that this alone may not fully mitigate these issues. Future research could enhance measurement reliability by incorporating reverse-coded items to detect inconsistent patterns, triangulating with behavioral data such as app usage logs or system interactions, and applying cross-method validation. These steps would strengthen the robustness of findings derived from perceptual constructs.

Third, the cross-sectional design does not allow for an examination of changes in user trust over time. Future research may attempt to overcome these limitations by examining different contexts and further embedding other factors, such as emotional influences, into their longitudinal designs to more accurately reflect the changing nature of trust in AI. Finally, as this study found that PCAI did not significantly impact TAIR, future research could examine this relationship in different contexts, such as e-commerce, where user autonomy may play a larger role. Additionally, cultural and regional factors could influence how users value control, with some cultures emphasizing autonomy more than others. Investigating these variations could provide deeper insights into the role of control in fostering trust and help tailor AI systems to meet diverse user expectations across contexts. Finally, this study employed convenience sampling and did not account for important contextual variables such as user experience (e.g., comfort with AI interfaces, digital literacy), emotional responses (e.g., anxiety, confidence in technology), and cultural and social influences (especially given the Iranian context), which may restrict the generalizability of the findings due to potential selection bias. Future studies would benefit from adopting more rigorous sampling methods, such as stratified or random sampling, to reach a broader and more diverse user base. Including underrepresented groups such as elderly users, individuals with limited technological experience, and patients with chronic health conditions while also examining emotional and cultural dimensions, would offer a richer and more comprehensive understanding of how trust in AI recommendations is shaped across various demographic and contextual settings.

Conflict of interest

The authors declare no conflict of interest.

Source of funding

The authors declare no funding received for this study.

Acknowledgements

We confirm that the manuscript has been primarily created by the authors, and all major sections, including the research design, data analysis, and key arguments, reflect the authors' original work. It is worth noting that minimal assistance was sought from AI tools or Large Language Models (LLMs) in specific areas, such as grammar refinement and formatting. All such outputs were carefully reviewed, validated, and revised by the authors to ensure their accuracy, reliability, and alignment with the academic standards of the manuscript. The authors take full responsibility for the integrity and originality of the manuscript, ensuring that it represents their intellectual contributions and adheres to the ethical guidelines of the journal.

References:

Abbas, A. and Khan, S. U. (2015), "E-health cloud: privacy concerns and mitigation strategies", *Medical Data Privacy Handbook*, pp. 389-421. https://doi.org/10.1007/978-3-319-23633-9_15

Abouelmehdi, K., Beni-Hessane, A. and Khaloufi, H. (2018), "Big healthcare data: preserving security and privacy", *Journal of big data*, Vol. 5 No. 1, pp. 1-18. https://doi.org/10.1186/s40537-017-0110-7

Adams, K. (2024), "Transparency in Health Care AI: A Conversation with Experts", Bipartisan Policy Center.

Ahmed, Z., Mohamed, K., Zeeshan, S. and Dong, X. (2020), "Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine", *Database*, Vol. 2020, p. baaa010. https://doi.org/10.1093/database/baaa010
Ajzen, I. (1991), "The Theory of planned behavior", *Organizational Behavior and Human Decision Processes*. https://doi.org/10.4135/9781446249215.n22
Anderson, J. C. and Gerbing, D. W. (1988), "Structural equation modeling in practice: A review and recommended two-step approach", *Psychological bulletin*, Vol. 103 No. 3, p. 411. https://doi.org/10.1037/0033-2909.103.3.411
Awotunde, J. B., Jimoh, R. G., Folorunso, S. O., Adeniyi, E. A., Abiodun, K. M. and Banjo, O. O. (2021), "Privacy and security concerns in IoT-based healthcare systems", *The fusion of internet of things, artificial intelligence, and cloud computing in health care*, Springer, pp. 105-134. https://doi.org/10.1007/978-3-030-75220-0_6
Bach, T. A., Khan, A., Hallock, H., Beltrão, G. and Sousa, S. (2024), "A systematic literature review of user trust in AI-enabled systems: An HCI perspective", *International Journal of Human–Computer Interaction*, Vol. 40 No. 5, pp. 1251-1266. https://doi.org/10.1080/10447318.2022.2138826
Bagozzi, R. P. and Yi, Y. (1988), "On the evaluation of structural equation models", *Journal of the academy of marketing science*, Vol. 16, pp. 74-94. https://doi.org/10.1007/BF02723327
Bansal, G., Zahedi, F. M. and Gefen, D. (2015), "The role of privacy assurance mechanisms in building trust and the moderating role of privacy concern", *European Journal of Information Systems,* Vol. 24 No. 6, pp. 624-644. https://doi.org/10.1057/ejis.2014.41
Belanche, D., Guinalíu, M. and Albás, P. (2022), "Customer adoption of p2p mobile payment systems: The role of perceived risk", *Telematics and Informatics*, Vol. 72, p. 101851. https://doi.org/10.1016/j.tele.2022.101851
Bernal, J. and Mazo, C. (2022), "Transparency of artificial intelligence in healthcare: insights from professionals in computing and healthcare worldwide", *Applied Sciences,* Vol. 12 No. 20, p. 10228. https://doi.org/10.3390/app122010228
Bhatt, U., Xiang, A., Sharma, S., Weller, A., Taly, A., Jia, Y., Ghosh, J., Puri, R., Moura, J. M. and Eckersley, P. (2020), "Explainable machine learning in deployment", in *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pp. 648-657. https://doi.org/10.1145/3351095.3375624
Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J. and Shadbolt, N. (2018), "'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions", in *Proceedings of the 2018 Chi conference on human factors in computing systems*, pp. 1-14. https://doi.org/10.1145/3173574.3173951
Castelo, N., Bos, M. W. and Lehmann, D. R. (2019), "Task-dependent algorithm aversion", *Journal of Marketing Research,* Vol. 56 No. 5, pp. 809-825. https://doi.org/10.1177/0022243719851788
Chaddad, A., Peng, J., Xu, J. and Bouridane, A. (2023), "Survey of explainable AI techniques in healthcare", *Sensors,* Vol. 23 No. 2, p. 634. https://doi.org/10.3390/s23020634
Chellappa, R. K. (2008), "Consumers' trust in electronic commerce transactions: The role of perceived privacy and perceived security", *under submission,* Vol. 13.
Chen, M. and Decary, M. (2020), "Artificial intelligence in healthcare: An essential guide for health leaders", in *Healthcare management forum,* Vol. 33, pp. 10-18. https://doi.org/10.1177/0840470419873123

Chenthara, S., Ahmed, K., Wang, H. and Whittaker, F. (2019), "Security and privacy-preserving challenges of e-health solutions in cloud computing", *IEEE access,* Vol. 7, pp. 74361-74382. https://doi.org/10.1109/ACCESS.2019.2919982
Culnan, M. J. and Armstrong, P. K. (1999), "Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation", *Organization science,* Vol. 10 No. 1, pp. 104-115. https://doi.org/10.1287/orsc.10.1.104
De Bellis, E. and Johar, G. V. (2020), "Autonomous shopping systems: Identifying and overcoming barriers to consumer adoption", *Journal of Retailing,* Vol. 96 No. 1, pp. 74-87. https://doi.org/10.1016/j.jretai.2019.12.004
Dinev, T. and Hart, P. (2006), "Privacy concerns and levels of information exchange: An empirical investigation of intended e-services use", *E-service,* Vol. 4 No. 3, pp. 25-60. https://doi.org/10.2979/esj.2006.4.3.25
Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., Scott, K., Schieber, S., Waldo, J. and Weinberger, D. (2017), "Accountability of AI under the law: The role of explanation", *arXiv preprint arXiv:1711.01134*. https://doi.org/10.48550/arXiv.1711.01134
Faul, F., Erdfelder, E., Buchner, A. and Lang, A.-G. (2009), "Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses", *Behavior research methods,* Vol. 41 No. 4, pp. 1149-1160. https://doi.org/10.3758/BRM.41.4.1149
Fehr, J., Citro, B., Malpani, R., Lippert, C. and Madai, V. I. (2024), "A trustworthy AI reality-check: the lack of transparency of artificial intelligence products in healthcare", *Frontiers in Digital Health,* Vol. 6, p. 1267290. https://doi.org/10.3389/fdgth.2024.1267290
Felzmann, H., Fosch-Villaronga, E., Lutz, C. and Tamò-Larrieux, A. (2020), "Towards transparency by design for artificial intelligence", *Science and engineering ethics,* Vol. 26 No. 6, pp. 3333-3361. https://doi.org/10.1007/s11948-020-00276-4
Fornell, C. and Larcker, D. F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of marketing research,* Vol. 18 No. 1, pp. 39-50. https://doi.org/10.1177/002224378101800104
Foroudi, P. and Dennis, C. (2023), *Researching and analysing business: Research methods in practice,* Taylor & Francis
Galiana, L. I., Gudino, L. C. and González, P. M. (2024), "Ethics and artificial intelligence", *Revista Clínica Española (English Edition)*. https://doi.org/10.1016/j.rceng.2024.02.003
Garver, M. S. and Mentzer, J. T. (1999), "Logistics research methods: employing structural equation modeling to test for construct validity", *Journal of business logistics,* Vol. 20 No. 1, p. 33.
Gefen, D. and Straub, D. (2003), "Managing user trust in B2C e-services", *e-Service,* Vol. 2 No. 2, pp. 7-24. https://doi.org/10.2979/esj.2003.2.2.7
Hair, J. (2009), "Multivariate data analysis", *Exploratory factor analysis*.
Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. and Tatham, R. L. (2006), *Multivariate Data Analysis: Pearson Prentice Hall,* Pearson Prentice Hall, Upper Saddle River, NJ
Hair Jr, J. F., Black, W. C., Babin, B. J. and Anderson, R. E. (2010), "Multivariate data analysis", *Multivariate data analysis*, pp. 785-785.
Karatas, M., Eriskin, L., Deveci, M., Pamucar, D. and Garg, H. (2022), "Big Data for Healthcare Industry 4.0: Applications, challenges and future perspectives", *Expert Systems with Applications,* Vol. 200, p. 116912. https://doi.org/10.1016/j.eswa.2022.116912

Kim, Y., Kim, S. H., Peterson, R. A. and Choi, J. (2023), "Privacy concern and its consequences: A meta-analysis", *Technological Forecasting and Social Change*, Vol. 196, p. 122789. https://doi.org/10.1016/j.techfore.2023.122789

Kiseleva, A., Kotzinos, D. and De Hert, P. (2022), "Transparency of AI in healthcare as a multilayered system of accountabilities: between legal requirements and technical limitations", *Frontiers in artificial intelligence*, Vol. 5, p. 879603. https://doi.org/10.3389/frai.2022.879603

Lambert, D. M. and Harrington, T. C. (1990), "Measuring nonresponse bias in customer service mail surveys", *Journal of business Logistics*, Vol. 11 No. 2, pp. 5-25.

Li, W. and Zheng, X. (2024), "Social Media Use and Attitudes toward AI: The Mediating Roles of Perceived AI Fairness and Threat", *Human Behavior and Emerging Technologies*, Vol. 2024 No. 1, p. 3448083. https://doi.org/10.1155/2024/3448083

Liao, Q. V. and Vaughan, J. W. (2023), "Ai transparency in the age of llms: A human-centered research roadmap", *arXiv preprint arXiv:2306.01941*, pp. 5368-5393. https://doi.org/10.1162/99608f92.8036d03b

Loh, H. W., Ooi, C. P., Seoni, S., Barua, P. D., Molinari, F. and Acharya, U. R. (2022), "Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022)", *Computer methods and programs in biomedicine*, Vol. 226, p. 107161. https://doi.org/10.1016/j.cmpb.2022.107161

Longoni, C., Bonezzi, A. and Morewedge, C. K. (2019), "Resistance to medical artificial intelligence", *Journal of Consumer Research*, Vol. 46 No. 4, pp. 629-650. https://doi.org/10.1093/jcr/ucz013

Malhotra, N. K., Kim, S. S. and Patil, A. (2006), "Common method variance in IS research: A comparison of alternative approaches and a reanalysis of past research", *Management science*, Vol. 52 No. 12, pp. 1865-1883.

Martin, K. D., Borah, A. and Palmatier, R. W. (2017), "Data privacy: Effects on customer and firm performance", *Journal of marketing*, Vol. 81 No. 1, pp. 36-58. https://doi.org/10.1509/jm.15.0497

Maseeh, H. I., Jebarajakirthy, C., Pentecost, R., Arli, D., Weaven, S. and Ashaduzzaman, M. (2021), "Privacy concerns in e□commerce: A multilevel meta□analysis", *Psychology & Marketing*, Vol. 38 No. 10, pp. 1779-1798. https://doi.org/10.1002/mar.21493

Mayer, R. C., Davis, J. H. and Schoorman, F. D. (1995), "An integrative model of organizational trust", *Academy of management review*, Vol. 20 No. 3, pp. 709-734.

Memarian, B. and Doleck, T. (2023), "Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI), and higher education: A systematic review", *Computers and Education: Artificial Intelligence*, p. 100152. https://doi.org/10.1016/j.caeai.2023.100152

Nunnally, J. C. (1978), "An overview of psychological measurement", *Clinical diagnosis of mental disorders: A handbook*, pp. 97-146. https://doi.org/10.1007/978-1-4684-2490-4_4

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. and Podsakoff, N. P. (2003), "Common method biases in behavioral research: a critical review of the literature and recommended remedies", *Journal of applied psychology*, Vol. 88 No. 5, p. 879.

Promberger, M. and Baron, J. (2006), "Do patients trust computers?", *Journal of Behavioral Decision Making*, Vol. 19 No. 5, pp. 455-468. https://doi.org/10.1002/bdm.542

Rohden, S. F. and Espartel, L. B. (2024), "Consumer reactions to technology in retail: choice uncertainty and reduced perceived control in decisions assisted by recommendation

agents", *Electronic Commerce Research*, pp. 1-23. https://doi.org/10.1007/s10660-024-09808-7
Samal, L., Fu, H. N., Camara, D. S., Wang, J., Bierman, A. S. and Dorr, D. A. (2021), "Health information technology to improve care for people with multiple chronic conditions", *Health services research*, Vol. 56, pp. 1006-1036. https://doi.org/10.1111/1475-6773.13860
Schomakers, E.-M., Lidynia, C. and Ziefle, M. (2019), "Listen to my heart? How privacy concerns shape users' acceptance of e-health technologies", in *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 306-311. https://doi.org/10.1109/WiMOB.2019.8923448
Schork, N. J. (2019), "Artificial intelligence and personalized medicine", *Precision medicine in Cancer therapy*, pp. 265-283. https://doi.org/10.1007/978-3-030-16391-4_11
Shick, A. A., Webber, C. M., Kiarashi, N., Weinberg, J. P., Deoras, A., Petrick, N., Saha, A. and Diamond, M. C. (2024), "Transparency of artificial intelligence/machine learning-enabled medical devices", *NPJ Digital Medicine*, Vol. 7 No. 1, p. 21. https://doi.org/10.1038/s41746-023-00992-8
Shin, D. (2021), "The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI", *International journal of human-computer studies*, Vol. 146, p. 102551. https://doi.org/10.1016/j.ijhcs.2020.102551
Shin, D. and Park, Y. J. (2019), "Role of fairness, accountability, and transparency in algorithmic affordance", *Computers in Human Behavior*, Vol. 98, pp. 277-284. https://doi.org/10.1016/j.chb.2019.04.019
Siala, H. and Wang, Y. (2022), "SHIFTing artificial intelligence to be responsible in healthcare: A systematic review", *Social Science & Medicine*, Vol. 296, p. 114782. https://doi.org/10.1016/j.socscimed.2022.114782
Singhal, A., Neveditsin, N., Tanveer, H. and Mago, V. (2024), "Toward fairness, accountability, transparency, and ethics in AI for social media and health care: scoping review", *JMIR Medical Informatics*, Vol. 12 No. 1, p. e50048. https://doi.org/10.2196/50048
Smith, H. J., Milberg, S. J. and Burke, S. J. (1996), "Information privacy: Measuring individuals' concerns about organizational practices", *MIS quarterly*, pp. 167-196. https://doi.org/10.2307/249477
Steenkamp, J.-B. E. and Van Trijp, H. C. (1991), "The use of LISREL in validating marketing constructs", *International Journal of Research in marketing*, Vol. 8 No. 4, pp. 283-299.
Suh, B. and Han, I. (2003), "The impact of customer trust and perception of security control on the acceptance of electronic commerce", *International Journal of electronic commerce*, Vol. 7 No. 3, pp. 135-161. https://doi.org/10.1080/10864415.2003.11044270
Tabachnick, B. (2007), "Experimental designs using ANOVA", *Thomson/Brooks/Cole*.
Tan, S.-K., Chong, S.-C., Wee, K.-K. and Chong, L.-Y. (2024), "Personalized Healthcare: A Comprehensive Approach for Symptom Diagnosis and Hospital Recommendations Using AI and Location Services", *Journal of Informatics and Web Engineering*, Vol. 3 No. 1, pp. 117-135. https://doi.org/10.33093/jiwe.2024.3.1.8
Thibaut, J. (1975), "Procedural justice: A psychological analysis", *Lawrance Erlbaum Associates*.
Topol, E. (2019), *Deep medicine: how artificial intelligence can make healthcare human again*, Hachette UK

Tucker, M., Jubb, C. and Yap, C. J. (2020), "The theory of planned behaviour and student banking in Australia", *International Journal of Bank Marketing*, Vol. 38 No. 1, pp. 113-137. https://doi.org/10.1108/IJBM-11-2018-0324

WHO (2021), *Ethics and governance of artificial intelligence for health: WHO guidance: executive summary*, World Health Organization, Switzerland

Xie, Q., Faust, K., Van Ommeren, R., Sheikh, A., Djuric, U. and Diamandis, P. (2019), "Deep learning for image analysis: Personalizing medicine closer to the point of care", *Critical Reviews in Clinical Laboratory Sciences*, Vol. 56 No. 1, pp. 61-73. https://doi.org/10.1080/10408363.2018.1536111

Age	Frequency	Percent	Status	Frequency	Percent
Male	228	56.2	Single	239	58.9
Female	178	43.8	Married	160	39.4
Age of respondents			Divorced	7	1.7
18-24	142	35.0	**Use the App**		
25-34	120	29.6	Less than 6 Months	184	45.3
35-44	92	22.7	Between 6 Months and 1 Yea	54	13.3
4.00	29	7.1	Between 1 year and 2 Years	65	16.0
45-54	23	5.7	More than 2 years	103	25.4
Education			Use Per Day		
Diploma	117	28.8	Once per day	237	58.4
Undergraduate	125	30.8	Twice per day	103	25.4
Postgraduate and above	164	40.4	Three times per day	40	9.9
			More than three times per day	26	6.4

Table 1- Demographic profile (No 406)

Source: Author's own creation

Variable name and Items		Factor loading	Mean	Std. Deviation	Cronbach alpha
Perceived Transparency of AI (Shin, 2021)					**.805**
	I think the evaluation criteria used by the AI system in [health platform/app name] is easily accessible and understandable to me (Understandability).	.779	5.3719	1.29657	
	Any outputs produced by the AI system in [health platform/app name] is explainable to me, especially when they affect my health decisions (Explainability)	.854	5.3645	1.19882	
	I think the AI system in [platform/app name] allow me to understand how well its internal processes can be inferred from its external outputs (Observability)	.795	5.3325	1.22334	
Perceived Fairness of AI (Shin, 2021)					**.755**
	I feel that the AI system in [health platform/app name] treats me fairly, without favoritism or discrimination against me (Nondiscrimination)	.798	5.8202	1.30135	
	I believe that the AI system in [health platform/app name] should make the sources of its data and algorithms clear to me, and ensure they are accurately logged and benchmarked (Accuracy).	.730	5.7291	1.25334	
	I think that the AI system in [platform/app name] follows fair and impartial processes, ensuring that I am not subjected to any prejudice (Due process).	.764	5.6749	1.28148	
Privacy Concern of AI (Chellappa, 2008)					**.815**
	I feel confident that I know all the parties involved in collecting the information I provide to the AI system in [health platform/app name].	.804	4.3596	1.71378	
	I am aware of the exact nature of the information that will be collected from me by the AI system in [health platform/app name].	.808	4.6576	1.59584	
	I know what information I need to provide during my interaction with the AI system in [health platform/app name]	.756	4.9852	1.59158	
	I believe I have control over how the information I provide is used by the AI system in [health platform/app name].	.639	4.7414	1.44858	
Trust in AI Recommendation (Shin, 2021)					**.886**
	I trust the recommendations provided by the AI system in [health platform/app name]	.764	5.3547	1.15569	
	I believe that the recommendations made through the AI system in [health platform/app name] are trustworthy.	.794	5.2291	1.21075	
	I believe that the recommendations generated by the AI system in [health platform/app name] are reliable.	.787	5.2956	1.13151	
Perceived Control Over AI (Belanche et al., 2022)					**.756**
	When I use the AI system in [health platform/app name], I feel that I have control over the actions I take.	.860	5.2808	1.29766	
	The use of the AI system in [health platform/app name] is under my control.	.884	5.1724	1.32211	
	When using the AI system in [health platform/app name], I do not feel confused.	.876	5.3916	1.21587	

Table 2- The domain and items of construct in extent literature, factor loadings, descriptive statistics and reliabilities

Items sources: Shin (2021), Belanche et al. (2022), Chellappa (2008)

Effect	Coeff	SE	t	P	LLCI	ULCI	Sig/Insignificant
Perceived Transparency of AI → Perceived Fairness of AI	0.3249	0.0506	6.4219	0.0000	0.2254	0.4243	Significant
Perceived Transparency of AI → Trust in AI Recommendation	0.0281	0.0529	0.5311	0.5956	-0.0759	0.1320	Insignificant
Perceived Transparency of AI → Perceived Control Over AI	0.3306	0.0491	6.7342	0.0000	0.2341	0.4271	Significant
Perceived Fairness of AI → Trust in AI Recommendation	0.2590	0.0535	4.8362	0.0000	0.1537	0.3643	Significant
Perceived Control Over AI → Trust in AI Recommendation	-0.0632	0.0516	-1.2240	0.2217	-0.1647	0.0383	Insignificant
Interaction							
Fairness of AI × Privacy Concerns of AI → Trust in AI Recommendation	0.2215	0.1072	2.0661	0.0395	0.0107	0.4323	Significant
Perceived Transparency of AI × Privacy Concerns of AI → Trust in AI Recommendation	-0.1612	0.1011	-1.5948	0.1116	-0.3600	0.0375	Insignificant
Perceived Control Over AI × Privacy Concerns of AI → Trust in AI Recommendation	-0.0032	0.1016	-0.0312	0.9751	-0.2030	0.1966	Insignificant
Direct Effect	**Effect**	**BootSE**	**BootLLCI**		**BootULCI**		
Perceived Transparency of AI → Perceived Fairness of AI → Trust in AI Recommendation	0.1157	0.0351	0.0520		0.1901		Significant
Perceived Transparency of AI → Perceived Control Over AI → Trust in AI Recommendation	-0.0203	0.0348	-0.0942		0.0408		Insignificant

Table 3- Results of hypothesis testing
Source: Author's own creation

	CR	AVE	Trust in AI Recommendation	Transparency of AI	Fairness of AI	Privacy Concerns of AI	Control Over AI
Trust in AI Recommendation	0.887	0.723	**0.850**				
Transparency of AI	0.813	0.593	0.262	**0.770**			
Fairness of AI	0.757	0.510	0.475	0.461	**0.714**		
Privacy Concerns of AI	0.819	0.532	0.400	0.421	0.551	**0.730**	
Control Over AI	0.757	0.510	0.157	0.471	0.420	0.524	**0.714**
CR (Composite Reliability) Threshold ≥ 0.70 (Hair et al., 2010; AVE (Average Variance Extracted) Threshold ≥ 0.50 (Foroudi and Dennis, 2023, Fornell and Larcker, 1981)							

Appendix 1: Discriminant validity, AVE and construct reliability
Source: Author's own creation