

**Patient Trust of an AI Medical Provider:
Investigating the impact of Competence, Benevolence and Integrity**

Abstract

Artificial Intelligence (AI) in healthcare has introduced disruptive advances with benefits and many unknown, sometimes negative issues echoing calls for additional research in this domain (Gómez-González & Gómez, 2020). This disruption has the capacity to impact millions of patients by changing the way medicine is practiced, drastically altering the patient-provider relationship. Our study therefore aims to understand the interactions between trust, loyalty and satisfaction in a patient-AI Medical provider relationship in order to identify key challenges that will be crucial in the design, implementation and utilization of this technology to benefit society.

Introduction

Artificial intelligence (AI), the ability of digital devices to perform tasks usually associated with humans is revolutionizing the healthcare industry providing many opportunities from reducing the burden of administrative tasks to increasing the accuracy of diagnosis and treatment of disease. This revolution is evident in the proliferation of healthcare applications that have evolved from very basic but effective symptom checking devices to increasingly autonomous robotic physicians and surgeons (González & Gutiérrez, 2020; Britt, 2018), wearables and mobile devices for tracking healthy habits (Raghuveer, 2018), Talkspace (Miner, Milstein & Hancock, 2017) mental-health care to name a few. Despite these benefits that AI is bringing to healthcare, researchers (González & Gutiérrez, 2020; Loh, 2018) have raised concerns about its potential to disrupt the way medicine is practiced as AI is projected to drastically alter the patient-provider relationship, a critical component for healthcare outcomes. Patient trust, loyalty and satisfaction are fundamental elements of this relationship (Pearson & Raeke, 2000; Platonova, Kennedy, & Shewchuk, 2008). We believe that these are the same elements that will shape patient-AI

medical provider interactions in healthcare. In our study we focus on patients for the following reasons.

(1) Patients are the largest group of healthcare consumers, and therefore patients will drive the successful implementation and usage of AI medical providers in healthcare and there is very limited research that examines interactions with AI medical providers from patient's perspectives (Longoni et al, 2019). (2) In primary care, the nature of patient-provider interaction has a major impact on healthcare outcomes. Given the crucial role of this relationship in improving quality of care, a deeper understanding of the patient–AI medical provider relationship and the factors that impact it is needed. (3) And finally, questions about the social effects and impacts of AI systems on humans usually arise after the technology has been deployed. AI-mediated technologies in healthcare usually become available, at very different stages and degrees of implementation. This study, therefore, aims to understand the interactions between trust, loyalty, and satisfaction in a patient-AI medical provider relationship in order to identify key challenges that will be crucial in the design, implementation, and utilization of this technology to benefit society.

Background and Research Model

Patient trust in their medical provider is highly correlated with patient satisfaction and loyalty as it leads to a stronger patient-doctor relationship and better wellness outcomes (Platonova, Kennedy & Shewchuk, 2008). Trust is, therefore, an essential element in the choice of medical provider. Trust is a key construct with multiple dimensions that may have an indirect but powerful effect on user interaction with the technology. It is therefore important to understand the role of trust in the increasingly complex relationship between the patient and AI Medical provider in order to implement effective strategies in designing and successfully implementing the technology. Prior research on trust in technology has mainly focused on technology-mediated interactions with less attention on user-technology interactions (Benbasat & Wang 2005; Söllner, Hoffmann, Hoffmann, Wacker, & Leimeister, 2012; Tripp, McKnight, & Lankton, 2011). Although some of these findings can be transferred to AI medical systems, these systems pose unique challenges compared to other technical domains due to their characteristics and the

nature of their interaction with humans. A framework for modelling trust in human computer interactions, which considers the human like attributes of competence, benevolence, integrity that are unique characteristics of AI is based on the premise that if technology has human-like characteristics, then human like attributes can be used to study trust in the technology (Tripp, McKnight, & Lankton, 2011; Gulati, Sousa & Lamas, 2017). In our study, we build on these prior works and propose a new framework to evaluate the impact of the human cognitive aspects of trust identified by Gulati et.al (2018) on the interaction between trust, loyalty and satisfaction in the patient-AI medical provider relationship. In trust research competence is conceptualized as the user's perception on how well the technology performs a function or provides system features the user needs to interact with the technology (Gulati et al. 2018). In the medical literature, patients have indicated that physician competence is an important characteristic to them. Therefore, we propose that the relationship between perceived competence associated with an AI medical provider and user trust will have a significant impact on the patient-AI medical provider interactions. Benevolence is the disposition to act in good faith. Gulati et al. (2017) proposed that when individuals perceive that technology would help them and act in their best interest, they were more likely to continue to use and trust the technology. Therefore, the relationship between perceived benevolence associated with AI medical provider and user trust will impact the patient-AI provider relationship. Integrity is attributed to the quality of honesty. In technology, integrity is understood to be a user's perception of the technology to properly manage and protect personal information (Gulati et al., 2017; Sousa, Lamas & Dias, 2014). Our study extends previous models (Gulati, S., Sousa, S., & Lamas, D., 2018) studying trust relationships in human-like technologies (HLT) to include some aspects of the patient-primary care physician model (Platonova et al., 2008) which has been used extensively in medicine to examine the patient-provider interactions. In their research Gulati et al., (2018) identified competence, benevolence, and honesty as important constructs in a human computer trust model (HCTM) to study user computer interactions. We adapt this model in our study to examine the impact of the elements of competence, benevolence and integrity as antecedents of trust, to get a better understanding

of patient-AI medical provider relationship interactions. The current study is guided by theories and techniques from previous work but adapts different aspects of the studies and proposes a new model shown in Figure 1 to study the patient-AI medical provider relationship and poses the following question: Research Question (RQ): To what extent do the elements of competence, benevolence, integrity impact the interactions between patient trust, satisfaction and loyalty in the patient-AI medical provider relationship?

Theoretical Model

Patient Trust has been identified as an important predictor of patient satisfaction and loyalty in a patient-human medical provider relationship. We believe that patient trust is a critical construct that will also shape patient-AI interactions in healthcare. Trust is viewed as a set of specific beliefs dealing with benevolence, competence, integrity, and predictability or a combination of these elements (Siau & Wang, 2018). However, the role of these antecedents of trust in the patient-AI provider interactions from a patient perspective has yet to be determined. It is against this setting that we present the following hypotheses:

H1: There is a positive relationship between patient satisfaction associated with an AI medical provider and patient loyalty.

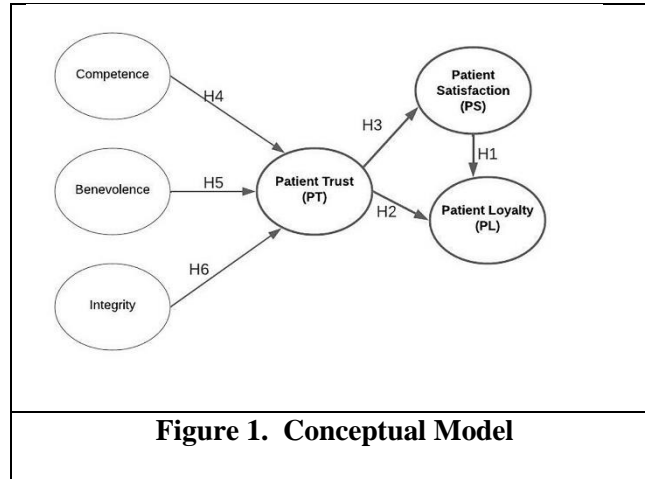
H2: There is a positive relationship between patient trust associated with an AI medical provider and patient loyalty

H3: There is a positive relationship between patient trust associated with an AI medical provider and patient satisfaction

H4: There exists a positive relationship between perceived competence associated with an AI medical provider and user trust.

H5: There exists a positive relationship between perceived benevolence associated with an AI medical provider and user trust.

H6: There exists a positive relationship between Integrity and trust in patient-medical AI interactions



Proposed Methodology

To explore this research question, we plan carry out a two-part study starting with a qualitative analysis of current theories and empirical results from the literature related to the variables depicted in our model.

Measures for the six variables in our model will be drawn from research in healthcare and academic literature on human computer interaction with minor modifications to adapt the measures for the patient–Medical AI in the context of mental Health and well-being. The second part will explore the interactions between these variables through an experimental vignette study comparing participants’ perceptions of using medical-AI technology versus human provider for assessing and diagnosing mental state.

Conclusion

The proposed study is very important in many ways. This study will add to our understanding of the human-machine interaction by examining the linkage in patient satisfaction, patient trust, and patient loyalty and exploring key antecedents to these critical variables in a patient-AI medical provider relationship. An understanding of the nature of interactions between medical AI and patients will help in designing these systems and will ultimately lead to systems with optimal AI-human collaboration.

References and Citations

Benbasat, I., & Wang, W. (2005). Trust in and adoption of online recommendation agents. *Journal of the association for information systems*, 6(3), 4.

Britt, P. (2018). How AI-Assisted Surgery Is Improving Surgical Outcomes. Retrieved from <https://www.roboticsbusinessreview.com/health-medical/ai-assisted-surgery-improves-patient-outcomes/>

Gómez-González, Emilio, and Emilia Gómez Gutiérrez. "Artificial Intelligence in Medicine and Healthcare: applications, availability and societal impact." (2020).

Gulati, S., Sousa, S., & Lamas, D. (2017, September). Modelling trust: An empirical assessment. In *IFIP Conference on Human-Computer Interaction* (pp. 40-61). Springer, Cham.

Gulati, S., Sousa, S., & Lamas, D. (2018, December). Modelling trust in human-like technologies. In *Proceedings of the 9th Indian conference on human computer interaction* (pp. 1-10).

Loh, E. (2018). Medicine and the rise of the robots: a qualitative review of recent advances of artificial intelligence in health. *BMJ Leader*, leader-2018.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4), 629-650.

Miner, A. S., Milstein, A., & Hancock, J. T. (2017). Talking to machines about personal mental health problems. *Jama*, 318(13), 1217-1218.

Pearson, S. D., & Raeke, L. H. (2000). Patients' trust in physicians: many theories, few measures, and little data. *Journal of general internal medicine*, 15(7), 509-513.

Platonova, E. A., Kennedy, K. N., & Shewchuk, R. M. (2008). Understanding patient satisfaction, trust, and loyalty to primary care physicians. *Medical Care Research and Review*, 65(6), 696-712.

Raghunandan, M. (2018). 5 Real-World Examples of AI in Healthcare. Retrieved from <https://www.kolabtree.com/blog/5-real-world-examples-of-ai-in-healthcare/>

Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter business technology journal*, 31(2), 47-53.

Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., & Leimeister, J. M. (2012, December). Understanding the formation of trust in IT artifacts. Association for Information Systems.

Sousa, S., Lamas, D., & Dias, P. (2014, June). A model for Human-computer Trust. In *International Conference on Learning and Collaboration Technologies* (pp. 128-137). Springer, Cham.

Tripp, J., McKnight, H., & Lankton, N. K. (2011). Degrees of humanness in technology: what type of trust matters?