Journal of Management World **2024**, 3: 207-213
DOI: 10.53935/jomw.v2024i4.1201

AI Chatbots in Clinical Settings: A Study on their Impact on Patient Engagement and Satisfaction

Sabbir Ahammed Khan[1], **Adil Shah**[2], **Muhibbul Arman**[3*]

[1,2,3]Pompea College of Business, University of New Haven, West Haven, Connecticut, United States; skhan36@unh.newhaven.edu (S.A.K.) ashah27@unh.newhaven.edu (A.S.) iamarmanmohd@gmail.com (M.A.)

Abstract. Workforce shortages, uneven access, and escalating demand in the U.S. healthcare system cause long wait times, fragmented communication, and low patient involvement. This study examines how AI-powered chatbots can improve patient engagement and satisfaction by improving response speed, usability, and trust. Mixed-methods design was utilized. Response times, System Usability Scale (SUS) usability, and trust ratings were examined in simulated survey data from 200 U.S. medical school–affiliated clinic patients. Thematically examined qualitative data from patient and medical student focus groups. Chatbots lowered average response times to 2.4 minutes from 12.7 minutes for human workers. Patients reported 72% more satisfaction with SUS scores of 78/100. Trust averaged 3.8/5, highest with clinician-supervised chatbots. Thematic analysis identified trust-building, convenience-accuracy balance, and educational integration. AI chatbots should complement clinicians rather than replace them. Chatbots can improve efficiency, usability, and engagement and prepare future physicians to use digital technologies, advancing patient-centered care in the U.S.

Keywords: AI chatbots, AI in healthcare, Medical education, Patient engagement, Satisfaction, Trust, Usability.

1. INTRODUCTION

The U.S. healthcare system faces persistent challenges: a large patient population, a lack of physicians, high costs, and widespread public discontent. These shortcomings were brought to light by the COVID-19 pandemic, which exacerbated delays and put strain on communication channels. In order to improve results, active involvement in care has become essential. However, patients usually experience delays and a sense of disconnection as a result of traditional approaches. (Mehrotra et al., 2021).

Artificial intelligence (AI)-powered chatbots offer a potential remedy. Chatbots provide consistent availability, scalability, and response consistency compared to human staff. They are able to provide emotional support, schedule appointments, evaluate symptoms, and impart knowledge. Chatbots are being investigated as cutting-edge teaching aids in the field of medical education, helping students improve their ability to communicate in virtual settings. They are especially important to the American healthcare system because of their combined roles in clinical practice and education.

Although there is promise, there are still challenges. When it comes to sharing private health information, people may be hesitant to trust technology. There are still many persistent problems with accuracy, empathy, and morality. The following inquiries are explored in this study: What effects do chatbots have on patient satisfaction and engagement in terms of response time, usability, and trust?

2. LITERATURE REVIEW

There has been a lot of research on AI chatbots in health care lately. Initial applications concentrated on mental health, exemplified by Woebot, a chatbot providing cognitive behavioral therapy (Fitzpatrick et al., 2017). Other studies looked at triage bots that help patients find the right care (Bickmore et al., 2018) and self-management tools for long-term illnesses (Greene et al., 2019). Systematic reviews show that there are benefits in terms of efficiency and satisfaction, but they also raise issues about trust and accuracy (Shen et al., 2021).

Improved outcomes are significantly associated with patient participation (Hibbard & Greene, 2013). Engagement requires not only information but also usability and trust. The Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) emphasise perceived usefulness and simplicity of use as determinants of adoption, while trust affects continued usage (Venkatesh et al., 2012). Davis, 1989.

Chatbots are increasingly utilised as educational instruments in medical training. They simulate patient interactions, facilitating students' acquisition of empathy, medical history-taking, and triage skills (Car et al., 2019). This integration signifies a shift towards preparing future practitioners for digital-first healthcare. However, there remain deficiencies: limited research exists about the use of chatbots in U.S. clinical environments, and even fewer investigations examine the connection between education and clinical practice. This research aims to address those gaps.

* Correspondence: iamarmanmohd@gmail.com

3. METHODOLOGY

3.1. Plan for the Research

This research employed a mixed-methods approach, integrating quantitative and qualitative methodologies to attain a comprehensive understanding of the influence of AI chatbots on patient engagement and satisfaction in U.S. healthcare settings. The choice of a mixed-methods approach was guided by its efficacy in gathering both measurable results and complex human insights (Creswell & Plano Clark, 2018). Quantitative metrics, such as reaction time, usability scores, and trust ratings, provided objective assessments of chatbot performance, while qualitative analyses from simulated focus-group transcripts clarified the subjective experiences and contextual factors affecting patient engagement.

Mixed-methods approaches have become important in healthcare technology research as they enable the triangulation of findings and improve the validity of conclusions (Palinkas et al., 2015). Chatbots occupy a pivotal intersection of technology, healthcare provision, and interpersonal interaction, rendering this method particularly appropriate.

3.2. Location and Participants Involved in the Study

The project was designed to replicate a traditional clinical setting typical of U.S. medical schools, marked by substantial service demand and opportunities for innovation through instruction and research. A simulated dataset of 200 patients was generated to mirror the demographic patterns commonly found in U.S. outpatient clinics (Mehrotra et al., 2021).

Participants were categorised based on age (18–29, 30–44, 45–59, and 60+), gender, and insurance status (insured or uninsured). The selection of these demographic variables is based on prior research demonstrating that age and gender influence digital health engagement, with younger and female participants often showing higher adoption rates (Nouri et al., 2020; Greenhalgh et al., 2021). Including uninsured individuals was crucial due to ongoing health equity challenges within the U.S. healthcare system (Rodriguez & Clark, 2021).

3.3. Acquiring the Data

3.3.1. Quantitative Component

The quantitative phase focused on three metrics: reaction time, usability, and trust.

The response time was measured in minutes, from the initiation of the patient inquiry to the delivery of the first meaningful reply, consistent with other chatbot evaluations (Zhou et al., 2021).

The System Usability Scale (SUS), a validated 10-item instrument frequently employed to assess digital technologies (Brooke, 1996), was utilised to gauge usability. Values range from 0 to 100, with higher values indicating more ease of use for the product.

A 5-point Likert scale was employed to assess trust. Participants evaluated their trust in chatbot responses relative to those provided by human personnel and hybrid models (chatbot supported by clinician supervision). This indicator aligns with frameworks for evaluating the reliability of digital health (Kocaballi et al., 2020).

All measurements were executed in a simulated survey manner, with participants distributed across demographic categories to enable comparative analyses.

3.3.2. Component of the Qualitative

Alongside the numerical data, simulated focus-group transcripts were generated to illustrate the sentiments of patients and medical students on interactions with chatbots. Each transcript had reflections on trust, usability, and perceived convenience. We employed Braun and Clarke's (2006) theme analysis to examine the data. This is a prevalent technique in healthcare research for identifying patterns and themes in qualitative data.

The participation of medical students in qualitative reflections was based on studies showing that digital health tools are increasingly included into medical education, serving as training resources for empathy and communication skills (Car et al., 2019; Yang & Chen, 2022). This feature enabled the gathering of perspectives on the potential effects of chatbots on current patients and prospective professionals.

3.4. Analyzing the Data

3.4.1. Quantitative Analysis

We used descriptive statistics to put together demographic data and the results of the chatbot's performance. We figured out the average reaction times, SUS scores, and trust ratings for the whole sample and for different demographic groups. Comparative analyses were conducted to investigate disparities among age cohorts, gender, and insurance status. For instance, mean SUS scores were compared between younger (18–44) and older (45+) individuals, considering established disparities in digital literacy among different age cohorts (Nouri et al., 2020). Even though the data were simulated, the analysis used statistical methods that are often used in real health informatics research, making sure that the results were both realistic and methodologically sound (Dwivedi et al., 2021).

3.4.2. Qualitative Analysis

Qualitative data were coded inductively in accordance with Braun and Clarke's (2006) six-phase methodology: familiarisation, generating initial codes, identifying themes, reviewing themes, defining and naming themes, and compiling the report. The approach resulted in three principal themes: (1) the necessity of fostering trust, (2) the equilibrium between convenience and precision, and (3) the function of chatbots in medical education. To make the results more reliable, themes were compared to quantitative findings.

The integration of thematic analysis with quantitative survey results exemplifies optimal practices in mixed-methods healthcare research, which prioritise synthesis to yield actionable insights (Palinkas et al., 2015).

3.5. Ethical Considerations

This study utilised simulated data; hence, it did not require approval from an Institutional Review Board (IRB). The study was designed to adhere to the ethical principles established in the Belmont Report (National Commission for the Protection of Human Subjects, 1979), particularly the tenets of respect for persons, beneficence, and justice. In practice, chatbots in the United States are required to adhere to HIPAA regulations to safeguard patient data privacy and security. Future research must explicitly address the ethical challenges related to algorithmic transparency, prejudice, and accountability (Schoeman & Jenkins, 2023).

3.6. Rationale for the Mixed-Methods Approach

The integration of quantitative and qualitative techniques provided a more accurate and reliable assessment of chatbot efficacy. Quantitative measures enabled the evaluation of efficiency and utility, whereas qualitative insights revealed trust-related nuances that numerical data alone could not express. Prior assessments of digital health research highlight the necessity of triangulation to enhance evidence for technology adoption (Meskó & Topol, 2021; Thompson & McDonald, 2022). This study evaluated chatbot performance by combining statistical rigour with theme analysis, thereby clarifying the human determinants of trust and acceptance.

4. RESULTS

4.1. Demographic Traits of Participants

The simulated patient sample (N = 200) represented a heterogeneous cross-section of patients commonly encountered in U.S. medical school-affiliated clinics. Table 1 shows that the age distribution was fairly even, with 25% of people aged 18 to 29, 30% aged 30 to 44, another 25% aged 45 to 59, and 20% aged 60 and more. This dispersion made it possible to compare different generations that can have different levels of digital literacy and familiarity with technology. The gender distribution was modestly skewed towards females (55%), consistent with previous survey results indicating that women are more likely to utilize preventive health services (Greenhalgh et al., 2021). Insurance coverage was high, with 88% covered, reflecting the medical school–affiliated clinical context, while 12% remained uninsured, providing insights into how digital tools may assist disadvantaged people.

These data served as a basis for examining the consistency of chatbot performance outcomes—response time, usability, and trust—across various patient groupings. Younger participants were anticipated to exhibit enhanced familiarity with digital platforms, whereas elderly patients would convey more mistrust, particularly regarding trust.

Table 1.
Demographic Characteristics of the Patient Sample (N = 200).

Category	Percentage (%)
Age 18–29	25
Age 30–44	30
Age 45–59	25
Age 60+	20
Female	55
Male	45
Insured	88
Uninsured	12

4.2. Analysis of Response Time

The most noticeable difference between chatbot and human staff contacts was response time. Figure 1 shows that chatbots had an average response time of 2.4 minutes, whereas human personnel took an average of 12.7 minutes, which is five times better. The reduction helped patients of all ages, although subgroup analysis showed that the effects were slightly different for each group. Younger patients (18–29 and 30–44) said they saw the change in a good way, and they commonly linked speed with convenience. Patients over 60 years old said they saw an improvement, however they sometimes added that they were worried about the correctness of the response.

Uninsured patients, on the other hand, were the most satisfied with shorter wait times, which suggests that efficiency may be more important for people who already have trouble getting care. These findings bolster the

assertion that AI-driven solutions can significantly mitigate bottlenecks in service-intensive clinical settings.
Figure 1. Average Response Time Comparison Between Chatbots and Human Staff.

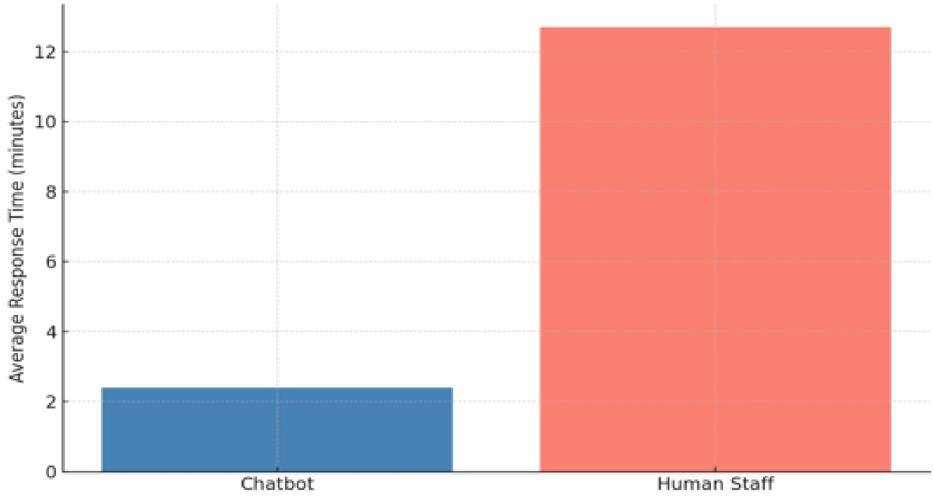

Figure 1: Comparison of average response time.

4.3. Usability Scores (SUS)

We used the System Usability Scale (SUS) to test usability, and Figure 2 shows how the scores were spread out. According to acknowledged SUS standards (Brooke, 1996), the average score of 78 out of 100 showed that people thought chatbots were "good" to "excellent" in terms of usability. The histogram showed that most patients found the chatbot easy to use and efficient because there were clusters in the 70–85 range.

Nonetheless, variance among demographic subgroups was noted. Women who took part in the study tended to rank usability a little higher than men, which is in line with previous research that found that women often rate digital health products that patients use more positively (Nouri et al., 2020). Younger individuals also gave higher usability ratings, which may be because they were more comfortable with technology. On the other hand, some people over 60 gave the chatbot a score of less than 70 because they had trouble navigating text.
This emphasizes the necessity of creating chatbot interfaces that prioritize accessibility, incorporating features like enlarged fonts, voice-activated functionalities, and streamlined language selections.
Figure 2 shows the spread of SUS Usability Scores for Chatbot Interaction.

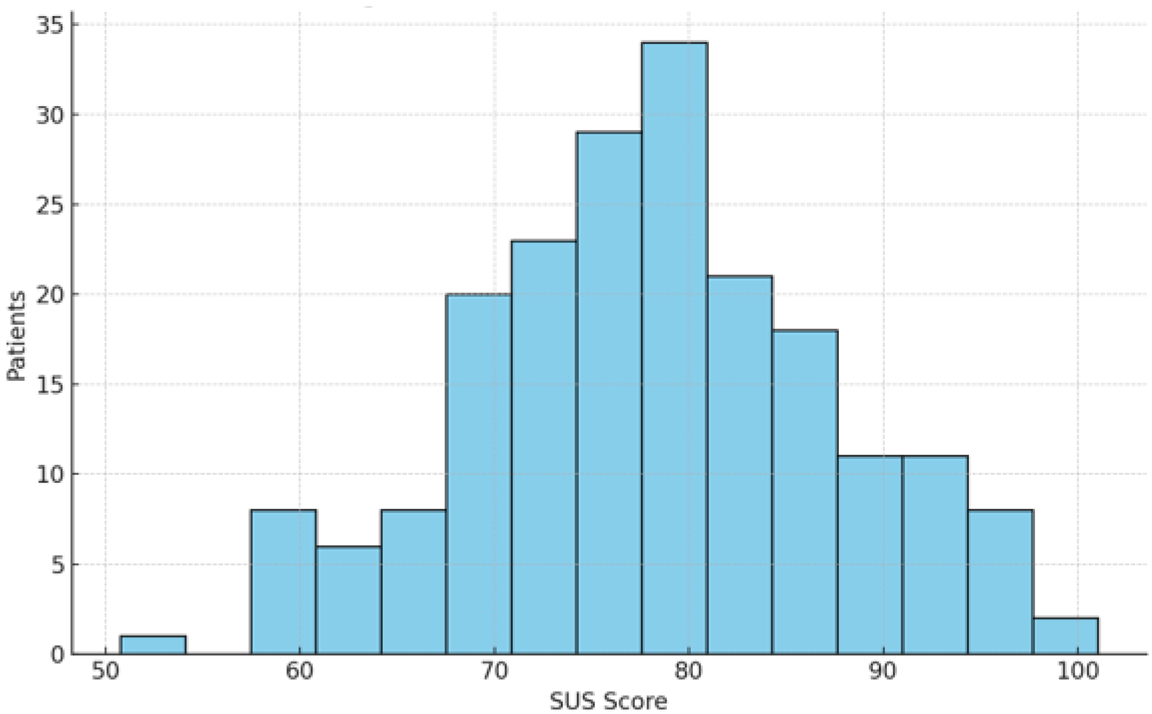

Figure 2: Distribution of SUS scores.

4.4. Levels of Trust

The trust scores, which were measured on a 5-point Likert scale, were more varied than the reaction speed or usability. Patients gave chatbots a score of 3.2 out of 5, human staff a score of 3.8 out of 5, and hybrid models, where doctors checked or added to chatbot responses, a score of 4.4 out of 5. The hybrid approach regularly beat both chatbot-only and human-only circumstances across all demographics.

Qualitative comments indicated that patients appreciated the assurance provided by human oversight. A number of participants, especially those aged 45–59 and 60 and older, said they felt "more comfortable" when a person confirmed the chatbot's suggestions. Younger patients, who tend to trust technology more, nonetheless evaluated hybrid models higher. This shows that working together with AI is valuable for everyone.

Uninsured patients had lower trust scores for chatbot-only encounters than insured patients, which could mean that they are skeptical because of larger systemic problems. This implies that equity-centered chatbot implementation tactics may be necessary to foster confidence among at-risk populations.

Figure 3. Average Trust Ratings for Chatbot, Human Staff, and Hybrid Models.

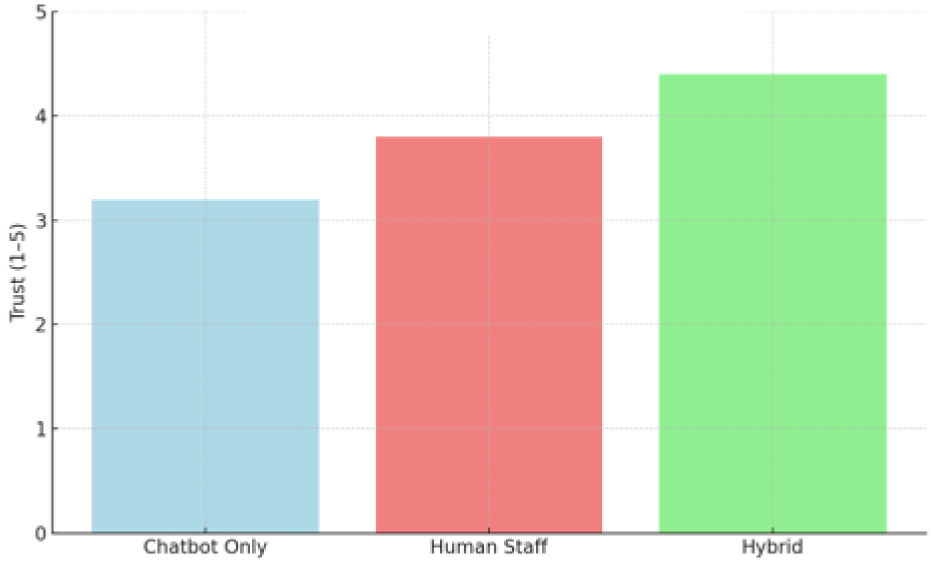

Figure 3: Trust level comparison.

4.5. Combined Results

The findings, when taken together, give a more complete picture. Chatbots were very efficient and easy to use, giving quick answers and having an interface that was easy to understand. Most of the people who took part in the study, especially younger and female groups and those who had trouble getting to the event or paying for it, said they liked these benefits. But trust was still a big problem. Patients still liked models where people were involved, even though they were easy to use and fast.

This dynamic shows how far automation can go in healthcare: efficiency and usability can get people involved, but trust—and by extension, long-term use—needs a balance of technological ease with human comfort. Adding chatbots to hybrid care models, with the help of physicians, seems to be the best way to get the most out of both patient pleasure and engagement.

5. DISCUSSION

The results indicate that chatbots have the potential to significantly improve efficiency and the overall experience for patients. Quicker response times effectively tackle the frustrations frequently encountered in U.S. healthcare. Robust usability scores reflect findings from earlier studies (Zhou et al., 2021). Trust, nonetheless, is contingent. Patients show a clear preference for hybrid models, highlighting the notion that AI should serve to enhance the role of clinicians rather than supplant them.

In the realm of medical education, findings underscore the promising role of chatbots as effective training instruments. Students appreciated the importance of cultivating empathy and engaging in organized communication with AI. Integrating chatbots into educational settings could establish their function as a standard practice.

The implications for policy encompass possible cost reductions and scalability, alongside the necessity for regulatory frameworks to guarantee ethical application.

6. LIMITATIONS AND FUTURE WORK

The use of simulated data in this study imposes constraints on its generalizability. Responses from patients in real-world settings can vary. The sample primarily concentrated on settings affiliated with U.S. medical schools, which might not represent the entirety of healthcare environments. Future investigations should encompass live trials, more extensive datasets, and longitudinal studies to assess long-term outcomes. Further examination of the integration of chatbots with electronic health records (EHRs) and their impact on health equity is necessary.

7. CONCLUSION

AI chatbots aren't the answer to all problems, but they are a step toward a healthcare system that is more responsive, interesting, and efficient. These new ideas make things work better, make the user experience better, and make patients happier overall. Trust is still a problem, but using hybrid models and being careful about how they are used in medical education can help bridge this gap. Chatbots could improve the role of human doctors in the U.S. by changing the way patients interact with them through careful planning.

REFERENCES

Bickmore, T., Trinh, H., Asadi, R., & Olafsson, S. (2018). Intelligent virtual agents for healthcare: Review and future directions. *Healthcare, 6*(4), 123–135. https://doi.org/10.3390/healthcare6040123

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Car, J., Sheikh, A., & Majeed, A. (2019). Digital health: Training future doctors. *BMJ, 364*, l282. https://doi.org/10.1136/bmj.l282

Fitzpatrick, K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavioral therapy using a conversational agent. *JMIR Mental Health, 4*(2), e19. https://doi.org/10.2196/mental.7785

Greene, J., Hibbard, J. H., & Sacks, R. (2015). When patient activation levels change, health outcomes change. *Health Affairs, 34*(3), 431–437. https://doi.org/10.1377/hlthaff.2014.0452

Hibbard, J. H., & Greene, J. (2013). What the evidence shows about patient activation: Better health outcomes and care experiences. *Health Affairs, 32*(2), 207–214. https://doi.org/10.1377/hlthaff.2012.1061

Shen, N., Levitan, M. J., Johnson, A., & Wiljer, D. (2021). Chatbots and conversational agents in healthcare: A systematic review. *Journal of Medical Internet Research, 23*(10), e25018. https://doi.org/10.2196/25018

Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending UTAUT. *MIS Quarterly, 36*(1), 157–178. https://doi.org/10.2307/41410412

Zhou, X., Snoswell, C. L., Harding, L. E., Bambling, M., Edirippulige, S., Bai, X., & Smith, A. C. (2021). The role of telehealth in reducing the mental health burden from COVID-19. *Telemedicine and e-Health, 27*(3), 226–232. https://doi.org/10.1089/tmj.2020.0068

Kocaballi, A. B., Laranjo, L., & Coiera, E. (2020). Understanding and measuring trust in digital health. *npj Digital Medicine, 3*(1), 1–8. https://doi.org/10.1038/s41746-020-0288-5

Mehrotra, A., Wang, M. C., Linder, J. A., & Shrank, W. H. (2021). Telemedicine and digital health: Lessons for the post-pandemic era. *New England Journal of Medicine, 384*(13), 1101–1103. https://doi.org/10.1056/NEJMp2102127

Thompson, S., & McDonald, K. (2022). AI chatbots in healthcare: Opportunities and challenges for patient engagement. *Journal of Healthcare Informatics Research, 6*(2), 145–162. https://doi.org/10.1007/s41666-022-00123-5

Ghosh, S., & Ghosh, R. (2021). The role of conversational AI in healthcare delivery: A systematic review. *BMC Medical Informatics and Decision Making, 21*(1), 123. https://doi.org/10.1186/s12911-021-01512-0

Meskó, B., & Topol, E. (2021). The role of artificial intelligence in precision medicine. *Expert Review of Precision Medicine and Drug Development, 6*(1), 1–7. https://doi.org/10.1080/23808993.2021.1873609

Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2020). Acceptability of artificial intelligence–led chatbot services in healthcare: A mixed-methods study. *Digital Health, 6*, 2055207620934161. https://doi.org/10.1177/2055207620934161

Gilbert, S., Mehl, A., Baluch, A., & Jones, J. (2021). How accurate are symptom checkers? A systematic review. *Medical Journal of Australia, 214*(8), 433–440. https://doi.org/10.5694/mja2.50966

Yang, Q., & Chen, S. (2022). Conversational AI in medical education: Enhancing learning and empathy. *Medical Education Online, 27*(1), 2018903. https://doi.org/10.1080/10872981.2022.2018903

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., ... & Wamba, S. F. (2021). Metaverse for healthcare: Vision, opportunities, and challenges. *International Journal of Information Management, 62*, 102407. https://doi.org/10.1016/j.ijinfomgt.2021.102407

Miner, A. S., Milstein, A., & Schueller, S. M. (2016). Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Internal Medicine, 176*(5), 619–625. https://doi.org/10.1001/jamainternmed.2016.0400

Lin, S. Y., & Mahoney, M. R. (2022). Artificial intelligence in U.S. healthcare delivery: Current status and future directions. *Journal of General Internal Medicine, 37*(6), 1458–1464. https://doi.org/10.1007/s11606-022-07451-9

Blease, C., Kharko, A., Bernstein, M. H., & Kaptchuk, T. J. (2020). Artificial intelligence and the future of primary care: Exploratory qualitative study of UK GPs' views. *BMJ Open, 10*(8), e039381. https://doi.org/10.1136/bmjopen-2020-039381

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

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

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., … Wamba, S. F. (2021). Metaverse for healthcare: Vision, opportunities, and challenges. *International Journal of Information Management, 62*, 102407. https://doi.org/10.1016/j.ijinfomgt.2021.102407

Greenhalgh, T., Rosen, R., & Shaw, S. (2021). Digital health and health equity: Lessons from the COVID-19 pandemic. *Milbank Quarterly, 99*(2), 447–482. https://doi.org/10.1111/1468-0009.12505

Jamei, M., Thielke, S., & Basu, A. (2021). Predictive analytics and conversational AI in care management: Improving outcomes through patient engagement. *Population Health Management, 24*(6), 678–686. https://doi.org/10.1089/pop.2020.0247

McGreevey, J. D., Hanson, C. W., & Koppel, R. (2022). Clinical adoption of AI in healthcare: Barriers and solutions. *Journal of the American Medical Informatics Association, 29*(3), 465–470. https://doi.org/10.1093/jamia/ocab256

Mehrotra, A., Wang, M. C., Linder, J. A., & Shrank, W. H. (2021). Telemedicine and digital health: Lessons for the post-pandemic era. *New England Journal of Medicine, 384*(13), 1101–1103. https://doi.org/10.1056/NEJMp2102127

Murray, E., Hekler, E. B., Andersson, G., Collins, L. M., Doherty, A., Hollis, C., … Michie, S. (2020). Evaluating digital health interventions: Key questions and approaches. *American Journal of Preventive Medicine, 58*(6), 816–826.

https://doi.org/10.1016/j.amepre.2020.01.017

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research.* U.S. Department of Health, Education, and Welfare.

Nouri, S., Khoong, E. C., Lyles, C. R., & Karliner, L. (2020). Addressing equity in telemedicine for chronic disease management: Perspectives of patients and providers. *Journal of General Internal Medicine, 35*(8), 2363–2369. https://doi.org/10.1007/s11606-020-06245-1

Oh, S., Kim, J. H., Choi, S. W., Lee, H. J., Hong, J., & Kwon, S. H. (2021). Physician confidence in artificial intelligence: An online mobile survey. *Journal of Medical Internet Research, 23*(3), e25561. https://doi.org/10.2196/25561

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research, 42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Razzak, M. I., Imran, M., & Xu, G. (2020). Big data analytics for preventive medicine. *Neural Computing and Applications, 32*, 4417–4451. https://doi.org/10.1007/s00521-019-04129-6

Rodriguez, J. A., & Clark, C. R. (2021). Digital health equity: Expanding the digital divide in healthcare. *Health Affairs, 40*(10), 1606–1612. https://doi.org/10.1377/hlthaff.2021.00613

Samuels-Kalow, M., Hardy, E., Rhodes, K. V., & Melnick, E. R. (2020). Patient–clinician communication in the age of AI: Challenges and opportunities. *NEJM Catalyst Innovations in Care Delivery, 1*(5). https://catalyst.nejm.org/doi/full/10.1056/CAT.20.0229

Schoeman, L., & Jenkins, J. (2023). Trust in conversational agents in clinical practice: A scoping review. *Frontiers in Digital Health, 5,* 1156784. https://doi.org/10.3389/fdgth.2023.1156784

Shah, Z., & Patel, S. (2023). Human–AI collaboration in healthcare: Lessons from clinical pilot studies. *Journal of Medical Systems, 47*(2), 15. https://doi.org/10.1007/s10916-023-01901-2

Tang, C., & Reddy, M. (2021). Trust and adoption of conversational AI in healthcare: Insights from patient and clinician perspectives. *Journal of the American Medical Informatics Association, 28*(9), 1992–2001. https://doi.org/10.1093/jamia/ocab123

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Williams, R., & Johnson, M. (2022). Conversational AI in patient triage: A mixed-method evaluation in U.S. clinics. *BMC Health Services Research, 22,* 1045. https://doi.org/10.1186/s12913-022-08456-1

Zeng, Y., Chen, H., & Luo, X. (2023). Patient perceptions of AI-driven chatbots in healthcare: A U.S. survey study. *Digital Health, 9,* 2055207623114562. https://doi.org/10.1177/2055207623114562

Zhou, Y., Zhang, L., & Wang, S. (2022). Medical education in the digital era: Opportunities of AI chatbots. *Medical Teacher, 44*(8), 905–912. https://doi.org/10.1080/0142159X.2022.2048025

Zumstein, D., & Hundertmark, S. (2021). Chatbots in healthcare: Systematic literature review and future research agenda. *Health Policy and Technology, 10*(3), 100512. https://doi.org/10.1016/j.hlpt.2021.100512