The Artificial Intelligence Dilemma: Navigating Ethics in Healthcare

Elen Mullaj
Bachelor of Computer Science, Artifical Intelligence Specialization & Bachelor of Science

ABSTRACT

The integration of Artificial Intelligence (AI) into healthcare represents a transformative shift towards more accurate and efficient patient care, highlighted by advancements in disease detection, treatment adherence, and patient interaction. However, this technological evolution introduces significant ethical dilemmas, including concerns over empathetic care provision, data security, and bias in clinical decision-making. This literature review critically examines these ethical challenges and proposes potential solutions to ensure AI's beneficial integration into healthcare without compromising patient well-being. The exploration into AI's capacity for empathetic care reveals its limitations to cognitive empathy, suggesting a complementary role to human providers rather than a replacement. Data privacy concerns underscore the imperative for secure handling and consent mechanisms in the utilization of patient information, amidst the risk of misuse by large corporations. Furthermore, the review addresses AI-induced biases, advocating for diverse data representation and algorithmic transparency to mitigate discrimination and enhance treatment efficacy across varied populations. While acknowledging the potential of AI to revolutionize healthcare, this paper advocates for a cautious and ethically informed approach to its adoption, emphasizing the need for comprehensive legislation, stakeholder engagement, and ongoing scrutiny to safeguard against the erosion of trust and equity in patient care.

INTRODUCTION

Artificial intelligence (AI), a term encompassing all computer systems that are able to perform tasks mimicking human intelligence, has ushered in a new era of healthcare by revolutionizing patient care with unparalleled accuracy and efficiency. Recent studies have shown revolutionary AI impacts, including 2-fold reductions in colorectal cancer detection, a 10% decrease in image analysis errors made by radiologists, the development of stress-reducing robotic pets, and more, positioning AI at the focal point of scientific discussions.[1-3] The body of literature on AI's revolutionary impacts in healthcare extends beyond early detection of diseases and improved diagnostic accuracy. For instance, AI has been instrumental in predicting patient outcomes by analyzing patterns in vast datasets that human experts may overlook. A notable study by Esteva et al. demonstrated AI’s ability to accurately predict the prognosis of skin cancer patients, outperforming traditional prediction models.4 Moreover, AI applications have shown significant promise in personalizing treatment plans, where algorithms tailor treatment recommendations based on a patient's unique genetic profile, lifestyle, and disease characteristics, leading to more effective and targeted therapies. This is exemplified in the work by Korach et al., where AI-driven analyses of patient data resulted in customized chemotherapy regimens for cancer patients, significantly improving survival rates compared to standard care.[5] AI has also been utilized in mental health care for addressing the global shortage of mental health professionals by offering innovative solutions such as chatbots and virtual therapists that provide psychological support and therapy.[6] These AI systems engage users in meaningful conversations, helping to alleviate symptoms of depression and anxiety with accessibility and anonymity that traditional therapy cannot always offer.

The diversity of AI's applications underscores its potential to revolutionize every aspect of healthcare, from preventive medicine and early detection to treatment personalization and mental health support. This highlights a future where AI will become a cornerstone of patient care and medical research. One question that needs to be asked, however, is how the benefits of AI can be harnessed while simultaneously addressing its associated ethical challenges such as the risk of apathetic care, biased solutions, and the unsafe handling of patient data,

the unauthorized access and leakage of personally identifiable information (PII), the misuse of sensitive health information for purposes not consented to by patients, security breaches).[7] The excitement surrounding AI has led to this field advancing faster than its ability to thoroughly consider these complexities. Thus, the purpose of this literature review is to highlight the ethical issues and potential solutions that need to be considered before incorporating AI into medical practices to preserve empathy, unbiased treatments, and safe data handling for patients.

METHODS

Literature Search Strategy

A focused literature search was conducted exclusively on PubMed, chosen for its comprehensive collection of biomedical literature. Our search aimed to identify original research articles published in English within the past ten years, highlighting the latest advancements and ethical considerations in artificial intelligence (AI) within healthcare. Keywords included "Artificial Intelligence", "Healthcare", "Ethics", "Patient Data Security", "Empathetic Care", and "Bias in AI". These were combined using Boolean operators to refine searches, for instance, "Artificial Intelligence AND Healthcare" to locate studies specifically on AI's application in healthcare settings.

Selection Process

A focused literature search was conducted exclusively on PubMed, chosen for its comprehensive collection of biomedical literature. Our search aimed to identify original research articles published in English within the past ten years, highlighting the latest advancements and ethical considerations in artificial intelligence (AI) within healthcare. Keywords included "Artificial Intelligence", "Healthcare", "Ethics", "Patient Data Security", "Empathetic Care", and "Bias in AI". These were combined using Boolean operators to refine searches, for instance, "Artificial Intelligence AND Healthcare" to locate studies specifically on AI's application in healthcare settings.

RESULTS

The results section of our review synthesizes insights from various studies focusing on the ethical concerns of empathy and data privacy in the use of artificial intelligence (AI) in healthcare (Table 1). Montemayor et al. emphasize the limitations of AI in performing empathetic roles within healthcare settings, highlighting the moral and ethical implications of substituting human empathy with AI-driven processes.8 This is complemented by the work of Suchman et al. and Kim et al., who respectively formulated a model of empathic communication in medical interviews and developed scales to measure empathy-related constructs, assessing their impact on patient satisfaction and compliance.[9,10] Their findings suggest that while AI can assist in healthcare, its capability to replace the human element in empathetic interactions remains questionable.

On the subject of data privacy, studies such as those by Trinidad et al. and Winkler et al. explore public comfort and ethical considerations associated with sharing patient data.[11,12] Trinidad et al. found varied levels of comfort among the U.S. public regarding the sharing of health data with third-party companies, influenced by the purpose of data sharing. Winkler et al. discuss the ethical conditions necessary for the reuse of such data, advocating for a balance between public health benefits and patient rights. These themes are echoed in the research by Gurevich et al. and Baowaly et al., who explore the ethical frameworks for AI in healthcare and the development of synthetic electronic health records, respectively.[13,15] These studies collectively highlight a critical tension between leveraging technological advancements for improved healthcare delivery and ensuring the protection of individual privacy and ethical standards.

DISCUSSION

Can AI provide empathetic care?

Empathetic care is one of the most essential pillars a healthcare professional can provide and is necessary for fostering patient trust and holistic well-being. Three types of empathy exist: emotional, cognitive, and motivational empathy.[8] Emotional and motivational empathy originate from instinctive or biological connections that fosters compassionate understanding and motivates people to support one another. These connections reflect our inherent capacity to respond to others' emotional states, driven by neurochemical reactions and evolutionary processes that have shaped human social interactions. For instance, witnessing someone in distress can trigger an empathetic response due to the activation of mirror neurons, which produces a shared feeling of concern and prompts actions aimed at alleviating the other's suffering.[8] This biological underpinning of empathy facilitates a depth of understanding and connection that is fundamental to compassionate care, underscoring the challenges AI faces in replicating these complex human emotions and motivations. On the other hand, cognitive empathy differs by focusing on the detection of individuals'

mental states through the representation of their circumstances and observable features in their expressions.[8] AI in its current state can only provide the latter due to its lack of biological instincts. This behavior, however, resembles that of psychopathic individuals who tend to excel in cognitive empathy, yet lack the emotional and motivational depth associated with human connection.[8] Thus, there is a risk of creating AI models that mimic a "psychopathic" lack of humanity.

The clinician's skill in resonating with a patient's perspective (i.e., emotional and motivational empathy) leads to more effective care for at least a few reasons. To begin with, a study by Suchman et al., that investigated 21 videotaped verbal exchanges showed that patients were hesitant to reveal information at first and instead gave emotional hints such as "my headache kept coming back" with great anxiety until they felt that their doctor understood the significance of this point in their narrative.[9] They were only willing to share additional information when this connection was perceived. Secondly, it is important to note that adhering to treatment plays a significant role in patient recovery. Patients are more likely to follow treatment suggestions when they get a sense that their doctor is worried and accompanying them empathetically, given that partnership and perceived affective empathy have the most significant impact on compliance and satisfaction.[10] Therefore, it is evident that with a proper comprehension of what "empathy" entails, abilities that are inherent in empathetic manifestations are beyond the reach of AI. Instead of attempting to replicate emotional understanding, AI can potentially focus on other areas such as minimizing clinicians' administrative burdens.

Will patient data be secure when used with AI models?

Data privacy is crucial for the wellbeing of patients as it ensures trust and confidentiality in healthcare interactions. However, safe data has become a significant challenge, especially with advanced AI technologies falling into the hands of big tech corporations. Despite only 28.9% of 1841 surveyed American adults being open to sharing their data with these companies, hospitals continue to share patient information with such corporations, giving them valuable market insights for tool development. [11,12] This exchange can lead to potential privacy breaches, where sensitive patient data might be mishandled or improperly accessed, posing serious risks to individual privacy and autonomy. Hence, it is imperative to obtain consent from patients at different stages of data usage, providing them with the flexibility to revoke with changing circumstances. While there is, of course, a tradeoff between keeping data private and simultaneously accessible for the development of research studies and innovative AI technologies, the lack of data sharing transparency poses a

significant risk of loss of public trust and even litigation against the implementation of AI in healthcare. The power imbalance between the government and tech companies fuels this issue even more given the latter's control over resources and data infrastructure.[12] Moreover, the lack of thorough AI legislation enables these companies to prioritize interests with few limitations. Although Canada is aiming to play a leading role in the Artificial Intelligence and Data Act global framework (planned to take effect in 2025), the deadline is still far away with AI currently advancing faster than the ability to regulate it.[13] Lastly, the issue of re-identification is another challenge of data security that needs to be tackled. It was discovered that an algorithm could be used to re-identify 85.6% of adults and 69.8% of children from a nutrition database with anonymous health information, indicating that current practices of de-identification might be insufficient.14 To address this issue, a potential solution involves the use of "generative models'' which generate synthetic data that retains the statistical properties of the original patient without exposing any sensitive information.[14] The model learns statistical trends from a dataset and generates a new dataset that is similar statistically but with no actual patient records, thus preserving patient identity.

Does AI provide biased suggestions?

Considering the health and wellbeing of a diverse society, it is essential to tackle the risk of bias to ensure that treatments are personalized, adaptable, and free from discrimination. When training AI algorithms, there are 3 main sources of bias that the model can be affected by: data-driven, algorithmic, and human bias.[16] The algorithm learns from the data provided, and the higher the level of accuracy, diversity, and representativeness of the data, the more reliable and less

biased its clinical decisions are. One example of data-driven bias is the use of genome-wide association studies (GWAS) to evaluate an individual's genetic susceptibility to a disease. The drawback is that 81% of GWAS studies concentrate on people with European backgrounds, which affects the wider applicability of disease risk scores across diverse populations.[13] Algorithmic bias, on the other hand, can occur due to inadequate regulation during the design phase which does not represent the data from different groups equally.[16] For instance, the hurry to implement AI solutions during COVID-19 without thorough testing in diverse groups may have unintentionally worsened health disparities and diagnoses for these populations.[18] This is distinct from data-driven bias, which arises directly from the characteristics of the data itself, such as missing data from certain demographic groups. Algorithmic bias, however, stems from the decision-making process within the algorithm's design—choices about which data to emphasize, which models to use, and how to weigh different inputs.

Lastly, human biases in AI can be among the most difficult to identify because they stem from ingrained societal prejudices that are often subtle but magnified by AI and big datasets16. However, there are many things that can be done to mitigate these effects in the long term. Standardizing data to make it compatible with different algorithms could assist in accessible data that captures diverse population information[19]. For example, if healthcare data from different hospitals use varying terms for the same medical conditions, standardizing this data would involve mapping these disparate terms to a single, consistent vocabulary. This would allow algorithms to analyze the data uniformly, improving accessibility and inclusiveness. Therefore, standardizing data is crucial for creating accessible datasets that accurately reflect diverse population information, ultimately mitigating long-term effects in data-driven fields. Sharing inclusive AI models can also promote transparency and collaboration by ensuring that code is tested across the globe, with different populations and developers to validate and ensure its efficiency across various clinical settings.[20] Additionally, including communities in the design process can not only mitigate biases but also fulfill the health needs of these communities directly through the input of their personalized experiences and ideas, which are especially important in the current era of precision health care. A notable example is the OpenAPS initiative which is a system that pumps insulin through a device at home based on the patients' self-collected glucose levels, providing safe and personalized diabetes care.[21]

Limitations

This paper intends to provide readers with a brief overview of the ethical challenges and solutions regarding empathy, bias, and data privacy in healthcare. It does not focus on areas such as transparency , accountability, and governance which can give the reader a multi-dimensional analysis of AI's impact on healthcare. Additionally, the solutions are meant to be introductory in nature and do not discuss details of the specific steps that need to be taken. Future research can explore the complexities of achieving this in practice. This includes how AI can enhance rather than replace patient-clinician relationships, along with specific data collection and design strategies and regulations that can ensure ethical practices.

CONCLUSION

To sum up, AI has enormous potential to shape the healthcare industry. It is important that its integration into the healthcare system is done so responsibly and in a way that does not impair the well-being of society along the journey. Being aware of AI's limitations, training caregivers to collaborate with it, ensuring strict legislation on designing unbiased models, and establishing security safeguards to protect patient data are all necessary steps to be taken to foster safe AI-driven healthcare. Given that the AI technologies we use today are capable of rapid advancement, we run the risk of falling behind very quickly.

Acknowledgements: I would like to express our special thanks to the scientists who have provided knowledge in their research. I declare that I have no conflict of interest.

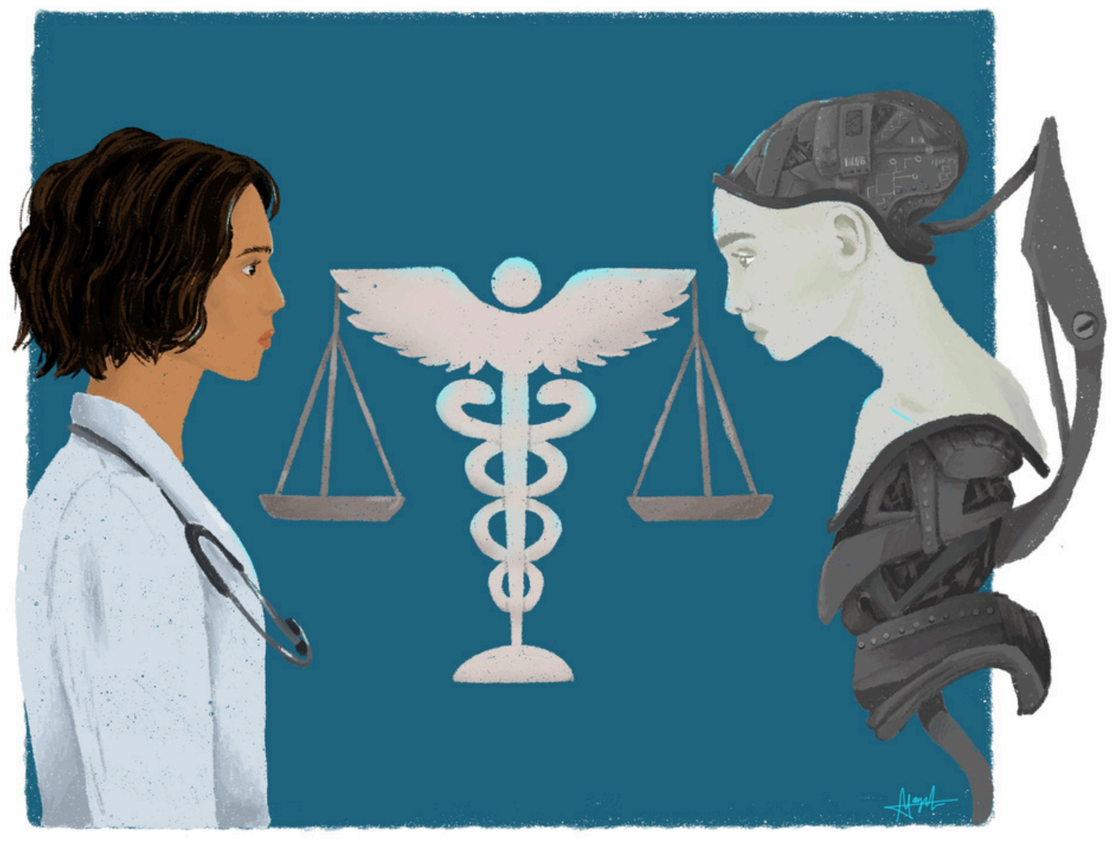

Table 1: Summary of sources selected in the review

Author	Origin	Ethical Concerns	Purpose	Type of Source	Research Design	Target Population	Framework Proposed	Major Themes
Montemayor et al. (2022)9	USA	Empathy	Assess AI's limits in empathetic roles in healthcare	Discussion	N/A	Public healthcare patients	No	Limitations of AI in empathetic healthcare, moral and ethical implications
Suchman et al. (1997)14	USA	Empathy	Formulate a model of empathic communication in medical interviews	Research	Descriptive, qualitative	Public healthcare patients	Model of empathetic communication	Empathic communication, emotional expression, physician-patient interaction
Kim et al. (2004)6	USA	Empathy	Develop scales for empathy-related constructs and test their effect on patient satisfaction and compliance	Research	Survey and modeling	University hospital outpatients	Empathy-satisfaction-compliance relationship	Physician empathy, patient satisfaction, patient compliance, communication skills
Trinidad et al. (2020)15	USA	Data Privacy	Assess public comfort with sharing health data with third-party companies for different purposes	Research	Survey and regression analysis	General U.S. public	No	Public comfort with data sharing, demographic factors, privacy concerns, healthcare data usage
Winkler et al. (2023)17	Germany	Data Privacy	Discuss the ethical considerations of sharing patient data from public healthcare with for-profit companies for research	Research	Conceptual analysis	Public healthcare patients	Ethical conditions for data reuse	Ethics of data sharing, patient rights, public health benefits
Gurevich et al. (2023)4	Canada	Data Privacy	Explore the interplay between health equity and AI, and how AI can either support or undermine health equity	Research	Conceptual analysis	Healthcare systems and patients	guidelines for ethically sharing patient data with for-profit companies	Health equity, AI in healthcare, ethical considerations, data sharing
Na et al.10	USA	Data Privacy	Evaluate the possibility of reidentifying individuals from deidentified physical activity data using machine learning	Research	Cross-sectional study	U.S. population	No	Data privacy, reidentification, machine learning, physical activity data

Author	Origin	Ethical Concerns	Purpose	Type of Source	Research Design	Target Population	Framework Proposed	Major Themes
Baowaly et al. (2019)1	Taiwan	Data Privacy	Develop and test advanced generative adversarial network models for creating synthetic electronic health records that are more realistic and efficient	Research	Experimental study using databases	Medical researchers	Improved data synthesis models	Synthetic data, model comparison
Norori et al. (2021)11	Switzerland	Bias/	Address algorithmic bias in AI for healthcare by promoting open science practices.	Research	Theoretical analysis	Healthcare professionals and AI researchers	Open science framework to mitigate AI bias	AI bias, data representation, open science approaches
Popejoy and Fullerton (2016)13	USA	Bias	Evaluate diversity in genomic studies and its implications for precision medicine.	Research	Analysis	Global population	Framework for inclusive genomic research	Genetic diversity, precision medicine, underrepresentation
Leslie et al. (2021)7	UK	Bias	Assess AI's impact on health inequities during the COVID-19 pandemic	Discussion	Conceptual analysis	Global healthcare systems and vulnerable populations	Recommendations for equitable AI use	AI and health equity, COVID-19, algorithmic bias
Kelly et al. (2019)5	USA	Bias	Explore the challenges and necessary steps for integrating AI into clinical healthcare practice	Discussion	N/A	Healthcare professionals and patients	AI integration in healthcare	AI in healthcare, clinical integration, machine learning challenges, regulatory considerations

Author	Origin	Ethical Concerns	Purpose	Type of Source	Research Design	Target Population	Framework Proposed	Major Themes
Tzovaras et al. (2019)3	USA	Bias	Explore challenges of integrating and sharing personal data for research via Open Humans platform	Research	Platform-based participatory research	Digital platform users	Open Humans platform	Personal data integration, ethical considerations, privacy, community-centric data sharing
Lewis et al. (2016)8	USA	Bias	Explore the real-world use and implications of open-source artificial pancreas systems by the diabetes community	Research	Descriptive analysis	Diabetes patients using open-source systems	No	Community-driven innovation, safety, user autonomy

REFERENCES

1. Bercean BA, Birhala A, Ardelean PG, et al. Evidence of a cognitive bias in the quantification of COVID-19 with CT: an artificial intelligence randomised clinical trial. Sci Rep. 2023;13(1):4887. doi:10.1038/s41598-023-31910-3
2. Petersen S, Houston S, Qin H, Tague C, Studley J. The Utilization of Robotic Pets in Dementia Care. J Alzheimers Dis. 2017;55(2):569-574. doi:10.3233/JAD-160703
3. Wallace MB, Sharma P, Bhandari P, et al. Impact of Artificial Intelligence on Miss Rate of Colorectal Neoplasia. Gastroenterology. 2022;163(1):295-304.e5. doi:10.1053/j.gastro.2022.03.007
4. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-118. doi:10.1038/nature21056
5. Dodington DW, et al. Analysis of tumor nuclear features using artificial intelligence to predict response to neoadjuvant chemotherapy in high-risk breast cancer patients. Breast Cancer Res Treat. 2021;186(2):379-389. doi:10.1007/s10549-020-06093-4
6. Khawaja, Z., & Bélisle-Pipon, J. C. (2023). Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots. Frontiers in digital health, 5, 1278186. https://doi.org/10.3389/fdgth.2023.1278186
7. Seh, A. H., Zarour, M., Alenezi, M., Sarkar, A. K., Agrawal, A., Kumar, R., & Khan, R. A. (2020). Healthcare Data Breaches: Insights and Implications. Healthcare (Basel, Switzerland), 8(2), 133. https://doi.org/10.3390/healthcare8020133
8. Montemayor C, Halpern J, Fairweather A. In principle obstacles for empathic AI: why we can't replace human empathy in healthcare. AI Soc. 2022;37(4):1353-1359. doi:10.1007/s00146-021-01230-z
8. Suchman AL, Markakis K, Beckman HB, Frankel R. A model of empathic communication in the medical interview. JAMA. 1997;277(8):678-682.
9. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. Eval Health Prof. 2004;27(3):237-251. doi:10.1177/0163278704267037.
10. Trinidad MG, Platt J, Kardia SLR. The public's comfort with sharing health data with third-party commercial companies. Humanities Soc Sci Communications. 2020;7(1):149. doi:10.1057/s41599-020-00641-5
11. Winkler EC, Jungkunz M, Thorogood A, Lotz V, Schickhardt C. Patient data for commercial companies? An ethical framework for sharing patients' data with for-profit companies for research. J Med Ethics. Published online 2023. doi:10.1136/jme-2022-108781
12. Gurevich E, El Hassan B, El Morr C. Equity within AI systems: What can health leaders expect? Healthc Manage Forum. 2023;36(2):119-124. doi:10.1177/08404704221125368
13. Na L, Yang C, Lo CC, et al. Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning. JAMA Netw Open. 2018;1(8):e186040. doi:10.1001/jamanetworkopen.2018.6040
14. Baowaly MK, Lin CC, Liu CL, Chen KT. Synthesizing electronic health records using improved generative adversarial networks. J Am Med Inform Assoc. 2019;26(3):228-241. doi:10.1093/jamia/ocy142
15. Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A. Addressing bias in big data and AI for health care: A call for open science. Patterns. 2021;2(10):100347. doi:10.1016/j.patter.2021.100347
16. Popejoy AB, Fullerton SM. Genomics is failing on diversity. Nature. 2016;538(7624):161-164. doi:10.1038/538161a
17. Leslie D, Mazumder A, Peppin A, Wolters MK, Hagerty A. Does "AI" stand for augmenting inequality in the era of covid-19 healthcare? BMJ. 2021;372:n304. doi:10.1136/bmj.n304

18. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195. doi:10.1186/s12916-019-1426-2

19. Greshake Tzovaras B, Angrist M, Arvai K, et al. Open Humans: A platform for participant-centered research and personal data exploration. GigaScience. 2019;8(6):giz076. doi:10.1093/gigascience/giz076

20. Lewis D, Leibrand S, #OpenAPS Community. Real-World Use of Open Source Artificial Pancreas Systems. J Diabetes Sci Technol. 2016;10(6):1411. doi:10.1177/1932296816665635