## Artificial or Human Judgement: Implications of Automating High-Stakes Decision Making in Healthcare

An STS Research Paper presented to the faculty of the School of Engineering and Applied Science University of Virginia

by

Allison Miller

May 11, 2023

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Allison Miller

STS Advisor: Peter Norton

# Artificial or Human Judgement: Implications of Automating High-Stakes Decision Making in Healthcare

By definition, artificial intelligence (AI) systems emulate human intelligence; they may support human performance or alleviate administrative burdens through automation (Kabalisa & Altmann, 2021). AI systems can boost productivity, optimize resource allocation, reduce labor costs, and increase GDP (Kabalisa & Altmann, 2021). These systems are now common in healthcare, for example in medical diagnostics, patient monitoring, and learning healthcare systems (Lysaght, 2019). Hospitals, clinics, and other care providers adopt medical AI to prevent human error and diminish the effect of biases, but supplementary reasons include cost reduction and liability risk management. Like other humans, medical professionals are subject to cognitive biases. Verghese et al. (2018) found that optimism bias can lead clinicians to overestimate life expectancies by a factor of 5. Yet automated systems are not immune from bias; indeed algorithms can further embed racial and socioeconomic inequalities in healthcare (BMJ, 2020). Patient concerns with the applications of AI in healthcare include data privacy, cybersecurity, accuracy of automated diagnostics, responsibility of errors produced by algorithms, and the lack of regulatory oversight over the creation of AI systems (Musbahi et al., 2021). Nevertheless, the proper place of medical AI in the healthcare system remains a matter of disagreement, and healthcare providers, medical professionals, medical technology companies, advocacies, and patients are competing to influence how it is applied.

Participants include digital advocacies such as the Algorithmic Justice League (AJL), which demand affirmative consent, procedural transparency, continuous oversight and accountability, and actionable assessments. AJL warns: "AI systems can perpetuate racism, sexism, ableism, and other harmful forms of discrimination, therefore, presenting significant

threats to our society - from healthcare, to economic opportunity, to our criminal justice system" (AJL, n.d.a). Healthcare technology companies, such as Optum Health, promote AI in healthcare: "The responsible use of AI can help health systems strengthen and scale essential functions and reduce administrative burdens, all of which helps clinicians focus on their core mission of patient care" (UnitedHealth Group, 2022). Patients who oppose or welcome AI in healthcare are also key participants. An anonymous patient in support stated: "I don't think you will cure a lot of diseases without that advanced intellect. Obviously, we've come a long way with the human brain, but we could probably go a lot farther and speed the process with AI" (Richardson, 2021). Some patients are apprehensive about AI in healthcare. One explains: "I believe the doctor always has the responsibility to be checking for you, and you're his responsibility, you know? The AI is not responsible; that's just a tool" (Richardson, 2021). Opinions on algorithm implementation also vary amongst physicians. Dr. Ziad Obermeyer, associate professor of health policy and management at the University of California, Berkeley has said, "When I work in the emergency department, learning about a new patient feels like drowning in information. What electronic medical records have done is convert a large, inaccessible stack of paper charts into a never-ending stack of virtual notes and results. It's not a good task for humans; we just don't have the bandwidth to process it." Dr. Danton Char, assistant professor of anesthesiology, perioperative, and pain medicine at Stanford, stated, "The first fear is that AI and machine learning may worsen the economic and racial disparities already inherent in U.S. health care. Second, U.S. health care is in a constant tension between profit and delivery of health. Those two agendas rarely line up. Consequently, the values of AI designers or the purchasing administrators are not necessarily the values of the bedside clinician or patient. Those value collisions and tensions are going to be sites of significant ethical conflict" (Ward,

2019). Beyond the individual and organizational perspective, the FDA generally welcomes AI in healthcare. In a statement, it asserted that "AI has the potential to transform health care by deriving new and important insights from the vast amount of data generated during the delivery of health care every day" (FDA, 2021).

While participants agree that AI has valuable applications in healthcare, they disagree about its optimum implementation. Most mainstream, healthcare service provider participants favor pervasive implementation of medical AI to improve reliability, prevent errors, and reduce costs. Yet many medical professionals and some patient advocates are more cautious, warning of hazards in the form of algorithmic biases, the vulnerability of patient data, and the compromising of expert human judgment and interpersonal care in successful medicine.

### **Review of Research**

Globally, there is an increasing deficit between labor demands of the healthcare industry and the workforce of healthcare professionals. The human resource crisis can be attributed to doctor shortages, burnout of physicians from being overloaded, and a higher demand for chronic care (Meskó et al., 2018). To ensure that patient needs are met while maintaining a feasible workload for physicians, there is a growing need for disruptive, yet helpful technology, especially as this crisis worsens year by year. While there are concerns surrounding artificial intelligence systems in healthcare, there are a number of beneficial opportunities associated with its utilization. As the collection of patient-specific information has gone almost entirely digital through the transition to electronic medical records (EMRs), the expansive workload and administrative burden now placed on healthcare professionals is leading to widespread burnout (Verghese et al., 2018). By synthesizing medical records and eliminating the need for manual analysis of essential patient information, the application of AI in healthcare can alleviate the

demanding duties of physicians (Aung et al., 2021). Many of the routine jobs performed by physicians or nurses can be automated by AI, freeing up more time for more intensive, complex tasks. Another useful extension of AI in healthcare is the development of ML-based chatbot applications that are used to reduce the number of unnecessary hospital admissions (Armstrong, 2018). These applications seek to offer screening capabilities that eliminate the need for in-house input from a physician to determine a medical diagnosis.

From a medical diagnosis standpoint, AI has shown to be on par with, or even surpass, human abilities in precision, speed, and accuracy. For example, a study was conducted on the cardiovascular Artificial Intelligence-Clinical Support System (AI-CDSS) towards its ability to assist physicians in heart failure diagnosis (Choi et al., 2020). AI-CDSS operates off of guidelines developed by domain-experts, or physicians, and rule generation by an ML algorithm. The accuracy of the AI system was evaluated against expert diagnoses and concordance, expressed as a percentage, was concluded if diagnoses between the AI and the expert matched (Choi et al., 2020). Results of this study found that between AI-CDSS and HF experts, there was a concordance rate of 98%, but between experts and non-HF specialists, there was a concordance rate of 76% (Choi et al., 2020). These results signify the utility of using an AI system when a domain-expert is unavailable, as it had comparable accuracy. Esteva et al. (2017) found that a deep convolutional neural network (CNN) was more effective than dermatologists in identifying cancerous skin lesions using image-based classifications. The CNN achieved an accuracy of 72% while the dermatologists in the study had an accuracy of around 66% (Esteva et al., 2017). Trained on a dataset of over 100,000 clinical images, the CNN was able to make generalizable inferences on unseen clinical images. The results of this study exhibit the power of AI and its potential impact on crucial fields like oncology. However, although research has shown there are

clear benefits to be reaped from the implementation of AI systems in healthcare, the reliance on datasets rather than physician expertise introduces the risk of algorithmic bias impacting outputs produced by algorithms.

The National Library of Medicine defines algorithmic bias as: "the instances when the application of an algorithm compounds existing inequities in socioeconomic status, race, ethnic background, religion, gender, disability or sexual orientation to amplify them and adversely impact inequities in health systems" (Panch et al., 2019). The problem is not limited to healthcare. As data analytics have proliferated, there is a growing need for mitigating embedded biases. However, the urgency for regulation of AI in healthcare is due to the potential harmful impacts of automated decision-making. For decision making and diagnostics, AI algorithms are trained with data derived from EMRs that contain anonymized patient information from previous diagnostics, procedures, and outcomes (Verghese et al., 2018). However, nonwhite people have long been underrepresented or misrepresented in these datasets, resulting in mistreatment and inequities (Norori et al., 2021). For example, Obermeyer et al. (2019) studied a widely used algorithm to determine if the risk score assigned to patients was distorted by racial bias. The algorithm assessed operates under the assumption that "those with the greatest care needs will benefit the most" and consequently relies on historical patient data as the key predictor towards future health care needs (Obermeyer et al., 2019). The conditional risk scores were derived from patients' health status and the number of chronic conditions by race. After analyzing the output of risk scores across Blacks and Whites, the researchers found that "less-healthy Blacks scored at similar risk scores to more-healthy Whites." These results indicate significant disparities in the output produced by the algorithm (Obermeyer et al., 2019). As this algorithm is used in programs nationwide, the harmful impacts of this racial bias may be unknowingly affecting millions of

Americans each year. Publication of studies like this have led healthcare participants to inquire upon the implications of the full-scale deployment of AI systems.

#### **Patient Apprehension Towards AI Implementation**

Controversy surrounds the implementation of AI in healthcare, or any industry, because it can alleviate administrative burden while simultaneously worsening issues of equity and racism prevailing in society. As a result, there is a lack of consensus amongst participants regarding the scale and purpose of using AI in healthcare. Patients, being directly impacted by the integration of AI systems through diagnoses and treatment, have varied perspectives on how to approach the issue. According to a Pew Research Study survey of 11,000 participants in 2022, a majority of adults in the United States would be uncomfortable if their healthcare provider relied on AI to diagnose disease and recommend treatments (Tyson et al., 2023). Forty-percent of Americans say that AI can help to mitigate human error, but the same majority emphasize the potential loss in value with patient-provider relationships. On algorithmic bias, 64 percent of Black adults, 42 percent of Hispanic adults, and 39 percent of English-speaking adults agree that bias introduced by racial or ethnic identity is a major problem in healthcare. However, 51 percent of adults surveyed believe that the use of AI in healthcare will help reduce bias and inequitable treatment, rather than worsening it (Tyson et al., 2023). Richardson et al. (2021) found that patients are intrigued about the future of AI in healthcare, but are apprehensive towards safety and preservation of patient autonomy. One participant from the study noted, "It's gonna do what it's gonna do. I think that I look forward with excitement... but I agree that there's definitely some flaws, and there definitely needs to be some markers in there, at least right now, that can also protect people. But it's going to be positive if we can get those safe markers in" (Richardson et al., 2021). Patients demand that clinicians ensure their safety when using AI to support medical

decision making and necessitate regulatory oversight against consumer harms. While patient perspectives on AI differ amongst demographics, the overall consensus is that great caution should be taken towards solely relying on algorithms for medical diagnoses and treatment. Automation can ease the workload of overworked medical professionals, but the personal relationship between the patient and the provider cannot be replaced by an AI interface.

#### Alleviating Physician Burnout through AI Automation

Physicians recognize the value of AI integration in healthcare while still being hesitant. In a survey of 77 medical professionals, participants were tasked with ranking a set of 40 statements regarding the ethics and implementation of AI in healthcare using q-methodology for analysis (Martinho et al., 2021). Analysis of the survey results gave way to 4 general physician perspectives on AI in healthcare: "AI is a helpful tool: let physicians do what they were trained for, rules and regulations are crucial: private companies only think about money, ethics is enough: private companies can be trusted, and explainable AI tools: learning is necessary and inevitable" (Martinho et al., 2021). Generally, physicians support the implementation of AI as a tool to aid in automatic processes, allowing more time to focus on their licensed skill sets and tasks. The physicians in this study acknowledge that AI should not be used as the decision-maker in healthcare, rather, it should only aid in the decision-making process. One participant noted, "The role of a skilled physician is to take into consideration what a machine/AI tells him and make the correct connection with clinical reality." AI systems offer the ability to provide datadriven advice towards medical diagnosis and treatment, but only the medical professionals can assess a specific patient's condition with clinical expertise. A concern that physicians have is the potential infringement of confidential patient data by private healthcare companies specializing in AI technologies. The aspirational pillars of patient respect and privacy that physicians must

adhere to do not typically align with the monetary-focused values of the health technology industry. Sarwar et al. (2019) found that amongst surveyed pathologists, 38 percent do not fear that AI will impact employment prospects. Furthermore, many believe that the implementation of AI will create new employment opportunities and improve diagnostic efficiency with appropriate training (Sarwar et al., 2019). In general, medical professionals advocate for AI in healthcare and have a positive outlook on its implementation, but only if they remain in charge of the medical decision process.

#### **Importance of Data Analytics for Healthcare Service Providers**

Healthcare service providers overwhelmingly support the implementation of AI systems into healthcare in order to improve patient outcomes and administrative processes. With the growing importance of data science encroaching on nearly every global industry, many health service companies are developing or acquiring data analytics subsidiaries. For example, Cigna Healthcare, a global health service company serving over 100 million people around the world, owns a clinical analytics branch called Evernorth (Cigna, n.d.). Gina Papush of Evernorth emphasizes the advantages of using predictive AI in healthcare: "Cancer, heart disease, diabetes - chronic conditions like these contribute to the majority of healthcare episodes and costs in the United States. To better support customers, Evernorth is using machine learning to predict a chronic diagnosis before it occurs - and providing patients and providers with intelligence to inform their decisions and lower costs" (Cigna, n.d.). From the perspective of service providers, the widespread implementation of AI into health systems would have tremendous benefits, not only for the physicians and patients, but for the providers as well with the reduction of costs accrued from chronic care. Optum Enterprise Analytics, a data analytics subsidiary of UnitedHealth Group, echoes Cigna representatives stating that AI will revolutionize the

administrative workload in healthcare. Sanji Fernando, head of the Center for Applied Data Science within OptumLabs, states: "While there is great opportunity to use machine learning for clinical decision support, there is just as much opportunity to optimize administrative processes. There is no doubt that AI will transform clinical care, but the lower hanging fruit is administrative challenges" (Davenport, 2020). Healthcare service providers argue that the biggest selling point for the implementation of AI is the alleviation of administrative burden physicians and nurses. With their workload lessened, service providers second the previously mentioned notion that medical professionals will have adequate time and energy to dedicate towards more complex, demanding tasks, helping the healthcare system as a whole.

#### Advocacies Call for Regulation of AI Systems in Healthcare

Advocacies, such as the American Civil Liberties Union (ACLU) and the AI Healthcare Coalition, agree that policy change is necessary to prevent the automation of discrepancies in healthcare. For example, the ACLU has laid out a framework detailing how key stakeholders should address the issue: requiring transparency of demographic information, impact assessments conducted by the FDA to determine if there are any differences across racial or ethnic subgroups before a device is approved, labels if a device passes the assessment, and establishing best practices guidelines for manufacturers not under the FDA to follow. The AI Healthcare Coalition, an alliance of healthcare AI innovators and stakeholders, offers another framework for AI implementation that is centered upon ethical principles of nonmaleficence, agency and accountability (AIHC, n.d.). The principles of the framework include improving patient outcome while ensuring safety, efficacy and equity, designing AI systems that are aligned with scientific knowledge of human clinician cognition, maximizing traceability of patient derived data, rigorous validation of outcomes using direct or linked evidence to clinical outcomes, and aligning liability (AIHC, n.d.). Members of the coalition include academic hospitals, government agencies, and private companies. In addressing the need for a national registry of clinical algorithms, Johan Halamaka, coalition affiliate and president of Mayo Clinic Platform, stated: "We don't have the tools today to understand whether machine learning algorithms and these new technologies are being deployed are good or bad for patients. The only way to change that is to more rigorously study their impacts and make the results transparent, so that users can understand the benefits and risks" (Ross, 2023). In July 2020, Joy Buolamwini, founder of the Algorithmic Justice League, co-founded the Community Reporting of Algorithmic System Harms (CRASH) project. The key addresses of the CRASH project are to prevent, report, and redress harms through bug bounties, coordinated bias disclosure, intersectional AI audits, and education. Advocacies stand to be the voice calling for regulation of AI systems as they are implemented throughout the healthcare system. They recognize that there is no stopping the development of artificially intelligent programs, and are working towards reducing consumer harms associated with their utilization. In order to ensure that the users only have access to ethical and accurate automated systems that optimize patient outcomes, the frameworks proposed by advocacies and agencies must be considered.

#### Conclusion

Regardless of patient, physician, advocacy, or healthcare service provider perspectives, AI systems will continue to be implemented. Given the accelerating pace of generative AI systems such as ChatGPT, there is an evolving need for an ethical framework of implementation. It is essential to understand the values held by each of these participants when evaluating the equitable application of automated systems. Patients are the core of the healthcare industry and their well-being must remain the top priority. Therefore, regulatory oversight and transparency

are critical requirements to ensure the integrity of the medical field is upheld. Only by understanding and respecting the values held by all stakeholders can a holistic framework for integrating AI in healthcare be created. Ultimately, it is crucial to maintain the trust and confidence of patients, who depend on healthcare providers to deliver safe and effective care, as challenges and opportunities are presented by AI technology in the field of medicine.

#### References

- AJL (n.d.a) The Algorithmic Justice League. Mission, Team and Story. www.ajl.org/about
- AJL (n.d.b) The Algorithmic Justice League. Algorithmic Vulnerability Bounty Project. https://www.ajl.org/crash-project
- Armstrong, S. C. (2018). The apps attempting to transfer NHS 111 online. *British Medical Journal*, k156. doi.org/10.1136/bmj.k156
- Aung, Y. Y., Wong, D., & Ting, D. S. J. (2021). The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. *British Medical Bulletin*, 139(1), 4–15. doi.org/10.1093/bmb/ldab016
- BMJ. (2020, Feb 13). Can we trust AI not to further embed racial health inequalities? British Medical Journal. www.bmj.com/company/newsroom/can-we-trust-ai-not-to-further-embed-racial-healthinequalities
- Choi, D., Park, J. B., Ali, T., & Lee, S. (2020). Artificial intelligence for the diagnosis of heart failure. *Nature Partner Journals Digital Medicine*, 3(1). <u>doi.org/10.1038/s41746-020-0261-3</u>
- Cigna (2021, July 28). Cigna Group. Cigna Uses Data and AI to Improve Patient Outcomes. Cigna Healthcare Newsroom. <u>newsroom.cigna.com/how-cigna-uses-data-and-ai-to-improve-patient-outcomes</u>
- Davenport, T. (2020, Oct 9). Optum Focuses On AI To Improve Administrative Decisions. *Forbes*. <u>www.forbes.com/sites/tomdavenport/2020/10/09/optum-focuses-on-ai-to-improve-administrative-decisions/?sh=6a1023336dad</u>
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J. S., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. doi.org/10.1038/nature21056
- AIHC (n.d.). AI Healthcare Coalition. Ethical AI Principles. ai-coalition.org/ethicalaiprinciples
- FDA (2021). U.S. Food and Drug Administration. Center for Devices and Radiological Health and Center for Devices and Radiological Health. (2021). Artificial Intelligence and Machine Learning in Software as a Medical Device. <u>www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-andmachine-learning-software-medical-device</u>

- Kabalisa, R., & Altmann, J. (2021). AI Technologies and Motives for AI Adoption by Countries and Firms: A Systematic Literature Review. In *Lecture Notes in Computer Science* (pp. 39–51). Springer Science+Business Media. doi.org/10.1007/978-3-030-92916-9\_4
- Lee, N. T., Resnick, P., & Barton, G. (2022, March 9). Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. *Brookings*. <u>www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-</u> <u>and-policies-to-reduce-consumer-harms</u>
- Lysaght, T., Lim, H. Y., Xafis, V., & Ngiam, K. Y. (2019). AI-Assisted Decision-making in Healthcare. *Asian Bioethics Review*, *11*(3), 299–314. doi.org/10.1007/s41649-019-00096-0
- Martinho, A., Kroesen, M., & Chorus, C. G. (2021). A healthy debate: Exploring the views of medical doctors on the ethics of artificial intelligence. *Artificial Intelligence in Medicine*, 121, 102190. doi.org/10.1016/j.artmed.2021.102190
- Meskó, B., Hetényi, G., & Győrffy, Z. (2018). Will artificial intelligence solve the human resource crisis in healthcare? *BMC Health Services Research*, 18(1). doi.org/10.1186/s12913-018-3359-4
- Musbahi, O., Syed, L., Feuvre, P. L., Cobb, J., & Jones, G. (2021). Public patient views of artificial intelligence in healthcare: A nominal group technique study. *Digital Health*, 7, 205520762110636. doi.org/10.1177/20552076211063682
- Norori, N., Qiyang, H., Aellen, F. M., Faraci, F., & Tzovara, A. (2021). Addressing bias in big data and AI for health care: A call for open science. *Patterns*, 2(10), 100347. doi.org/10.1016/j.patter.2021.100347
- Obermeyer, Z., Powers, B. W., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. doi.org/10.1126/science.aax2342
- Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: implications for health systems. *Journal of Global Health*, 9(2). <u>doi.org/10.7189/jogh.09.020318</u>
- Rai, T. S. (2019). Racial bias in health algorithms. *Science*. doi.org/10.1126/science.366.6464.440-e
- Rainie, L., Anderson, J., & Rainie, L. (2022a, Sep 15). Code-Dependent: Pros and Cons of the Algorithm Age. *Pew Research Center: Internet, Science & Tech.* <u>www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age</u>

- Rainie, L., Anderson, J., & Rainie, L. (2022b, Sep 15). Theme 4: Biases exist in algorithmicallyorganized systems. *Pew Research Center: Internet, Science & Tech.* www.pewresearch.org/internet/2017/02/08/theme-4-biases-exist-in-algorithmicallyorganized-systems
- Richardson, J. P., Smith, C., Curtis, S., Watson, S. E., Zhu, X., Barry, B., & Sharp, R. R. (2021). Patient apprehensions about the use of artificial intelligence in healthcare. *Npj Digital Medicine*, 4(1). doi.org/10.1038/s41746-021-00509-1
- Ross, C. (2023, April 27). A new coalition aims to close AI's credibility gap in medicine with testing and oversight. *STAT*. <u>www.statnews.com/2022/12/07/artificial-intelligence-hospitals-coalition-health-ai</u>
- Sarwar, S., Dent, A., Faust, K., Richer, M., Djuric, U., Van Ommeren, R., & Diamandis, P. (2019). Physician perspectives on integration of artificial intelligence into diagnostic pathology. Npj Digital Medicine, 2(1). doi.org/10.1038/s41746-019-0106-0
- Tyson, A., Pasquini, G., Spencer, A., Funk, C., & Nadeem, R. (2023, Mar 6). 60% of Americans Would Be Uncomfortable With Provider Relying on AI in Their Own Health Care. *Pew Research Center Science & Society*. <u>www.pewresearch.org/science/2023/02/22/60-of-</u> <u>americans-would-be-uncomfortable-with-provider-relying-on-ai-in-their-own-health-care</u>
- UnitedHealth Group. (2022). AI Survey: Health Care Organizations Continue to Adopt Artificial Intelligence to Help Achieve Better, More Equitable and Affordable Patient Outcomes. UnitedHealth Group. www.unitedhealthgroup.com/newsroom/2021/2021-12-15-optum-ai-survey-for-betterequitable-affordableoutcomes.html?cid=EM:SubscribetoNews:OA:000:standard:NAT:Newsroom
- Verghese, A., Shah, N. H., & Harrington, R. A. (2018). What This Computer Needs Is a Physician. Journal of the American Medical Association, 319(1), 19. doi.org/10.1001/jama.2017.19198
- Ward, L. (2019, Oct 14). The Ethical Dilemmas AI Poses for Health Care. Wall Street Journal. www.wsj.com/articles/the-ethical-dilemmas-ai-poses-for-health-care-11571018400