Artificial Intelligence in Healthcare: Applications and Cautions

Joyce Hu

15 June 2022

"Health is one of the most important and wide-spread social sectors in which algorithms are already used at scale today, unbeknownst to many" (Obermeyer et al. 2019, 447). With the development of artificial intelligence and the plethora of possible applications, the healthcare system may face a revolution that will overcome current barriers to access. In countries like India where there are only around 12 thousand ophthalmologists for over 60 million citizens suffering from diabetic retinopathy—a major preventable cause of blindness—seeking expert care is often not an option (Pandey and Sharma 2018, 1653). With the development of AI that can quickly and inexpensively diagnose illnesses, specialized care that was only available to the affluent could be made available to all who require quality care. This memo will discuss the current and potential applications of AI in healthcare, and evaluate the potential risks associated with this rapidly growing field.

AI Applications in Healthcare: Present and Future

Medical Imaging

Millions die from cancer every year, and the best approach to saving their lives is to perform early detection and diagnosis—an approach that can be expedited with the development of AI techniques in medical imaging. Spurred by the development of MRI, CT, and PET scans, medical imaging allows clinicians to diagnose and treat illnesses by studying dense, data-packed images. This process is often challenging and time-consuming, even for experienced clinicians. Machine learning and deep learning AI algorithms can assist clinicians in making faster and more accurate diagnoses. In one study, researchers applied a deep learning algorithm to diagnose lung cancer from CT scans (Ardila et al., 2019). Typically, radiologists pore over hundreds of images to look for small discrepancies in lung tissue and patterns of cancer. In contrast, the AI model used the patient's current and prior CT scans to quickly locate those discrepancies. The model found stellar success, with a 94.4 percent AUC (Ardila et al., 2019). Compared to six radiologists, the model had an eleven percent reduction in false positives and a five percent reduction in false negatives. AI in medical imaging has also been seen in diagnosing breast cancer, finding cardiovascular abnormalities, and identifying diabetic retinopathy from eye images (Bresnick, 2018; Peng, 2020). For many people in developing countries—and industrialized countries as well—the traditional approach to cancer screening is not fiscally practical. Especially with the rising global demand for radiologists and other medical specialists, AI in medical imaging can save lives by making specialized care more accessible and expediting the diagnosis of serious illnesses.

AI-Assisted Surgery

A potential clinician-oriented medical AI application is AI-assisted surgery, which will use deep learning methods to support surgeons in the operating room (Bodenstedt et al., 2020). Currently, this application is challenged by the need for large amounts of labeled training data. Unlike typical training sets where a layperson can easily define terms, training data sets for assisted surgery requires real-time expert knowledge from a clinical perspective, which is especially difficult to obtain during surgery (451). When these challenges are overcome, AI can be used to

help surgeons make data-driven decisions and offer "context-aware assistance" (451). For example, AI can use data from hundreds of patients to predict complications and offer additional information at the optimal time during surgery by studying typical operating room workflow (452). This can also alert operating room clinicians to dangerous deviations from normal workflow.

Medical Scribing

AI-assisted medical scribing is a solution for addressing the demand for medical experts by reducing the time burden of documentation. In a 2016 study published in *Annals of Internal Medicine*, physicians reported that they only spent 27 percent of their day directly communicating with patients, and 49 percent on their day working on the electronic health record and desk work (Medical Economics, 2021). 86 percent of physicians surveyed by *Medical Economics* cited paperwork as the number one reason for feeling burnt out (2021). The burden of completing paperwork and updating electronic health records detracts from a physician's role to care for their patients. Companies like Suki and Notable utilize voice recognition and natural language processing (NLP) to assist physicians in verbally updating electronic health records (EHR) (Nanalyze, 2020). Instead of manually inputting patient-reported information, AI medical scribes will input the information verbally reported to it. Some systems also use NLP to help doctors narrow down relevant information to save time pursuing dense patient records (Nanalyze, 2020). AI-assisted medical scribing helps make the EHR system more efficient and allows physicians to spend more time in conversation with the patient rather than inputting data.

Virtual Assistants in Healthcare

During the COVID-19 pandemic, hospitals were flooded with patients requiring emergency care and focused their resources on expanding intensive care units and reallocating clinical staff. The unexpected public health crisis placed hospitals under tremendous burden to provide normal clinical care while meeting the influx of demand for COVID-19 care. As a result, routine or elective procedures for patients in noncritical conditions were postponed (Jadczyk, 2021). There was a reduction in hospital admissions for patients without COVID-19 because of limited resources and patients' fear of being infected with the illness. Moreover, the pandemic exposed existing public health gaps and exacerbated access to care issues in low-income populations. For a healthcare system that relied primarily on face-to-face health delivery, the demand to meet patients' needs while preventing the spread of disease motivated healthcare providers to turn to telehealth service options.

To reduce the volume of in-person non-essential hospital visits, save resources for COVID-19 patients, and reduce the possibility of spreading the disease, many providers relied on video conferencing and other telehealth services (Jadczyk, 2021). Although effective for mitigating in-person contact, telehealth delivery is time and resource-intensive—patients still require individual assessment (Jadczyk, 2021). This limitation in telehealth delivery motivated healthcare providers to implement AI-driven conversational agents and virtual assistants. Both forms of intelligent communication technology serve similar goals of relieving the burden on healthcare resources by making patient screening, disseminating medical information, and other functions more efficient and accessible.

Specifically, conversational agents, or “chatbots,” generally use machine learning to engage with users through text and multi-turn dialogue. Startups like Babylon Health and Buoy Health have been incorporating these functions as the first line of contact between patient and physician (Nanalyze, 2020). Their chatbots serve as symptom checkers and can provide education-level services by answering patients' questions about their conditions. They also aid in relieving the burden on hospital systems by either advising patients to seek care or stay home, and by answering educational-level questions. Other chatbots address access to care issues through care triaging, which can prediagnose patients’ health status before meeting with a clinician (Sezgin, 2020).

Voice assistants (VA), popularized by Apple’s Siri and Amazon’s Alexa software, use natural language processing methods to address health delivery limitations through voice communication rather than text (Sezgin, 2020). Voice assistants may also seem like a welcome alternative for patients who would prefer to disclose information to a VA rather than their healthcare practitioner. A study on veterans with PTSD found that in clinical interviews, patients were more likely to disclose personal information when they believed they were conversing with a “virtual human” (Lucas, 2014). Based on this study, there is evidence for combining the benefits of face-to-face interviews and medical questionnaires. With further development, artificial intelligence voice assistants can potentially provide human-like rapport that encourages patients to share while removing hesitation from fear of judgment.

VA technology is also compelling for serving senior citizens and those affected by technological barriers. A study conducted on the viability of the Healthy Coping in Diabetes application, which utilizes a conversational bot downloaded onto the Google Home device, found that 80 percent of senior citizens in the study preferred to use the voice interface software over a smartphone application (Chang, 2018). For seniors who find it difficult to work with smaller screens because of physical barriers such as arthritis and poor eyesight, or because of the complexity of mobile applications, having this technology available over text interfaces increases ease of use for screening and survey processes. The functionality of voice assistants can also be expanded to serve hard-to-reach patients, for example, those who would find it more helpful to ask and hear information about medical conditions through voice commands rather than pursuing medical websites for reliable information. Especially when considering the scarcity of stable healthcare in low-income countries, virtual healthcare has the potential to overcome global barriers to access.

In a clinical setting, VA also has great potential for decreasing physician burden on completing routine tasks and can expedite care. Instead of having a nurse or physician ask screening or standard exam questions, VAs can communicate with patients and update their electronic health records (Sezgin, 2020). Coupled with triaging applications from chatbots, the use of VAs in the hospital setting can address hospital demand. However, although VA technology has the potential to bridge public health gaps in care, there should still be hesitation surrounding current technology.

In light of the surge of COVID-19 cases, healthcare professionals and policymakers must mediate the tension between adopting AI technology designed to address the emergency and ensure that new interventions do not spread misinformation. Casey Ross, a National Technology

Correspondent for STAT Health Tech, tested eight different chatbots for their reliability in providing advice for COVID-19 symptoms. Although Ross's team provided the same set of symptoms to each AI bot, they received a variety of responses ranging from low risk of COVID-19 infection to being advised to "start home isolation immediately" (Ross, 2020). Ross's experience with commonly used chatbots reveals the current limitations of VAs, especially during a public health crisis. With rapidly changing information about COVID-19 symptoms and varying professional opinions on how patients should be treated, it is imperative for the health of the nation for chatbot programmers to stay current on new medical information.

Predictive Bias in Healthcare AI

As AI technologies become more prevalent in the healthcare field, patients and providers should remain cautious about potential biases that may arise. In "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations," researchers revealed hidden algorithmic biases in AI software used to predict which patients would benefit from enrollment in additional healthcare programs (Obermeyer, 2019, 9). The algorithm that was implemented in hospitals nationwide was trained on past data to generate a risk score for each patient. Patients who were placed in the 97th percentile were automatically identified by the algorithm for enrollment in a program that would provide additional health resources. However, after calculating a "comorbidity score" by race that would give an overview of each patient's overall health, researchers found that, "at the same level of algorithm-predicted risk, Blacks have significantly more illness burden than Whites." (2) Simply put, at the 97th percentile, black patients suffered from 26.3 percent more chronic illnesses than white patients. Moreover, if the algorithm simulation did not contain a predictive gap between blacks and whites, "the fraction of Black patients would rise from 17.7 to 46.5%" (2). This discrepancy points to a serious issue where a significant proportion of Black patients were not afforded the same access to healthcare programs and potentially increased their susceptibility to preventable illnesses.

Why does this discrepancy exist? Upon examination of the algorithm, researchers found that the discrepancy was caused by issues with label choice. Algorithm programmers made the assumption that those who incur the highest healthcare costs would be the ones who benefit the most from additional health programs. At face value, this reasoning makes sense, as health care costs and health needs are closely correlated. In fact, using "cost" as a predictor is an industry-wide approach implemented by many academic groups and non-profit organizations (551). However, hidden in this assumption is the fact that in general, black patients spend less on healthcare than white patients, holding the number of chronic illnesses constant (550). Even though all participants in the study were enrolled in health insurance, other barriers to health, such as transportation, job demand, and knowledge of reasons to seek care, may prevent black patients from spending on care. Moreover, as a result of historical instances where the healthcare system failed to protect its patients, data shows that in general, black patients have less trust in physicians (550).

The findings in this study illuminate the importance of well-chosen labels in AI algorithms. As researchers comment, "Health is, by nature, holistic and multidimensional, and there is no single, precise way to measure it" (50). Choosing the right label to predict patients' healthcare needs is

no easy task. However, some approaches can point researchers in the right direction. First, take into account structural and historical inequalities that may introduce algorithmic bias. Since predictive algorithms are trained on past data, the labels chosen are often reflective of structural inequalities. The algorithm "specifically excludes race," but as seen by various racial and socioeconomic factors that prevent black patients from receiving the equal quality of care, this color-blind approach to health care may not be the best. To avoid predictive bias in the future, programmers should assess cultural fit and the adaptivity of the algorithm through community engagement (Chou, 2021). Programmers need to conduct user research through discourse with community leaders and members to catch potential discrepancies in their algorithms.

Another possible solution is for greater transparency in how AI processes data. Although it is understandable that AI companies want to protect the inner workings of their programs from competitors, and that there are privacy concerns when it comes to health data, there should still be an effort to overcome the "black box" issue. The EU General Data Protection Regulation (GDPR) defines informed consent as the right for patients to know how AIs make their diagnostic decisions (Dalton-Brown, 2019). This begins with programmers making training data, objective function, and prediction methodology more accessible to the public and/or other researchers (Dalton-Brown, 2019).

This finding reveals a serious warning for using algorithms in isolation without expert input. If providers blindly relied on the algorithm to identify patient needs, a large population of black patients would not have access to the health services they need. Thankfully, for this algorithm, patients at the 55th percentile were referred to a doctor for further consideration (Obermeyer, 452). Even then, doctors are not without their own biases. Researchers found that though doctors' input partially redressed the algorithmic biases, a reconstructed algorithm created to address the bias performed better. Predictive screening algorithms can be an extremely powerful tool that has the potential to overcome human bias. However, as it pertains to healthcare, the industry must be cautious when adopting new AI-based technology.

Conclusion

The development and integration of artificial intelligence into our healthcare system has the potential to overcome barriers to access, lower cost of care, and expedite care delivery. As Dr. Kim from Aravind Eye Hospital commented, "Even one person screened with AI is incredible—it means one person we can help prevent from going blind" (Peng, 2020). For many people, seeking treatment and obtaining good health seems like a privilege only available to those who can afford it, and for many physicians, burn-out is the norm as they are overwhelmed by the sheer volume of patients to see, data to sift through, and routine tasks to perform (Peng, 2020). Artificial intelligence can address those issues and beyond. However, healthcare providers and patients must be vigilant in ensuring that the biases of society past and present do not affect the health outcomes of future generations. By seeking regulatory standards for transparent algorithms and inclusive datasets, artificial intelligence can empower global citizens to live happier, healthier lives.

References

Ardila, Diego, Atilla P. Kiraly, Sujeeth Bharadwaj, Bokyung Choi, Joshua J. Reicher, Lily Peng, and Daniel Tse. 2019. “End-to-end Lung Cancer Screening with Three-dimensional Deep Learning on Low-dose Chest Computed Tomography.” *Nature Medicine* 25:954–961. doi.org/10.1038/s41591-019-0447-x.

Bresnick, Jennifer. 2018. “Top 5 Use Cases for Artificial Intelligence in Medical Imaging.” *Health IT Analytics*, healthitanalytics.com/news/top-5-use-cases-for-artificial-intelligence-in-medical-imaging.

Bodenstedt, Sebastian, Martin Wagner, Beat Peter Müller-Stich, Jürgen Weitz, and Stefanie Speidel. 2020. “Artificial Intelligence-Assisted Surgery: Potential and Challenges.” Visceral Medicine 36(6): 450-455. doi.org/10.1159/000511351.

Cheng, Amy, Vaishnavi Raghavaraju, Jayanth Kanugo, Yohanes P. Handrianto, and Yi Shang. 2018. “Development and Evaluation of a Healthy Coping Voice Interface Application Using the Google Home for Elderly Patients with Type 2 Diabetes.” 2018 15th IEEE Annual Consumer Communications & Networking Conference. doi: 10.1109/CCNC.2018.8319283.

Chou, Katherine. 2021. “AI in Healthcare.” Massachusetts Institute of Technology, www.youtube.com/watch?v=cvXVK8oqU4Q.

Dalton-Brown, Sally. 2019. “The Ethics of Medical AI and the Physician-Patient Relationship.” Cambridge University Press 29(1): 115-121. doi.org/10.1017/S0963180119000847.

Jadczyk, Tomasz, Wojciech Wojakowski, Michal Tendera, Timothy D. Henry, Gregory Egnaczyk, and Satya Shreenivas. 2021. “Artificial Intelligence Can Improve Patient Management at the Time of a Pandemic: The Role of Voice Technology.” *Journal of Medical Internet Research* 23(5). doi.org/10.2196/22959.

Lucas, Gale M, Jonathan Gratcha, Aisha King, and Louis-Philippe Morency. 2014. “It’s Only a Computer: Virtual Humans Increase Willingness to Disclose.” *Computers in Human Behavior* 37:94-100. doi.org/10.1016/j.chb.2014.04.043.

Medical Economics Staff. 2021.“Top Challenges 2021: #1 Administrative Burdens and Paperwork.” *Medical Economics Journal* 98(1). www.medicaleconomics.com/view/top-challenges-2021-1-administrative-burdens-and-paperwork.

Nanalyze. 2020. “9 AI Healthcare Startups Deploying Virtual Assistants.” www.nanalyze.com/2020/02/9-ai-healthcare-startups-deploying-virtual-assistants/.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 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.

Pandey, Suresh K. and Vidushi Sharma. 2018. “Battling the Emerging Epidemic of Diabetic Retinopathy.” *Indian Journal of Ophthalmology* 66(11): 1652-53. doi.org/10.4103/ijo.IJO_1681_18.

Peng, Lily. 2020. “Democratizing Healthcare with AI.” Tedx Talks, www.youtube.com/watch?v=MNp26DgKxOA.

Ross, Casey. 2020. “I Asked Eight Chatbots Whether I Had Covid-19. The Answers Ranged from ‘Low Risk’ to ‘Start Home Isolation.’” STAT, www.statnews.com/2020/03/23/coronavirus-i-asked-eight-chatbots-whether-i-had-covid-19/.

Sezgin, Emre, Yungui Huang, Ujjwal Ramtekkar, and Simon Lin. 2020. “Readiness for Voice Assistants to Support Healthcare Delivery During a Health Crisis and Pandemic.” *npj Digit. Med* 3(122). doi.org/10.1038/s41746-020-00332-0.