

A Guide to AI Challenges and Barriers

Rosemarie Burynski¹; Bernice L. Hausman²

1. Penn State College of Medicine, Hershey, Pennsylvania.
2. Department of Humanities, Penn State College of Medicine, Hershey, Pennsylvania.

Correspondence: rburynski@pennstatehealth.psu.edu

Conflict of interest: The authors declare no potential conflicts of interest.

Keywords

artificial intelligence; social challenges; technology; medical bias; insurance

Abstract

Artificial intelligence (AI) tools are developing quickly and prominently within the U.S healthcare system. It therefore seems essential to understand their weaknesses to practice medicine that is fully informed. The goal of this paper is to overview several key concerns surrounding healthcare AI, as well as some anticipated barriers to its implementation. AI's generalizability is currently limited due to a widely fragmented Electronic Health Record (EHR) and inaccessibility to training data. AI tools are therefore at risk of acting on incomprehensive knowledge and generating inaccurate outputs. They are also extremely susceptible to several different forms of bias. Such bias can result in preferences towards diagnosing some diseases over others and recommending interventions that are only beneficial to certain populations. Privacy and transparency are also of great concern, especially when dealing with private medical data. While "black box" algorithms are criticized for their lack of transparency, innovators are working towards explainable AI (XAI) tools that can "show their work." Developing guidelines make it difficult to predict how liability for AI malpractice may be distributed across parties but has interesting implications for how physicians will change their practice in response. Finally, the current U.S. payment structure does not easily accommodate

healthcare AI tools. This challenge raises questions surrounding healthcare AI's reimbursement mechanism as it becomes more widely utilized. While this paper does not provide solutions for the outlined concerns, it emphasizes the importance of understanding and anticipating the shortcomings of new healthcare technologies.

Introduction

As artificial intelligence (AI) becomes increasingly prevalent, so do concerns regarding its ability to accurately and equitably supplement the medical field. Therefore, it is the duty of providers to be aware of both the benefits and harms that healthcare AI may pose towards their patients. This guide overviews some major talking points surrounding healthcare AI's anticipated challenges and implementation barriers.

Generalizability

A major concern within the field of AI research is the generalizability of results. There are a few reasons for this. One is that there is no universal Electronic Health Record (EHR) within the U.S. Most current healthcare AI is supervised, meaning that it requires training on large data sets that are labeled by humans. This training is difficult to accomplish when health data is scattered throughout different

systems. Many programs must settle for smaller amounts of training data. This limitation raises questions regarding the applicability of those systems to larger or different populations of patients. This concern is also prevalent on a smaller scale - AI with great performance using data from one U.S. hospital may fail in another U.S. hospital due to its lack of generalizability (1). One exception to the fragmented American EHR is the VA, the country's biggest integrated healthcare system (2). The vast amount of data stored within the VA EHR makes for an ideal AI training ground. However, the data demographics still leave plenty of room for debate - How generalizable is VA data to the rest of the country?

AI training seems to risk spectrum bias, which refers to a test that is performed and evaluated within a population that is different from the intended population (3). This incomprehensive training leads to limited, incomprehensive knowledge. If medical students are only taught to recognize and treat diseases that are prevalent in their school's state, they will misdiagnose and mistreat the vast array of diseases that are less geographically common. If they are taught only to identify infections on lighter skin tones, they are more likely to miss presentations on darker skin tones. In fact, they may fail to identify the presence of disease altogether because they have had no exposure to it in training. The same problem applies to AI.

Other Biases

In general, a program that is trained on biased data will inherit and reinforce inequalities in its own algorithms. Echoing the concerns of generalizability is diagnosis bias. COVID-19 diagnostic tools that are trained in the U.S. may not have much exposure to lung-related diseases such as tuberculosis and types of pneumonia associated with HIV/AIDS that are more prevalent in other countries. The algorithm

then runs the risk of misdiagnosing these diseases as COVID-19 due to their similarities and lack of knowledge of their differences (4).

Bias may also be introduced in disease modeling scenarios. A big mitigation initiative during the peak of the COVID-19 pandemic was disease "mapping" to track spread. These modeling techniques often require specified data inputs that are more challenging to obtain from underrepresented populations. Additionally, the models were used in recommending interventions (such as quarantining and social distancing) that were a lot less attainable in crowded and/or poor sanitary environments. Similarly, treatment selections made through AI tools were less likely to account for social determinants of health that are underrepresented and less accurately documented in the EHR (4).

Infodemic bias is especially prevalent in the age of social media and mass information. AI has increasingly been used to help fact-check and combat the spread of misinformation. However, this integration is mostly done within "easy to mine" data sources such as Twitter and Facebook. While it may be having a positive impact on these platforms, other information sources such as radio and TV are less likely to be fact-checked by AI. Simultaneously, these alternate channels may act as primary information sources for certain countries and populations (4).

Privacy, Transparency and Mistrust

By design, medical AI is going to guide and influence clinical decision-making. As always, it is important for a physician to be able to explain how and why a recommendation is made. This process is made more complex when considering the "black box" tendencies of certain AI technologies. The term "black box" is used to refer to the lack of transparency regarding AI output. It is not always as clear how an AI algorithm generated a particular

output given a specific input. This lack of clarity can lead to an overall mistrust of AI, and especially so within the healthcare field. How can a physician use the medical advice of an algorithm when its reasoning is not available?

These concerns have led to the push for explainable AI (XAI) that is able to reason through its output in an understandable way. There are a wide variety of XAI methods, many of which include visual representations (decision trees, graphs, etc.) of the decision-making process (5). These methods ensure that when a clinician inputs data, they receive the output they are looking for and an explanation for that output. The clinician can use this information in their own professional evaluation of the AI's performance before utilizing its advice in their decision-making process. Given the importance of physician transparency, it will be unsurprising to see XAI continue to grow throughout the field.

Rules, Regulations, and Malpractice

AI has been leveraged by developers through the promise of its increased accuracy, and hence, its ability to reduce medical mistakes. However, the rules and regulations surrounding responsibility for AI mishaps are still developing. Historically, physicians are typically liable for their actions even when under the influence of third-party information. For example, if a physician follows an insurer's recommendation for a procedure plan and harm occurs, the physician is still responsible. Similarly, the physician is responsible for appealing coverage denials if they believe a service to be medically necessary (2). In the case of AI, then, it would seem that physicians should maintain liability for all decision-making. However, the lack of clear guidelines may still leave room for liability to be potentially shared by AI producers. This possibility combined with the hope of increased accuracy puts malpractice in an interesting economic position.

On the one hand, malpractice pressure increases the demand for AI. Physicians may shift towards "defensive medicine," through which they'd push certain decision-making onto AI in hopes of avoiding some amount of liability. This increased demand for AI would result in raised AI prices. Competition increases and product differentiation decreases. Subsequently, AI producers would have to rely more on prices in order to compete and would likely reduce their prices in response. Therefore, malpractice seemingly has two opposite effects on price-setting and profit-making⁶. While the long-term economic effect is still up for questioning, it does certainly raise social concerns regarding how malpractice will affect physician reliance on AI.

What about insurance?

The already complicated world of health insurance is made more complex when considering how AI can or should be billed for. In general, medical procedures and services are defined by Current Procedural Terminology (CPT) codes that are developed by the American Medical Association (AMA). These codes are divided into three categories (7):

- Category I → Describes a procedure or service that must meet specific criteria. These codes are typically reimbursed by both Medicare and commercial payers.
- Category II → Used for tracking and performance measurement purposes. These codes are not generally reimbursed by Medicare or commercial payers.
- Category III → Codes for developing technology, services, and procedures. These codes are temporary and may be later placed in Category I if the criteria is met. While there are no fees assigned to these codes, reimbursement may be available on a case-by-case basis (8).

To be billed for, a specific AI technology must fall into a CPT category and be defined by a CPT code. AI does not fit well into this type of payment structure. New AI technologies are performing countless and various tasks. It would take a long and tedious amount of time to create a special CPT code for each (7). Simultaneously, one CPT code cannot overlap with another. This rule provides a unique challenge for AI, as many algorithms do work that has historically been performed by humans and is likely already defined by an existing CPT code. For example, an AI tool that detects pulmonary hypertension in medical imaging is performing work that is already covered by CPT code CPT71275: CT angiography, chest (noncoronary) w/ contrast material(s), including noncontrast images, if performed, and image postprocessing (7). One

suggestion has been to bundle AI services with their complementary services (in this case, the new pulmonary hypertension AI tool and the already-existing imaging service would be bundled) (9). However, it is still unclear as to what billing trajectory AI will end up following.

Conclusion

The rapid rise of healthcare AI promises big change for the medical field. The associated challenges outlined in this paper are non-exhaustive and subject to change over time. Still, understanding them will guide learning and decision-making when implementing new AI tools into the healthcare system.

References

1. Cossio, M., & Gilardino, R. E. (2021). Would the Use of Artificial Intelligence in COVID-19 Patient Management Add Value to the Healthcare System? *Frontiers in Medicine*, 8. <https://doi.org/10.3389/fmed.2021.619202>
2. Agrawal, A., Gans, J., Goldfarb, A., & Tucker, C. (2024). *The Economics of Artificial Intelligence*. University of Chicago Press.
3. Gupta, A., Slater, J., Boyne, D. J., Mitsakakis, N., Béliveau, A., Drużdżel, M. J., Brenner, D. R., Hussain, S., & Arora, P. (2019). Probabilistic Graphical Modeling for Estimating Risk of Coronary Artery Disease: Applications of a Flexible Machine-Learning Method. *Medical Decision Making*, 39(8), 1032–1044. <https://doi.org/10.1177/0272989x19879095>
4. Luengo-Oroz, M., Bullock, J., Pham, K. H., Lam, C. S. N., & Luccioni, A. (2021). From Artificial Intelligence Bias to Inequality in the Time of COVID-19. *IEEE Technology and Society Magazine*, 40(1), 71–79. <https://doi.org/10.1109/mts.2021.3056282>
5. Sarp, S., Catak, F. O., Kuzlu, M., et al (2023). An XAI approach for COVID-19 detection using transfer learning with X-ray images. *Heliyon*, 9(4), e15137–e15137. <https://doi.org/10.1016/j.heliyon.2023.e15137>
6. Chopard, B., & Musy, O. (2023). Market for artificial intelligence in health care and compensation for medical errors. *International Review of Law and Economics*, 75, 106153. <https://doi.org/10.1016/j.irle.2023.106153>
7. Smetherman, D., Golding, L., Moy, L., & Rubin, E. (2022). The Economic Impact of AI on Breast Imaging. *Journal of Breast Imaging*, 4(3), 302–308. <https://doi.org/10.1093/jbi/wbaco12>
8. Dotson, P. (2013). CPT® Codes: What Are They, Why Are They Necessary, and How Are They Developed? *Advances in Wound Care*, 2(10), 583–587. <https://doi.org/10.1089/wound.2013.0483>
9. Zink, A., Chernew, M. E., & Neprash, H. T. (2024). How Should Medicare Pay for Artificial Intelligence? *JAMA Internal Medicine*, 184(8). <https://doi.org/10.1001/jamainternmed.2024.1648>