ARTIFICIALLY INTELLIGENT?
Machine Learning in Healthcare and Why It May Not Be As Advanced As You Think

By Avishek Choudhury, PhD & Caitlyn Allen, MPH

DOI: 10.33940/001c.77632
Submitted: May 11, 2023 / Accepted: May 11, 2023

Machine learning: What exactly is it, and how is it being used in healthcare? Are machines always better than a person? How do we know? Managing editor, Caitlyn Allen, sat down with Dr. Avishek Choudhury, artificial intelligence healthcare researcher, to answer these questions and more.

*Corresponding author
West Virginia University
Patient Safety Authority
Disclosure: The authors declare that they have no relevant or material financial interests.
Caitlyn Allen: How do you define artificial intelligence [AI] and what are the different types?

Dr. Avishek Choudhury: AI is any technology mimicking how humans think and process information: any technology that can mine data, understand patterns, and then propose a conclusion based on previous experiences.

Much of your research surrounds AI use in healthcare. Is this a relatively new phenomenon or something that we’ve been using in healthcare for a while?

While AI’s potential in healthcare is impressive, its implementation remains largely in the research stage. Very few practitioners have fully integrated AI into their clinical routines. This limited adoption is driven by several key concerns, including accountability, the risk of over-reliance, and challenges with usability.

Firstly, the question of accountability is paramount. When an AI aids in decision-making, who bears the responsibility for that decision, particularly if the outcome is not favorable? This is an issue that hasn’t been thoroughly addressed yet. Secondly, there’s the fear of blind trust. While AI has shown promise in data processing and pattern recognition, it’s critical to remember that these systems are not infallible. They rely on the quality and accuracy of the data they’re given. There’s a danger that over-reliance on AI may lead to overlooking its limitations and potential errors. Thirdly, usability is another substantial hurdle. The integration of AI into existing workflows in a manner that is seamless and user-friendly remains a challenge. Moreover, there’s a lack of comprehensive research detailing the safe and effective integration of AI into clinical workflows.

We have seen numerous studies demonstrating the performance of specific algorithms in research settings. Yet, there’s a dearth of information on how these tools impact patient outcomes when implemented in the chaos of a real-world clinical environment. The gap between AI’s potential and its real-world application in healthcare is substantial. We have yet to address these concerns fully or explore how AI can be safely and effectively used by doctors and nurses—those on the front lines of patient care—in their day-to-day operations.

When an AI aids in decision-making, who bears the responsibility for that decision, particularly if the outcome is not favorable?

The perception often associated with AI is that it transcends human capabilities, demonstrating almost miraculous abilities. However, this isn’t the entire truth. AI shines in pattern recognition and data processing—areas where the cognitive workloads far exceed what humans can handle. This edge doesn’t make AI superior to humans, but simply more efficient in certain tasks, akin to a crane being able to lift heavier weights than humans.

Where AI truly demonstrates its value is in its ability to enhance human efforts. In healthcare, for instance, AI can be an invaluable tool. Imagine the case of a patient with multiple comorbidities, devoid of insurance, and without familial support. To plan optimal treatment in such cases, the amount of information to process is immense. This is where AI steps in, distilling vast amounts of data and providing healthcare professionals with a more manageable, efficient decision-making process. Of course, when we talk about raw speed and consistency in processing, AI excels. Yet, it’s important to understand that AI’s performance is tethered to its predefined parameters and quality of data. AI requires explicit directives about intended outcomes in different scenarios.

Now, consider the scenario of a healthcare professional at the end of a grueling 12-hour shift. Under such circumstances, fatigue might cloud their judgment or lead to oversights. AI, on the other hand, doesn’t experience fatigue. It remains consistent in its operations and won’t overlook critical patient information, such as high blood pressure, due to exhaustion. Therefore, if we were to characterize AI as “better,” it would be in this context: AI excels not by superseding humans but by amplifying human capabilities and mitigating human errors, especially in high-pressure environments like healthcare.

It’s often assumed that a computer will automatically outperform a human, but it sounds like that has not really been tested.

Correct. The perception often associated with AI is that it transcends human capabilities, demonstrating almost miraculous abilities. However, this isn’t the entire truth. AI shines
What about areas where humans often outperform AI?

Indeed, while AI excels at pattern recognition and data processing, its abilities fundamentally depend on the data it’s been trained on. Therefore, in scenarios where there’s no preexisting data, humans will often outperform AI. For instance, consider a blood test that reveals a hitherto unknown pattern. AI would be at a loss, unable to classify this pattern without a reference point. For AI to recognize, “This is disease X,” it would need a vast amount of data—perhaps thousands of data points—highlighting different scenarios that lead to disease X. Only then can AI learn and identify the associated patterns. Given a new pattern, the AI could potentially relate it to a known pattern, suggesting, “This pattern resembles one seen 20 years ago, and there’s a 60% likelihood that it corresponds to disease X.” However, should disease X be a novel or rare condition, the AI will be unable to reliably identify it. It can only highlight the presence of a new pattern. The onus then falls on a human clinician to deduce what this novel pattern signifies. This underlines one of the inherent limitations of AI—it can’t independently discover or invent, but rather is fundamentally reliant on the data it’s been trained on. This is why the human element in healthcare will always be essential, to interpret and investigate when AI encounters the unknown.

What kind of quality control, if any, goes into vetting the data?

It depends on whoever created the AI. A developer can ask the algorithm to weigh sources differently, but it’s optional. Some automatically assign higher weights to some data points, but many do not.

Could that cause the algorithms to eventually become biased?

AI algorithms are inherently neutral—they are mathematical constructs devoid of bias. However, if they are trained on biased data, they will reflect and propagate that bias.

For example: An AI system trained on data from Indian patients, many of whom have a diet rich in spicy foods, might develop an association between being Indian and having digestive issues. If the data predominantly showcases such cases, the algorithm may erroneously predict digestive problems even for an Indian patient who doesn’t consume spicy food. It’s important to clarify that this isn’t a bias in the algorithm, but rather a reflection of bias in the data.

Similarly, disparities in healthcare also can influence data and, consequently, AI. In low- and middle-income countries, healthcare access may be limited, and record-keeping may be less digitized, leading to fewer data points for AI to learn from. Furthermore, if there are systemic biases in how different demographic groups are treated, these biases will be reflected in the data, skewing the AI’s analyses accordingly.

And like you said, AI can do well what you tell it to do, but it’s still operating within those constraints.

Correct.

You mentioned in a recent study that drug safety is one of the most common areas in patient safety where AI is being used. Why might that be the case?

Absolutely, drug safety is a prime area where AI research is extensively used because it’s not only safer but also feasible, given the accessibility and nature of the data involved. AI is adept at identifying potential drug interactions—a critical aspect of patient safety. In a typical clinical scenario, a doctor might be unaware of the full range of medications a patient is taking, or a patient might unintentionally omit certain medications during their consultation. This could potentially lead to harmful drug interactions, such as prescribing drug X that negatively interacts with drug Y.

The nature of drug interaction data—structured, comprehensive, and readily accessible—makes it particularly amenable to AI study and implementation. This contrasts with other areas of healthcare that may involve unstructured data or require nuanced human interactions, which are more complex for AI to handle.

However, AI has the capability to overcome this human limitation. AI algorithms can be trained on expansive databases that encapsulate a wide range of possible drug interactions. This allows them to predict and alert healthcare professionals about harmful drug combinations, like drugs X and Y. Every time this drug combination is prescribed for a patient, the AI system would generate an alert, allowing the doctor to adjust the prescription accordingly.

Additionally, the nature of drug interaction data—structured, comprehensive, and readily accessible—makes it particularly amenable to AI study and implementation. This contrasts with other areas of healthcare that may involve unstructured data or require nuanced human interactions, which are more complex for AI to handle. Thus, due to the feasible study design, the accessibility of data, and the tangible impact on patient safety, AI has become an integral tool in enhancing drug safety.

In that paper, you also mentioned that an AI-attributable error might lead to mass patient injuries compared to those attributable to a single provider’s error. Tell me more about that.

Indeed, the broad impact of AI tools in healthcare can potentially amplify errors in a way that is not seen with individual providers. The critical factor here is the scale at which AI operates and the delay in feedback that might occur.
For example, consider an AI-powered recommendation system used by a doctor. This AI tool, even if highly competent, can become biased if exposed predominantly to a specific patient type over a period. Suppose the AI system is self-learning or adaptive; in such a scenario, it might gradually become more tailored to that patient population. Now, when a different patient type presents, the AI system’s recommendation might not be as accurate or appropriate. The doctor, having trusted the AI system over the past months, follows the recommendation and prescribes a particular medication. However, this prescription might not be suitable for the patient, which is not immediately apparent. The impact of the erroneous recommendation may not be detected until the patient has been on the medication for a few weeks. During this lag time, the AI system may have made similar recommendations for other patients. Thus, by the time the initial error is discovered, multiple patients may have received incorrect treatment. This potential for mass patient impact is a unique challenge associated with AI use in healthcare. It highlights the importance of rigorous testing, continuous monitoring, and safeguards to prevent the propagation of errors in AI systems.

Because the AI is the tool that everybody is dipping into. So, if the pool is tainted, it’s worse than just a tainted cup of water.

Correct.

What do you think is the next iteration of AI in healthcare?

Looking ahead, I see several exciting possibilities for the next iteration of AI in healthcare, with three areas standing out: digital twins, mental health diagnosis, and analysis of clinical notes.

A digital twin is essentially a real-time, virtual clone of a patient. With the continuous exchange of data between the patient and their digital twin, we can simulate various health scenarios and interventions. For example, if a patient is experiencing certain symptoms, their digital twin can illustrate what’s occurring within their body and suggest modifications to optimize health. It allows physicians to see a computerized version of the patient’s health status and predict the likely outcomes of different treatment options. For instance, administering medication to the digital twin can simulate the patient’s potential reactions, offering valuable insights for treatment planning.

So, if you’re experiencing chest pains, your doctor can show you what’s causing the pain and what will happen if they give you drug X as a treatment?

Correct. A real-time data transformation. It already exists in manufacturing, but people are working on digital human twins. At the 2023 National Academies workshop on integrated diagnostics, we discussed how AI-based digital twins can help with oncology.

AI also has the potential to revolutionize the field of mental health. By analyzing patterns in speech, language use, facial expressions, and even social media activity, AI could help identify signs of mental health conditions much earlier than currently possible. This could greatly improve the prognosis for many conditions by enabling earlier intervention.

Next, clinical notes are a treasure trove of valuable patient information, but their unstructured nature makes it difficult for healthcare providers to extract insights manually. AI algorithms, particularly those using natural language processing (NLP), can help analyze these notes, identify relevant information, and present it in a structured format for clinicians. This could significantly enhance patient care by making it easier for providers to access and understand a patient’s full medical history.

Taking notes is one of the most time-consuming things any nurse or doctor has to do. There’s a huge opportunity that generative AI [e.g., ChatGPT] can allow doctors and nurses to have more patient time versus time spent documenting.

I suspect that will make clinicians ecstatic if they could spend less time on documentation.

Certainly, easing the burden of documentation could be a significant boon for clinicians, allowing them to devote more time to direct patient care. However, transitioning this from an exciting possibility into a practical reality does require careful consideration of numerous factors, including policy and accountability.

These advancements represent just a few ways AI can further improve healthcare. The key will be to ensure these technologies are developed and deployed responsibly, with patient safety, usability, privacy, and equity always in mind.

The goal should be to build a framework where AI can augment human skills and judgement, improve healthcare outcomes, and do so in a manner that is ethically sound, accountable, and financially sustainable.
In our current healthcare system, if a clinician makes an error, they bear the responsibility and potentially face penalties. This clear line of accountability becomes more complex when AI is involved. If a mistake occurs while using an AI system, should the clinician be held responsible? Or should the blame be attributed to the AI, the developers, or the institutions that implemented it? These are important questions that need answers to promote the safe and effective use of AI in healthcare. Financial considerations also play a crucial role in this. Developing, validating, implementing, and maintaining AI systems in healthcare is an expensive process. Ensuring these systems are reliable, safe, and accountable requires significant investment, which can be a barrier to their widespread adoption.

Addressing these issues will be critical in shaping the future of AI in healthcare. The goal should be to build a framework where AI can augment human skills and judgment, improve healthcare outcomes, and do so in a manner that is ethically sound, accountable, and financially sustainable. But for this to be a real-life thing, there has to be policy and accountability in place, which is not there. If you make a mistake as a person, you get penalized. If you make a mistake while using an AI, regardless of whether it was the AI’s fault, you still get penalized. That’s a big thing people are not working on because of many factors, mostly money.

As you mentioned, AI in healthcare is currently in its research phase. Maybe as this becomes more of a part of our day-to-day experience, the policies will start to catch up. Tell me about “vertical standards” and how they come into play in patient safety.

Indeed, the current landscape of AI in healthcare lacks concrete benchmarks or vertical standards that define the level of accuracy or performance required for different healthcare settings and tasks. Without these, it’s challenging to gauge whether an AI system is “good enough” for use in clinical practice.

For instance, let’s consider an AI tool that has a 90% accuracy rate in detecting drug reactions. Is this satisfactory? Should we deploy this tool in a clinical setting? There are no clear answers to these questions currently. The prevalent trend seems to be a competitive race among researchers to incrementally improve accuracy, but without a defined threshold of acceptability, it remains unclear when an AI tool is ready for clinical use. This situation underscores the need for setting vertical standards in healthcare AI. We need guidelines tailored to the specific requirements of different departments and tasks, as the risk and acceptable margin of error may vary significantly. For example, the acceptable error margin for AI systems analyzing clinical notes might be higher than for those diagnosing critical conditions such as pancreatic cancer. Without these standard benchmarks, it’s challenging to determine the performance level that an AI system should achieve for a specific task to be considered safe and effective for use. Creating these standards will provide much-needed clarity and confidence in deploying AI tools in healthcare, helping to ensure patient safety and optimal care outcomes.

And it sounds like this goes back to quality control: Is 90% more accurate than how a human would perform?

Correct. And that 90% is tested on the research dataset. Will it perform the same in a clinic in Monongalia County in rural West Virginia? We don’t know. It may or it may not.

What about when it comes to scale: using AI on a micro level [e.g., an individual facility] versus on a macro level [e.g., across multiple health systems]?

At the micro level, such as in a single facility or a specific clinical specialty, the use of AI can be highly tailored to the unique needs of that setting. For example, if a clinic primarily serves a particular patient cohort, an AI system could be trained specifically on that population’s data. This would allow the AI to become very adept at understanding that population’s unique health characteristics and trends, leading to potentially higher accuracy and effectiveness. However, it also means that the AI might not perform as well when faced with patient data outside of its training set.

Conversely, implementing AI on a macro level, such as across multiple health systems, allows for the analysis of much larger and diverse datasets. This broad perspective can reveal patterns and trends that would be impossible to discern at a smaller scale, potentially leading to more generalized insights. However, the diversity and complexity of these large datasets can also introduce challenges. There may be numerous missing or inconsistent data points, and variations in how data is collected and recorded across different systems could lead to discrepancies. Moreover, patient privacy and data security become even more critical issues at this scale.

Overall, whether you use AI at a micro or macro level depends on your specific goals and constraints. The key is to carefully consider the unique advantages and challenges of each approach and choose the one that best fits your needs.

And I would think similarly for larger datasets, that it would be important for the information to be uniform and uniformly collected. If the information from different hospitals looks different, that would create a challenge to analyze it.

Absolutely, uniformity in data collection and documentation is crucial for successful AI analysis, especially on a larger scale. The lack of standardized procedures or documentation formats across different healthcare providers or institutions is indeed a significant challenge in healthcare data analysis.

Let’s say we’re dealing with a symptom as common as a stomachache. Different doctors may order different diagnostic tests based on their own medical judgement and experiences. This leads to varied datasets even for the same symptoms, creating a challenge for AI systems that need consistent data to function effectively.

If an AI system is trained on a dataset that includes certain diagnostic tests, but is then implemented in a setting where these tests are not typically conducted, this could lead to incomplete data inputs. The AI system may not perform optimally in this new setting due to the missing data.

Therefore, to maximize the effectiveness of AI systems in healthcare, efforts should be made to standardize data collection and documentation practices across different healthcare providers. This would ensure that AI systems are trained and tested on datasets that accurately reflect the diversity and complexity of real-world healthcare scenarios, increasing their robustness and generalizability.
To maximize the effectiveness of AI systems in healthcare, efforts should be made to standardize data collection and documentation practices across different healthcare providers. This would ensure that AI systems are trained and tested on datasets that accurately reflect the diversity and complexity of real-world healthcare scenarios, increasing their robustness and generalizability.

And when you have those holes across multiple patients in multiple hospitals, that’s another way that the data could get biased?

Correct. That’s the difference between real-world data versus university-collected data. Datasets collected for research from research institutes are good because everything is there. But in most places, that might not be the case.

Most healthcare data involves protected information. How secure is the information fed into the algorithms?

One of the foundational tenets of data privacy in healthcare is that individual patient data is de-identified before being used for analysis or machine learning. This process ensures that the algorithm, on its own, cannot identify individuals from the data it is analyzing. However, the potential for re-identification, especially with the presence of unique characteristics or outliers, is a complex and sensitive issue.

Consider the example where we have a patient cohort largely composed of South Asians with one exception of an East Asian. If the ethnic background data was used in the model, the lone East Asian patient could potentially be re-identified, especially if the users of the AI system have access to the original data source. In this respect, it’s crucial to have robust data privacy policies and technologies in place to protect individuals’ health information. This is especially important as we leverage AI and machine learning more in healthcare, where the use of large and diverse datasets is integral. Sometimes, certain personal information may be necessary for tailoring healthcare services to an individual’s needs. For instance, knowing whether a patient has insurance could enable an AI system to recommend treatments that the patient can afford. In such cases, patients should be clearly informed about how their data will be used and protected, and their consent should be obtained. This way, we can strike a balance between personalized healthcare delivery and data privacy.

That goes back to the eventual need to reconcile policy. Is there anything else about AI or machine learning that we didn’t cover?

Trust, workload, and accountability. Take workload. Developers don’t often understand the end user and their digital literacy. When we see a healthcare provider who may be struggling with Epic [electronic health record software] and then, if you add another AI module in that already complex software, it overcomplicates things for that end user. Instead of reducing the workload, it’s an extra thing they’re doing. You must consider, if there is an AI that works well, how do you integrate it in the clinical workflow? You cannot just disrupt everything that’s going on and then say, “Okay. From tomorrow we’ll do this.” That will not go well. That ties back to trust.

If you’re using something that’s working well and you are blindly trusting it and then something goes wrong and there’s a disaster, you’ll stop trusting it. Or you just don’t trust AI because of all the myths and hypes, you’ll miss that opportunity to use that good technology. So, we need balance and some policy to build around that.

And accountability. If the end user is responsible for everything, why would that person use AI? What’s the point? If anything goes wrong, then their license will be at stake. So, why invest and learn a new technology if there is no reward for the end user?

That makes a lot of sense. Well, it sounds like AI is going to continue to play a supporting role in healthcare, at least for a while. Do you think that will always be the case?

The notion of AI replacing human roles in healthcare is a complex and nuanced issue, primarily due to the importance of accountability. Theoretical scenarios where AI could replace a nurse or a doctor hit a wall when we consider liability related to medical errors. In our current understanding, if something goes wrong, we trace it back to the source—which could lead us to the data, the algorithm, the data collecting agency, and the AI developer. These organizations are often large entities, and imposing accountability on them could lead to many complications.

Thus, the reality is that AI in healthcare will likely play a significant supportive role rather than a replacement one. By supplementing human decision-making with AI, we hope to boost efficiency, reduce human errors, and free up valuable time for healthcare professionals. This way, doctors and nurses can focus on more nuanced aspects of patient care or perhaps even devote more time to innovation and discovery.

The key will be to ensure AI tools are reliable, accurate, and accountable, and that they are used in a way that enhances the role of healthcare professionals rather than attempts to replace them. This balanced approach will likely yield the greatest benefits for healthcare providers and patients.

Staffing shortages in healthcare are pervasive, but perhaps AI may be a way to free up clinicians’ time and provide a stopgap.

Correct. It can help the doctor integrate all the information and summarize it. It can help the patients to learn about what’s going on with their health. It can help us identify a patient who is prone to commit suicide, for example. It’s very common in a cancer setting when you deliver a diagnosis, and you see that patient went home and committed suicide. There have been cases like that for pancreatic or liver cancer, which are high-risk and have

“
higher mortality rates. They often commit or attempt suicide. AI can be used to identify those at-risk patients based on their brain activity or facial expression. It’ll never be a replacement, because the patient wants a doctor, and everything is around the patient. If future patients say, “I don’t want a doctor,” then maybe. It depends on what patients need.

Are there any other less-obvious uses for AI?

Identifying burnout in healthcare workers. That can then reduce human error. Consider nurses working 12 hours who then go home, with a one-hour travel time, then sleep for four hours, wake up, one-hour travel time, back on duty three times or four times a week. AI can identify those nurses or doctors who are prone to human error because they’re too tired.

You mean like looking at schedules to identify where there might only be limited time for sleep?

Not just schedules, but actually the people there. If we can link AI to a smartwatch or something like that, you can monitor heart rate, rate of perspiration. There are EEG [electroencephalography] monitors, glasses that analyze pupil dilations and facial expression, and could say, “This person is tired, the brain is not working as well. Maybe he should be given a two-hour break.” AI can do that. It’s very simple, because everything exists, you just have to adapt and use it.

Do we have existing datasets for this type of thing, or would we need to build them first?

It can be done in parallel because some of the things that identify burnout are known. If heart rate is elevated, then you know that person is anxious. If the EEG signals a certain pattern, you know that brain is not functioning well. AI can be used to detect that. That’s it. You don’t need to train anything because you’re not predicting anything. You’re just saying, “The brain activity for this doctor is 10% lower.” Then the manager or attending will be able to identify potential healthcare workers who are more prone to commit an error because they’re too tired.

That is fascinating. What about near misses? We often learn the most by trying to determine why something did not occur.

AI has significant potential in learning from “near misses” in healthcare. Near misses, or close calls that could have resulted in harm but didn’t, are a gold mine of information because they provide insights into areas of vulnerability that otherwise might not be noticeable.

Reference


About the Authors

Dr. Avishek Choudhury, an assistant professor at West Virginia University, is renowned for his contributions to the field of systems engineering. He focuses his research on the intersections of patient safety, artificial intelligence (AI), cognitive human factors, neuroergonomics, and clinical decision-making. He is a pioneer of efforts to incorporate the principles of human factors and systems thinking to seamlessly integrate AI into clinical workflows. His research delves into the complex world beyond mere algorithmic performance, striving to humanize technology to enhance usability and drive adoption. Dr. Choudhury also collaborates with international nonprofits to achieve the global Sustainable Development Goals; his work targets public health issues and digital divide concerns, particularly in low-income countries, leveraging the transformative power of human factors science and AI.

Caitlyn Allen (caiallen@pa.gov) is director of External Affairs for the Patient Safety Authority and managing editor for Patient Safety, the PSA’s peer-reviewed journal. Before joining the PSA, she was the project manager for Patient Safety at Jefferson Health, where she also was the only nonphysician elected to serve on the House Staff Quality and Safety Leadership Council.

This article is published under the Creative Commons Attribution-NonCommercial license.