AMA Journal of Ethics®
November 2020, Volume 22, Number 11: E945-951

STATE OF THE ART AND SCIENCE: PEER-REVIEWED ARTICLE

# How Might Artificial Intelligence Applications Impact Risk Management?

John Banja, PhD

**Abstract**

Artificial intelligence (AI) applications have attracted considerable ethical attention for good reasons. Although AI models might advance human welfare in unprecedented ways, progress will not occur without substantial risks. This article considers 3 such risks: system malfunctions, privacy protections, and consent to data repurposing. To meet these challenges, traditional risk managers will likely need to collaborate intensively with computer scientists, bioinformaticists, information technologists, and data privacy and security experts. This essay will speculate on the degree to which these AI risks might be embraced or dismissed by risk management. In any event, it seems that integration of AI models into health care operations will almost certainly introduce, if not new forms of risk, then a dramatically heightened magnitude of risk that will have to be managed.

*To claim one* AMA PRA Category 1 Credit™ *for the CME activity associated with this article, you must do the following: (1) read this article in its entirety, (2) answer at least 80 percent of the quiz questions correctly, and (3) complete an evaluation. The quiz, evaluation, and form for claiming* AMA PRA Category 1 Credit™ *are available through the* AMA Ed Hub™.

## AI Risks in Health Care

Artificial intelligence (AI) applications in health care have attracted enormous attention as well as immense public and private sector investment in the last few years.[1] The anticipation is that AI technologies will dramatically alter—perhaps overhaul—health care practices and delivery. At the very least, hospitals and clinics will likely begin importing numerous AI models, especially "deep learning" varieties that draw on aggregate data, over the next decade.[2]

A great deal of the ethics literature on AI has recently focused on the accuracy and fairness of algorithms, worries over privacy and confidentiality, "black box" decisional unexplainability, concerns over "big data" on which deep learning AI models depend, AI literacy, and the like.[3,4] Although some of these risks, such as security breaches of medical records, have been around for some time, their materialization in AI applications will likely present large-scale privacy and confidentiality risks. AI models have already posed enormous challenges to hospitals and facilities by way of

cyberattacks on protected health information, and they will introduce new ethical obligations for providers who might wish to share patient data or sell it to others.[5] Because AI models are themselves dependent on hardware, software, algorithmic development and accuracy, implementation, data sharing and storage, continuous upgrading, and the like, risk management will find itself confronted with a new panoply of liability risks. On the one hand, risk management can choose to address these new risks by developing mitigation strategies. On the other hand, because these AI risks present a novel landscape of risk that might be quite unfamiliar, risk management might choose to leave certain of those challenges to others. This essay will discuss this "approach-avoidance" possibility in connection with 3 categories of risk—system malfunctions, privacy breaches, and consent to data repurposing—and conclude with some speculations on how those decisions might play out.

## System Malfunctions

Every human performance specialist knows that the introduction of a novel, powerful, and complex technology into an already complex and dynamic workspace presents a ripe opportunity for errors and system breakdowns.[6] It is bad enough when computerized systems go down in health care facilities. AI-involved crashes or malfunctions might prove much worse. AI forecasters predict that clinicians will eventually come to rely heavily on AI applications, which, over time, will likely become thickly integrated with coding, billing, medical records, scheduling, contracting, medication ordering, and administrative functions.[7] It is easy to imagine how a breakdown or virus affecting any one element of an AI chain could wreak havoc with the entire system.[8] For example, if AI models ultimately come to schedule patients, interpret laboratory specimens or radiographs, generate a report to the referring entity, and send a bill to the insurer, then a malfunction at any point in this continuum could result in a high volume of errors and adverse events. One is reminded of the 2010 article by Dudzinski and colleagues that examined single-point failures—such as infection control lapses, malfunctioning disinfection technology, laboratory errors, and incompetent clinicians—that went on to affect thousands of patients.[9] Within the past few years, one such single-point failure—weaknesses and vulnerabilities in data storage programs—enabled hackers access to health records, resulting in ransomware crimes and identity theft that affected millions of patients.[10]

Clinicians have only to reflect on their day-to-day experience with information technology and its frequent breakdowns—eg, disabled access to servers, computerized systems that freeze up, programs that are hard to navigate or easy to misuse, malware attacks—to appreciate how vulnerable workflow (and the liabilities that attach to it) could become to AI malfunctions. Moreover, none of these technologies and their related operations will remain static. Given the need for constant upgrading, the potential for new system failures is always present, frequently unpredictable, and sometimes impossible to prevent.

## Privacy

While a recurrent problem for health care facilities has been their failure to protect massive data repositories from cyber predators, another risk-laden problem has involved hospitals and clinics simply sharing their data with other health care entities or uploading their data onto publicly accessible servers. Reports in the *Washington Post* and other media have described how Google partnerships for the purpose of training AI algorithms inadvertently resulted in some data with protected health information being uploaded in ways that exposed the data to anyone with basic search engine

capability.[11,12] Data used for research purposes must be appropriately de-identified or scrubbed of various items that can identify the subjects.[13] But, in certain instances, personnel have either failed to remove items that identified subjects—in one of the Google partnerships, by failing to notice x-ray images that showed patients' jewelry[11]—or exposed patients' identities by failing to delete common identifiers like treatment dates or doctors' notes[12] or social security numbers or addresses.

The kind of big data use that is typical of AI exponentially heightens the risk of data exposure. In 2020, Zack Whittaker reported that hundreds of hospitals, medical offices, and imaging centers were found to have insecure storage systems that allowed "anyone with an internet connection and free-to-download software to access over 1 billion medical images of patients across the world."[14] In 2019, a diagnostic medical imaging services company paid $300 million to the Office for Civil Rights to settle a data breach suit that exposed over 300 000 patients' protected health information.[15] Certain US hospitals and imaging centers perpetrated some of the most notorious breaches, which can make patients, in Dirk Schrader's words, "perfect victims for medical insurance fraud."[14]

### Consent to Data Repurposing

Even if data are properly de-identified and protected from privacy intrusions, securing patients' informed consent for the use or reuse of their data can be ethically challenging. Typically, patients consent to their data being used upon admission, such as for their treatments and hospital operations like billing and insurance, or for public health (as well as public security or law enforcement) programs, as permitted under the Privacy Rule of the Health Insurance Portability and Accountability Act (HIPAA).[16] But beyond those uses—especially for research purposes—additional and explicit consent is required.[13] Once patients consent to their deidentified data being used for purposes beyond those specified in the HIPAA regulations, however, HIPAA regulations no longer apply because HIPAA doesn't recognize deidentified patient information as protected.[17] As such, health care facilities can use that data however they want, including sharing it or selling it to data brokers or companies in the private sector.[13,18]

It is well recognized, however, that when deidentified data are coupled with other data streams, especially social media, it becomes easier to reidentify individuals and then classify them according to whatever an interested party's wishes are.[19] For example, multiple data sets have been compiled that identify individuals who might be considerably harmed from identity exposure—eg, lists of rape victims or persons afflicted with genetic or neuropsychiatric illnesses, substance use disorders, or erectile dysfunction.[20] The moral question then becomes whether health care facilities should engage in sharing or selling data in light of these privacy concerns because, once a facility does so, it cannot control how that data will be subsequently repurposed unless there are explicit and agreed-upon use limitations.

A variation of this problem that affects risk management more directly involves sharing or selling data with personally identifying information without patient consent. At least 2 university health care systems have been sued for failing to inform patients that their records might be shared with or sold to the private sector when the shared data involved personally identifying information.[12] In 2013, the US Department of Health and Human Services filed a protective objection in Delaware bankruptcy court, arguing that health care facilities facing bankruptcy cannot sell their patient data for debt relief without explicit patient consent.[21]

Consequently, an interesting and evolving legal problem these cases present is how exacting must the language of patient consent be to allow a facility to use *even deidentified health data*? The federal government recently imposed a requirement for researchers participating in the 1000 Genomes Project to obtain informed consent for the use of deidentified data. Researchers would have to pledge that data would only be used for the approved research; there would be no attempt to (re)identify individual participants; and data obtained from National Institutes of Health data repositories would not be sold, nor would the data be shared with anyone other than authorized persons.[22] Similarly, the California Consumer Privacy Act now requires businesses that collect consumer information to tell consumers how their data will be used and to inform them upon request with whom the data might be shared. Consumers also have the right to refuse to have their data sold.[23] Examples like these signal changing public attitudes toward the privacy of online data that will surely give health facilities pause. The question with which this essay will conclude is the extent to which risk management might find itself charged with managing developments like these.

## A New Era

This discussion has largely focused on 2 varieties of risk from AI technologies: those attaching to data, especially big data, and those attaching to certain technologies immediately bearing on or functioning as patient care interventions. If we now ask which one is likely to have the greater impact on risk management operations, the answer would seem to be the latter. Although data repurposing and security might pose some liability considerations and therefore be of interest to risk managers, the discipline's attention historically has been focused more on the intersection of humans and their environments. Thus, because AI technologies are anticipated to increasingly replace the human element of that intersection, it seems inevitable that risk managers in clinical environments will increasingly find themselves contemplating strategies to mitigate the risks these new technologies pose.

There is certainly a positive, risk management side to these developments, as various diagnostic and prognostic AI models are being touted as at least—if not more—accurate than their human counterparts.[24] Furthermore, AI technologies do not suffer cognitive lapses from fatigue nor do they encumber employers with the costs of employee benefits. On the negative side, however, history has taught that the introduction of novel, powerful, and complex technologies always comes with risks that oftentimes are not appreciated until they materialize.

Anticipating the extent of that threat might pose the greatest challenge for risk managers because of the way AI technologies can precipitate large-scale disasters. As long as AI models remain relatively decoupled from one another and each one performs a discrete or narrow task—eg, does a first read of mammograms but nothing else—the risk of large-scale events is reduced.[8] But as these models become "smarter" and begin "talking to one another"—a technological development that will likely be irresistible among AI developers—risk magnitude will exponentially increase.[25]

If the importation of AI technologies for diagnosis or treatment is very rapid, risk managers could find themselves enrolling in crash courses that familiarize them with AI models and their vulnerabilities. It should not be surprising if some larger health systems have some of their risk managers specialize in AI applications to manage their attendant risks. In any event, risk management will not be able to expect "business as usual" in the coming decades for the simple reason that AI systems will dramatically

change the delivery of health care operations. Those changes will usher in a new era of and for risk management.

## References


1. Spencer M. Artificial intelligence hype is real. *Forbes*. February 25, 2019. https://www.forbes.com/sites/cognitiveworld/2019/02/25/artificial-intelligence-hype-is-real/#4a0e618e25fa. Accessed May 6, 2020.
2. International Data Corporation. IDC expects Asia/Pacific artificial intelligence systems spending to reach nearly USD 5.5 billion in 2019. https://www.idc.com/getdoc.jsp?containerId=prAP45089819. Published May 21, 2019. Accessed May 6, 2020.
3. Schönberger D. Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. *Int J Inf Technol*. 2019;27(2):171-203.
4. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med*. 2018;378(11):981-983.
5. Banja J. Welcoming the "Intel-ethicist." *Hastings Cent Rep*. 2019;49(1):33-36.
6. Banja JD. *Patient Safety Ethics: How Vigilance, Mindfulness, Compliance, and Humility Can Make Healthcare Safer*. Baltimore, MD: Johns Hopkins University Press; 2019.
7. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94-98.
8. Webb A. *The Big Nine: How the Tech Titans and Their Thinking Machines Could Warp Humanity*. New York, NY: PublicAffairs; 2019.
9. Dudzinski DM, Hébert PC, Foglia MB, Gallagher TH. The disclosure dilemma—large-scale adverse events. *N Engl J Med*. 2010;363(10):978-986.
10. Alder S. Largest health data breaches of 2018. *HIPAA Journal*. December 27, 2018. https://www.hipaajournal.com/largest-healthcare-data-breaches-of-2018/. Accessed May 6, 2020.
11. MacMillan D, Bensinger G. Google almost made 100,000 chest x-rays public—until it realized personal data could be exposed. *Washington Post*. November 15, 2019. https://www.washingtonpost.com/technology/2019/11/15/google-almost-made-chest-x-rays-public-until-it-realized-personal-data-could-be-exposed/. Accessed March 24, 2020.
12. Wakabayashi D. Google and the University of Chicago are sued over data sharing. *New York Times*. June 26, 2019. https://www.nytimes.com/2019/06/26/technology/google-university-chicago-data-sharing-lawsuit.html. Accessed March 24, 2020.
13. Metcalf J, Crawford K. Where are human subjects in big data research? The emerging ethics divide. *Big Data Soc*. 2016;3(1):1-14.
14. Whittaker Z. A billion medical images are exposed online, as doctors ignore warnings. *TechCrunch*. January 10, 2020. https://techcrunch.com/2020/01/10/medical-images-exposed-pacs/?renderMode=ie11. Accessed March 24, 2020.
15. Tennessee diagnostic medical imaging services company pays $3,000,000 to settle breach exposing over 300,000 patients' protected health information [press release]. Washington, DC: US Department of Health and Human Services; May 6, 2019. https://www.hhs.gov/about/news/2019/05/06/tennessee-diagnostic-medical-imaging-services-company-pays-3000000-settle-breach.html. Accessed May 6, 2020.
16. Alder S. What is HIPAA authorization? *HIPAA Journal*. February 9, 2018. https://www.hipaajournal.com/what-is-hipaa-

authorization/#:~:text=HIPAA%20authorization%20is%20consent%20obtained,by%20the%20HIPAA%20Privacy%20Rule. Accessed August 25, 2020.
17. Snell E. De-identification of data: breaking down HIPAA rules. *HealthITSecurity*. April 3, 2015. https://healthitsecurity.com/news/de-identifcation-of-data-breaking-down-hipaa-rules. Accessed March 24, 2020.
18. Charette KR, DeAngelis TM. The use of de-identified patient information: understanding the scope of the HIPAA Privacy Rule. *Fitzpatrick Lentz & Bubba Blog*. June 20, 2017. https://www.flblaw.com/de-identified-patient-information-hipaa/. Accessed March 24, 2020.
19. Na L, Yang C, Lo CC, Zhao F, Fukuoka Y, Aswani A. Feasibility of reidentifying individuals in large national physical activity data sets from which protected health information has been removed with use of machine learning. *JAMA Netw Open*. 2018;1(8):e186040.
20. Martin KE. Ethical issues in big data industry. *MIS Q Exec*. 2015;14(2):67-85.
21. Brino A. Company needs patient OK to sell PHI. *Healthcare IT News*. December 20, 2013. https://www.healthcareitnews.com/news/patient-consent-required-sell-phi. Accessed March 24, 2020.
22. US Department of Health and Human Services. Attachment A: human subjects research implications of "big data" studies. https://www.hhs.gov/ohrp/sachrp-committee/recommendations/2015-april-24-attachment-a/index.html. Reviewed April 25, 2015. Accessed March 24, 2020.
23. Knowledge@Wharton. Your data is shared and sold ... what's being done about it? https://knowledge.wharton.upenn.edu/article/data-shared-sold-whats-done/. Published October 28, 2019. Accessed March 24, 2020.
24. Schier R. Hello AI, goodbye radiology as we know it. *Radiology Business*. February 18, 2020. https://www.radiologybusiness.com/topics/artificial-intelligence/hello-ai-goodbye-radiology-we-know-it. Accessed March 24, 2020.
25. Bostrom N. *Superintelligence: Paths, Dangers, Strategies*. Oxford, UK: Oxford University Press; 2017.

John Banja, PhD is a professor and medical ethicist at Emory University in Atlanta, Georgia. He is the editor of *AJOB Neuroscience*, and his most recent book is *Patient Safety Ethics: How Vigilance, Mindfulness, Compliance, and Humility Can Make Healthcare Safer* (Johns Hopkins University Press, 2019).

Citation
*AMA J Ethics*. 2020;22(11):E945-951.

DOI
10.1001/amajethics.2020.945.

Acknowledgements
Support for this article was provided by an unrestricted grant from the Advanced Radiology Services Foundation for research on understanding the ethics of artificial intelligence and radiology.

Conflict of Interest Disclosure
The author(s) had no conflicts of interest to disclose.

*The viewpoints expressed in this article are those of the author(s) and do not necessarily reflect the views and policies of the AMA.*


ISSN 2376-6980