




The need for health AI ethics in medical school education

Gali Katznelson¹ · Sara Gerke² 

Received: 29 June 2020 / Accepted: 16 February 2021 / Published online: 3 March 2021
© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

Abstract

Health Artificial Intelligence (AI) has the potential to improve health care, but at the same time, raises many ethical challenges. Within the field of health AI ethics, the solutions to the questions posed by ethical issues such as informed consent, bias, safety, transparency, patient privacy, and allocation are complex and difficult to navigate. The increasing amount of data, market forces, and changing landscape of health care suggest that medical students may be faced with a workplace in which understanding how to safely and effectively interact with health AIs will be essential. Here we argue that there is a need to teach health AI ethics in medical schools. Real events in health AI already pose ethical challenges to the medical community. We discuss key ethical issues requiring medical school education and suggest that case studies based on recent real-life examples are useful tools to teach the ethical issues raised by health AIs.

Keywords Artificial intelligence · Ethics · Health · Curriculum · Medical education

Introduction

Artificial Intelligence (AI) is transforming health care. In April 2018, the U.S. Food and Drug Administration (FDA) permitted marketing of the first “autonomous” AI software system, called IDx-DR, to provide a diagnostic decision for the eye disease diabetic retinopathy (Digital Diagnostics, 2020; FDA, 2018a). In September 2019, the medical neuro-technology company BrainScope received FDA clearance for its AI-based medical device BrainScope TBI (model: Ahead 500) that helps in the diagnosis of concussion and mild traumatic brain injury by providing results and measures (FDA, 2019a). Other health AIs that have already been cleared or approved by the FDA include Arterys, the first medical imaging platform powered by AI for cardiac patients (FDA, 2017), OsteoDetect, an AI software designed to detect wrist fractures (FDA, 2018b) and Viz.AI, a clinical decision

✉ Sara Gerke
sgerke@law.harvard.edu

¹ Schulich School of Medicine and Dentistry, Western University, 1151 Richmond St, London, ON N6A 5C1, Canada

² The Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics at Harvard Law School, Harvard University, 23 Everett Street, Cambridge, MA 02138, USA

support software that uses AI to analyze computed tomography results for indicators associated with a stroke (FDA, 2018c).

Over 160 health AI-based devices have been cleared or approved by the FDA so far (Ross, 2021), and there are currently many more products in the development pipeline. It is difficult to predict what the future of health AI will bring, but virtual assistants, personalized medicine, and autonomous robotic surgeries are not outside the realm of possibilities of how the health care system could look like in the future. Health AI has tremendous potential to transform medicine for the better, but it also raises ethical issues, ranging from informed consent, bias, safety, transparency, patient privacy to allocation.

Although health AI is already shaping health care and will continue to do so, health AI ethics is not yet a standard course in medical school curriculums. This gap may also be reflected in the current literature: Only few publications focus on AI and medical education, and even fewer mention ethics at all (Kolachalama & Garg, 2018; Wartman & Combs, 2018; Wartman & Combs, 2019; Paranjape et al., 2019; AMA, 2019).

We will first explain what health AI ethics is and will then argue why this education is needed in medical schools. Our focus will then be devoted to a discussion of six key ethical issues raised by health AI that require medical school education: informed consent, bias, safety, transparency, patient privacy, and allocation. These ethical issues form a framework for AI that is pertinent to clinical medical ethics. We suggest that it may be useful to teach Health AI Ethics in an applied context using real-life examples.

What is health AI ethics?

We define *Health AI Ethics* as the ‘application and analysis of ethics to contexts in health in which AI is involved’. Just as *health* has been defined in several ways (Huber et al., 2011), there is also no single definition for the term *artificial intelligence (AI)*, which was first coined in 1955 by McCarthy et al. (2006). In medicine, it is becoming increasingly important for health care professionals to understand AI and its applications.

For example, the American Medical Association (AMA) defines AI as:

The ability of a computer to complete tasks in a manner typically associated with a rational human being—a quality that enables an entity to function appropriately and with foresight in its environment. True AI is widely regarded as a program or algorithm that can beat the Turing Test, which states that an artificial intelligence must be able to exhibit intelligent behavior that is indistinguishable from that of a human (2018).

The understanding of what constitutes AI has evolved over the years. In the 1970s, AI in health care consisted of rule-based algorithms, such as electrocardiogram interpretation, which remain commonplace today but are not typically those referred to as “true AI” (Kundu et al., 2000; Yu et al., 2018). Recent developments in AI have consisted of Machine Learning (ML), a branch of computer science that finds patterns in data using algorithms (Murphy, 2012; Yu et al., 2018). ML can be supervised or unsupervised. Supervised ML is used to predict outputs based on correlations of input–output data during training, while unsupervised ML identifies patterns in data that is not labelled (Yu et al., 2018). For example, while supervised ML would be trained to detect the presence or absence of a certain type of cancer on an image, unsupervised ML would evaluate a large number of images to detect patterns and group images accordingly. In particular, developments within Deep

Learning, a field within ML that uses multi-layered artificial neural networks on large datasets, has propelled the field of AI forward (Yu et al., 2018). Most AI that is used in clinical care has yet to surpass human intelligence, and remains narrow and task-specific (Hosny et al., 2018). The AMA uses the term “Augmented Intelligence” to highlight current AI’s ability to augment the physician’s intelligence rather than to replace it (2018).

Ethics is the area of philosophy that, in simplest terms, considers questions of right and wrong. Ethics can be divided into normative ethics, metaethics, and applied ethics (Fieser, 1999). Normative ethics tries to answer questions about the right way to act. Normative ethical theories include, but are certainly not limited to deontology, consequentialism, and virtue ethics. Metaethics is the branch of philosophy that addresses questions about the nature of right and wrong. Applied ethics deals with applying ethical theories or principles to specific, real-life issues. Many medical students will be familiar with Beauchamp and Childress’ (2012) four principles of biomedical ethics: autonomy, non-maleficence, beneficence, and justice. This pluralistic theory, known as principlism, is meant to guide a pragmatic approach to resolving ethical dilemmas in medicine. A Health AI Ethics education should approach ethics through an applied, case-based method, but should also include elements from normative ethics and metaethics. We suggest that a framework to study Health AI Ethics consisting of the following six issues is most pertinent for medical learners: informed consent, bias, safety, transparency, patient privacy, and allocation.

Why is health AI ethics needed in medical school education?

Several signs point toward the increasing impact that health AI will have in medicine. First, there are gigantic amounts of data being collected across the global health care system. According to one estimate, the “global datasphere” is expected to grow to 175 zettabytes by 2025 (Reinsel et al., 2018). These unprecedented amounts of data may lead to revolutionary developments in fields that rely on data, such as health AI.

Second, it is estimated that the market for health AI will increase from 4.9 billion USD in 2020 to 45.2 billion USD in 2026 (Markets & Markets, 2020). In particular, tech giants that are traditionally not health-related, such as Google, Microsoft, and IBM, are key players in the health AI market and are probably motivated by the economic incentives of this rapidly growing market.

Third, patients have found themselves to be living in a connected world that uses AI. Health care is no longer confined to the doctor’s office—patients are targeted for health adds, are quick to consult ‘Dr. Google’ for advice, and use fitness trackers, health AI apps, and digital pills (Gerke et al., 2019) to help them monitor and manage their health. Many of these activities collect patient data, and data has been dubbed “the new oil” in today’s economy for its economic value (The Economist, 2017). Meanwhile, clinics have been encouraged to become “learning healthcare systems,” systems in which data and evidence are intended to drive the development of improved patient care (Olsen et al., 2007). Health AI has the potential to help hospitals achieve this goal through its capacity to analyze large quantities of data in novel ways. The massive generation of data, increasing market for data, and changing patient care landscape serve as welcoming indications for health AI’s potential to transform medicine.

Such changes, however, bring with them many ethical issues (see below). For example, both Anita Ho (2019) and Junaid Nabi (2018) have written in the Hastings Center Report about how AI should be designed and implemented ethically in medicine. Ho argues that

health AI should be developed within a culture of quality improvement in health care that responds to current needs and remains open to issues that arise along the way. Nabi argues that health AI should be designed with bioethical principles as a guidance. We argue that, for medical students, as important future stakeholders in medicine, it will be difficult to participate in the design, implementation, and use of health AI without adequate training in health AI ethics.

Some scholars have already called for a medical school curriculum reform in the age of AI (Kolachalama & Garg, 2018; Wartman & Combs, 2018; Wartman & Combs, 2019; Paranjape et al., 2019). For example, Kolachalama and Garg (2018) encourage medical schools to strive for “machine learning literacy,” which includes an emphasis on “the benefits, risks, and the ethical dilemmas” of AI. In June 2019, the AMA released a policy regarding the integration of augmented intelligence into medical education. The AMA recommends, among other things, the inclusion of data scientists and engineers on medical school faculties, as well as educational materials that address bias and disparities of augmented intelligence applications. The Association also calls on accreditation and licensing bodies to consider how augmented intelligence should be integrated into accreditation and licensing standards.

We support these calls and highlight the need for a medical school curriculum that comprehensively addresses the ethical dimensions of health AI. To ensure that health care professionals are adequately prepared for health AIs, medical school students do not only need to become knowledgeable in data science but also need to develop a nuanced understanding and awareness of the ethical issues raised by them. Today’s medical students will be involved in the development, implementation, and evaluation of AI that is integrated into health care. They require a solid foundation in AI ethics to engage meaningfully in these processes.

What are the ethical issues requiring medical school education?

There are, in particular, six key ethical issues raised by health AI that require medical school education: 1. informed consent, 2. bias, 3. safety, 4. transparency, 5. patient privacy, and 6. allocation. We believe that all health professionals should have an ethical understanding of how these issues interplay with AI. Although medical students may already come across these issues in their studies, AI complicates these issues and presents unique concerns that warrant special attention. For each issue, we provide examples with AI that can be incorporated into the curriculum in a case-based method.

Informed consent

The deployment of new health AIs in clinical practice raises urgent questions about informed consent (Gerke et al., 2020c). If a doctor uses a health AI to determine a treatment plan, does the patient have a right to know that the AI was involved in the decision-making process, and if so, what specific information should the patient be told? (Cohen, 2020). For example, does the clinician need to inform the patient about the data that is being used to train the algorithm, such as whether it is electronic health record data or artificially created (synthetic) data?

In an effort to respect patient autonomy, the medical community needs to decide if there is something essential about using health AIs that requires informed consent (in addition to signing the general treatment consent form). The answer to this issue will likely depend on the particular health AI and its application. Perhaps a hypothetical future AI that fully autonomously decides on the treatment course of a patient will warrant a specific informed consent. We believe that only by understanding the ethics surrounding informed consent, as well as how a specific health AI is developed, how it works, and how it is intended to be used, can physicians formulate a critical opinion on this matter. Stakeholders should start the discussion today on the ethical issue of informed consent before even more health AIs enter clinical practice. As a second step (after having this discussion), guidelines and communication plans for physicians on informed consent that are tailored to specific health AIs could then be developed and serve as a useful tool in clinical practice. Medical students need to watch this space and can also play a role in shaping this discussion.

Bias

There are different types of biases, and it is important for health care professionals to be aware of the types of biases that exist with health AI, and to think about how to mitigate such biases both within their patient encounters and on a health care system level. If health AIs are trained on biased data, existing disparities in health may be augmented rather than reduced. For example, the scenario could be that a Black patient's acral lentiginous melanoma, the most common type of melanoma in individuals with darker skin (Villines, 2019), is missed because a physician trusts a health AI cancer screening device that was primarily trained on white skin (Adamson & Smith, 2018). Whether this AI is based on supervised or unsupervised ML, if it is only trained on white skin as input data, it cannot learn how to diagnose cancer in any other skin type.

In other cases, the algorithm may exhibit bias. For example, Obermeyer et al., (2019) recently showed a major racial bias against Black patients in an algorithm used by large health systems and payers to guide health decisions applied to about 200 million people in the U.S. every year. The algorithm falsely assigned the same level of risk to Black as to White patients, even though Black patients in the dataset were much sicker (Obermeyer et al., 2019). Such a racial bias arose because the algorithm used health care costs (instead of illness) as a proxy for the level of health needs (Obermeyer et al., 2019). Since less money was spent on Black patients' health, the algorithm incorrectly concluded that Black patients were healthier (Obermeyer et al., 2019). Algorithmic biases may be mitigated by carefully thinking about the labels used (Obermeyer et al., 2019), but new algorithms are also needed that can be trained on unbiased models (Wiens et al., 2020).

Consequently, the goal should be “ethics by design—rather than after a product has been designed and tested” (Gerke et al., 2019). Stakeholders, particularly AI makers, must be aware of the types of biases that can exist alongside health AI and try as early as possible in the development process of their products to mitigate biases. With an awareness of the types of biases that can arise, medical students may be encouraged to join AI development teams, and contribute to the design of future health AIs. Moreover, they will be prepared for their future profession and will be ready to critically scrutinize a health AI before using it in clinical practice for the treatment of their patients.

Safety

Consider an online chatbot that assumes that a woman with a sensation of heartburn has gastroesophageal reflux disease, commonly known as acid reflux, and simply recommends an antacid. A trained physician should know the importance of immediately ruling out a heart attack for this patient. Many AI apps and chatbots are designed to limit unnecessary doctor visits, but some of them can also cause serious harm to consumers if they are not continuously updated, checked, or regulated.

Regulators like the FDA have recently initiated steps toward new approaches of regulating some health AIs—those that are classified as Software as a Medical Device (i.e., SaMD, “software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device”) (IMDRF, 2013)—to improve their performance and ensure their safety and effectiveness. Examples include the FDA’s effort to develop a Software Precertification Program (2019b), its recent discussion paper (2019c) to address the so-called “update problem” and the treatment of “adaptive” versus “locked” algorithms (Babic et al., 2019), and its recent Action Plan (2021). In particular, regulators like the FDA currently face the problem of whether they should permit the marketing of algorithms that continue to learn and change over time (FDA, 2019c; Babic et al., 2019). If they decide to permit the marketing of such adaptive algorithms, they face the follow-up issue of how they can safeguard that such AI/ML-based SaMD are continuously safe and effective. Regulators still need to figure out the details of such new innovative regulatory models, but it is essential that a key component of such models will be a continuous risk monitoring approach that focuses on risks due to features that are specific to AI/ML systems as well as a system view rather than a device one (Babic et al., 2019; Gerke et al., 2020a).

It is important for students to understand these efforts, with particular attention to the gaps and barriers to health AI safety that remain. Even with ongoing regulatory initiatives, future health professionals need to know that many health AIs are currently not subject to FDA review, such as certain clinical decision support software and many AI apps and chatbot (U.S. Federal Food, Drug, and Cosmetic Act, s. 520(o)(1)). Of those health AIs that do need to undergo FDA premarket review, they can be subject to different pathways requiring different controls to provide reasonable assurance of their safety and effectiveness (FDA, 2018d). It is important for medical students to be aware of the complicated regulatory landscape and to remain engaged with such processes and new developments in the field.

As patients embrace health care solutions that can be increasingly found outside of the doctor’s office, health professionals will need to educate patients about the potential risks of consumer health AIs. They will need to be able to recommend AI apps and chatbots that are designed with safety in mind. This will be a practical challenge for health professionals since there is already an overwhelming number of apps and chatbots available on the market, and more are added every day. In addition, these apps and chatbots are frequently updated, which makes it even harder to determine whether they are still trustworthy and recommendable. There are also significant concerns regarding who is liable when AI is involved in patient care (Price et al., 2019, 2021). By gaining an awareness of these problems, medical students might be encouraged to participate in solving them, such as by engaging with experts in the field to develop robust ethical and legal frameworks that evaluate the safety and effectiveness of health apps and chatbots.

Transparency

To complicate the matters of informed consent and safety, the use of health AIs that are ‘black boxes’ raises the question of how physicians can remain transparent with patients while working with systems that are, by nature, not fully transparent. ‘Black boxes’ can be described as software that is usually designed to help physicians with patient care, but that does not explain how the input data is analyzed to reach its decision (Daniel et al., 2019). This inexplicability may result from complicated AI/ML models that cannot be easily understood by physicians or due to the algorithm being considered proprietary (Daniel et al., 2019). If physicians cannot comprehend how or why a health AI/ML has arrived at a decision for a particular patient, it is important to consider whether they should be relying on the software, let alone what they should be informing their patients. Although data scientists are working on opening the black box of AI/ML (Lipton, 2016), the issue of how much information physicians and their patients should have will remain with explainable AI.

Consider a situation in which a newly deployed health AI/ML predicts the date of a patient’s death, but the physician has no way of understanding how the health AI/ML has calculated this date. ML techniques to predict mortality have already been described by several groups (Motwani et al., 2017; Shouval et al., 2017; Weng et al., 2019). Imagine that the health AI/ML has calculated that the patient will not survive a few more hours. Meanwhile, based on clinical experience, albeit limited, with other patients, the physician would estimate a much better prognosis for the patient. Should the physician inform the patient and their family about the health AI’s calculation? What exactly would the physician say? Now consider that the health AI/ML turns out to be correct in most such cases. Should a physician now rely on this AI/ML fully? Should physicians use black-box health AI/ML models in clinical practice at all?

The answers to such questions may depend on different considerations, such as the physician’s liability risk. They may also depend on whether the AI/ML maker has shown sufficient proof that the device is safe and effective such as through randomized clinical trials. As medical students might know, many of the drugs used in medicine may not have been fully understood initially. For example, clinicians prescribed Aspirin for about 70 years, knowing that the drug had antipyretic, analgesic, and anti-inflammatory effects, but without knowing its underlying mechanism (London, 2019). The pathway by which aspirin binds cyclooxygenase to inhibit prostaglandin production to produce these effects was only later understood (Vane & Botting, 2003). Thus, one might argue that physicians can (and perhaps should) use some black-box health AI/ML models in clinical practice, as long as there is sufficient proof that they are reliable and accurate and are not used for allocating scarce resources such as organs (Babic et al., 2020). If medical students are to be working alongside black-box health AI/ML models as physicians in the near future, the discussions about how to do so ethically should begin before they find themselves at the bedside disagreeing with a black-box health AI/ML that the hospital purchased for a considerable amount of money and thus is likely encouraging them to use.

Patient privacy

With the large amounts of health data collected from patients, beyond clinical settings and in daily life through wearables and health apps, patient privacy has emerged as an

important consideration. The accumulation of unprecedented amounts of health data may compromise patient privacy, without patients even realizing to what extent.

For example, the recent lawsuit, *Dinerstein v. Google*, has reflected the emerging concern for individuals' data protection and privacy (2019). The lawsuit was by Matt Dinerstein, a patient of the University of Chicago Medical Center, individually and on behalf of all other patients similarly situated, against Google, the University of Chicago, and the University of Chicago Medical Center. In 2017, the University of Chicago Medical Center and Google proclaimed a partnership to use new ML techniques to predict medical events, such as hospital readmissions (Wood, 2017), and the study results were published a year later (Rajkomar et al., 2018). Dinerstein claimed that between 2009 and 2016, the University of Chicago Medical Center transferred hundreds of thousands of medical records to Google, which included free-text notes and timestamps without obtaining patients' express consent (*Dinerstein v. Google*; Rajkomar et al., 2018). However, this lawsuit was dismissed in September 2020 by a federal judge in Illinois on the grounds that Dinerstein failed to demonstrate damages. This case highlights the challenges of pursuing claims against hospitals that share patient data with tech giants such as Google, and shows the insufficient protection of health data privacy (Becker, 2020).

Through this example, students can learn about data sharing issues such as how data should be handled and with whom it should be shared, as well as the types of rights patients should have regarding their data. This example helps to facilitate a discussion on emerging patient privacy concerns, such as the issue of reidentification through data triangulation (Cohen & Mello, 2019; Price & Cohen, 2019; Gerke et al., 2020b), the interplay between health AI innovation and sensitive patient data, and the relationship between hospital systems and third party health AI developers.

Patient privacy has been a longstanding principle in medicine, and should remain a key consideration with the development and implementation of all health AIs. Medical students should be confronted with new health AI developments alongside an understanding of the ethical and legal dimensions of data privacy in health care so they can critically appraise the implementation of new health technologies that threaten patient privacy.

Allocation

Finally, there is the issue of allocation. The just allocation of resources is an ongoing issue in health care, made more complicated by the advent of health AI. One real-life example involves the allocation of caregiving resources based on an algorithm. In 2016, Tammy Dobbs, a woman from Arkansas with cerebral palsy who was initially allocated 56 h of care per week as part of a state program, was suddenly allotted just 32 h after the state decided to rely on an algorithm to allocate its caregiving (Lecher, 2018). According to Ms. Dobbs, these hours were insufficient, and she was not given any information about how the algorithm reached its decision (Lecher, 2018). AI complicates allocation decisions because it can make these decisions invisible. Medical students must gain an awareness of how AI interfaces with resource allocation so they can advocate for future patients, who, like Tammy Dobbs, may find their access to health care resources undermined by automated systems.

The use of health AIs in clinical settings also depends on the reimbursement of such tools. Many future patients—perhaps the most vulnerable ones—may not have access to health AI diagnostic or health AI treatment tools if their insurance does not cover them. Questions that should be discussed with students include: What is justice in the context

of health AI? How can health AI be designed and used to promote justice rather than to subvert it? How can physicians, health insurance companies, and public health authorities ensure the just allocation of health AI?

Conclusion

As important future stakeholders in the health care system, today's medical students will have to make important decisions related to the use of health AI, both in every day clinical interactions with patients, and in broader policy discussions about the emergent integration of AI into health care. Recent events are highlighting the need to teach health AI ethics to medical students. We must begin the ethical discussion about health AI in medical schools so that students will not only be encouraged to pay attention to the ongoing developments in health AI but also will develop the tools needed to understand these developments through an ethical lens and be able to deal with new emerging ethical issues raised by health AI or other digital health technologies.

The ethical issues that should constitute this curriculum are informed consent, bias, safety, transparency, patient privacy, and allocation. There is no shortage of current health AI events that raise challenging questions about these issues, and that can be presented in the context of case studies. With exposure to case studies based on real-life examples that raise these issues, medical students can gain the skills to appreciate and solve the ethical challenges of health AI that are already affecting the practice of medicine.

Acknowledgements We thank I. Glenn Cohen for valuable comments on an earlier version of this manuscript.

Author contributions GK and SG contributed equally to the analysis and drafting of the paper. SG supervised the work.

Funding GK was supported by the 2019 Ontario Medical Student Association (OMSA) Medical Student Education Research Grant. S.G. was supported by a grant from the Collaborative Research Program for Biomedical Innovation Law, a scientifically independent collaborative research program supported by a Novo Nordisk Foundation Grant (NNF17SA0027784). S.G. received funding from the German Federal Ministry of Education and Research (BMBF), from April 1, 2016, to March 31, 2018, outside the submitted work.

References

- Adamson, A. S., & Smith, A. (2018). Machine learning and health care disparities in dermatology. *JAMA Dermatology*, *154*, 1247–1248.
- AMA (American Medical Association). (2018). Report of the Council on Long Range Planning and Development. Retrieved from <https://www.ama-assn.org/system/files/2018-11/a18-clrpd-reports.pdf>
- AMA (American Medical Association). (2019). AMA Adopt Policy, Integrate Augmented Intelligence in Physician Training. Retrieved from <https://www.ama-assn.org/press-center/press-releases/ama-adopt-policy-integrate-augmented-intelligence-physician-training>
- Babic, B., Gerke, S., Evgeniou, T., & Cohen, I. G. (2019). Algorithms on regulatory lockdown in medicine. *Science*, *366*, 1202–1204.
- Babic, B., Cohen, I. G., Evgeniou, T., Gerke, S., & Trichakis, N. (2020). Can AI fairly decide Who gets an organ transplant? *HBR*. Retrieved from <https://hbr.org/2020/12/can-ai-fairly-decide-who-gets-an-organ-transplant>.
- Beauchamp, T. L., & Childress, J. F. (2012). *Principles of biomedical ethics*. Oxford University Press.

- Becker, J. (2020). Insufficient Protections for Health Data Privacy: Lessons from *Dinerstein v. Google*. Retrieved from <https://blog.petrieflom.law.harvard.edu/2020/09/28/dinerstein-google-health-data-privacy>
- Cohen, I. G. (2020). Informed consent and medical artificial intelligence: What to tell the patient? *Georgetown Law Journal*, 108, 1425–1469.
- Cohen, I. G., & Mello, M. M. (2019). Big data, big tech, and protecting patient privacy. *JAMA*, 322, 1141–1142.
- Daniel, G., Sharma, I., Silcox, C., & Wright, M. B. (2019). Current State and Near-Term Priorities for AI-Enabled Diagnostic Support Software in Health Care. Retrieved from <https://healthpolicy.duke.edu/sites/default/files/atoms/files/dukemargolisaienabledxss.pdf>.
- Digital Diagnostics (2020). Retrieved from <https://dxs.ai>
- Dinerstein v. Google*. (2019). No. 1:19-cv-04311.
- FDA (U.S. Food and Drug Administration). (2017). K163253. Retrieved from https://www.accessdata.fda.gov/cdrh_docs/pdf16/K163253.pdf
- FDA (U.S. Food and Drug Administration). (2018a). De Novo Classification Request for IDX-DR. Retrieved from https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN180001.pdf
- FDA (U.S. Food and Drug Administration). (2018b). FDA Permits Marketing of Artificial Intelligence Algorithm for Aiding Providers in Detecting Wrist Fractures. Retrieved from: <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-algorithm-aiding-providers-detecting-wrist-fractures>
- FDA (U.S. Food and Drug Administration). (2018c). FDA Permits Marketing of Clinical Decision Support Software for Alerting Providers of a Potential Stroke in Patients. Retrieved from: <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-clinical-decision-support-software-alerting-providers-potential-stroke>
- FDA (U.S. Food and Drug Administration). (2018d). Step 3: Pathway to Approval. Retrieved from <https://www.fda.gov/patients/device-development-process/step-3-pathway-approval>
- FDA (U.S. Food and Drug Administration). (2019a). K190815. Retrieved from https://www.accessdata.fda.gov/cdrh_docs/pdf19/K190815.pdf
- FDA (U.S. Food and Drug Administration). (2019b). Developing a Software Precertification Program: A Working Model, v1.0 – 2019. Retrieved from <https://www.fda.gov/media/119722/download>
- FDA (U.S. Food and Drug Administration). (2019c). Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD). Retrieved from <https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf>
- FDA (U.S. Food and Drug Administration). (2021). Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan. <https://www.fda.gov/media/145022/download>.
- Fieser, J. (1999). *Metaethics, normative ethics, and applied ethics: Contemporary and historical readings*. Wadsworth Publishing.
- Gerke, S., Minssen, T., Yu, H., & Cohen, I. G. (2019). Ethical and legal issues of ingestible electronic sensors. *Nature Electronics*, 2, 329–234.
- Gerke, S., Babic, B., Evgeniou, T., & Cohen, I. G. (2020a). The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *NPJ Digital Medicine*, 3, 53.
- Gerke, S., Yeung, S., & Cohen, I. G. (2020b). Ethical and legal aspects of ambient intelligence in hospital. *JAMA*, 323, 601–602.
- Gerke, S., Minssen, T., & Cohen, I. G. (2020c). Ethical and legal challenges of artificial intelligence-driven health care. In A. Bohr & K. Memarzadeh (Eds.), *Artificial intelligence in healthcare* (pp. 295–336). Elsevier Inc.
- Ho, A. (2019). Deep ethical learning: Taking the interplay of human and artificial intelligence seriously. *Hastings Center Report*, 49, 36–39.
- Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, 18, 500–510.
- Huber, M., Knottnerus, J. A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., Leonard, B., Lorig, K., Loureiro, M. I., van der Meer, J. W., Schnabel, P., Smith, R., van Weel, C., & Smid, H. (2011). How should we define health? *BMJ*, 343, d4163.
- IMDRF (International Medical Device Regulators Forum). (2013). Software as a Medical Device (SaMD): Key Definitions. Retrieved from <http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf>
- Kolachalama, V. B., & Garg, P. S. (2018). Machine learning and medical education. *NPJ Digital Medicine*, 1, 54.

- Kundu, M., Nasipuri, M., & Basu, D. K. (2000). Knowledge-based ECG interpretation: A critical review. *Pattern Recognition*, 33, 351–373.
- Lecher, C. (2018). What Happens When an Algorithm Cuts Your Health Care. The Verge. Retrieved from <https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy>
- Lipton, Z.C. (2016). The Myths of Model Interpretability. ICML Workshop on Human Interpretability in Machine Learning. Retrieved from <https://arxiv.org/abs/1606.03490>
- London, A. J. (2019). Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Center Report*, 49, 15–21.
- Markets and Markets. (2020). Artificial Intelligence in Healthcare Market with Covid-19 Impact Analysis by Offering (Hardware, Software, Services), Technology (Machine Learning, NLP, Context-Aware Computing, Computer Vision), End-Use Application, End User and Region - Global Forecast to 2026. Retrieved from <https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-healthcare-market-54679303.html>
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27, 12–14.
- Motwani, M., Dey, D., Berman, D. S., Germano, G., Achenbach, S., Al-Mallah, M. H., Andreini, D., Budoff, M. J., Cademartiri, F., Callister, T. Q., Chang, H. J., Chinnaiyan, K., Chow, B. J. W., Cury, R. C., Delago, A., Gomez, M., Gransar, H., Hadamitzky, M., Hausleiter, J., ... Slomka, P. J. (2017). Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A 5-year multicentre prospective registry analysis. *European Heart Journal*, 38, 500–507.
- Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT Press.
- Nabi, J. (2018). How bioethics can shape artificial intelligence and machine learning. *Hastings Center Report*, 48, 10–13.
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366, 447–453.
- Olsen, L. A., Aisner, D., & McGinnis, J. M. (2007). *The learning healthcare system: Workshop summary*. National Academies Press.
- Paranjape, K., Schinkel, M., Panday, R. N., Car, J., & Nanayakkara, P. (2019). Introducing artificial intelligence training in medical education. *JMIR Medical Education*, 5, e16048.
- Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. *Nature Medicine*, 25, 37–43.
- Price, W. N., Gerke, S., & Cohen, I. G. (2019). Potential liability for physicians using artificial intelligence. *JAMA*, 322, 1765–1766.
- Price, W. N., Gerke S., & Cohen, I. G. (2021). How much can potential jurors tell us about liability for medical artificial intelligence? *Journal of Nuclear Medicine*, 62, 15–16.
- Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., Liu, P. J., Liu, X., Marcus, J., Sun, M., Sundberg, P., Yee, H., Zhang, K., Zhang, Y., Flores, G., Duggan, G. E., Irvine, J., Le, Q., Litsch, K., ... Dean, J. (2018). Scalable and accurate deep learning with electronic health records. *NPJ Digital Medicine*, 1, 18.
- Reinsel, D., Gantz, J., & Rydning, J. (2018). The Digitization of the World From Edge to Core. Retrieved from <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-data-age-white-paper.pdf>.
- Ross, C. (2021). As the FDA Clears a Flood of AI Tools, Missing Data Raise Troubling Questions on Safety and Fairness. Retrieved from <https://www.statnews.com/2021/02/03/fda-clearances-artificial-intelligence-data>.
- Shouval, R., Hadanny, A., Shlomo, N., Iakobishvili, Z., Unger, R., Zahger, D., Alcalai, R., Atar, S., Gottlieb, S., Matetzky, S., Goldenberg, I., & Beigel, R. (2017). Machine learning for prediction of 30-day mortality after ST elevation myocardial infarction: An Acute Coronary Syndrome Israeli Survey data mining study. *International Journal of Cardiology*, 246, 7–13.
- The Economist. (2017). The World's Most Valuable Resource is No Longer Oil, But Data. Retrieved from <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
- U.S. Federal Food, Drug, and Cosmetic Act, as amended through P.L. 116–260, enacted December 27, 2020, s. 520(o)(1).
- Vane, J. R., & Botting, R. M. (2003). The mechanism of action of aspirin. *Thrombosis Research*, 110, 255–258.
- Villines, Z. (2019). What is Acral Lentiginous Melanoma? MedicalNewsToday. Retrieved from <https://www.medicalnewstoday.com/articles/320223>
- Wartman, S. A., & Combs, C. D. (2018). Medical education must move from the information age to the age of artificial intelligence. *Academic Medicine*, 93, 1107–1109.

- Wartman, S. A., & Combs, C. D. (2019). Reimagining medical education in the age of AI. *AMA Journal of Ethics*, *21*, 146–152.
- Weng, S. F., Vaz, L., Qureshi, N., & Kai, J. (2019). Prediction of premature all-cause mortality: A prospective general population cohort study comparing machine-learning and standard epidemiological approaches. *PLoS ONE*, *14*, 3.
- Wiens, J., Price, W. N., & Sjoding, M. W. (2020). Diagnosing bias in data-driven algorithms for healthcare. *Nature Medicine*, *26*, 25–26.
- Wood, M. (2017). UChicago Medicine Collaborates with Google to Use Machine Learning for Better Health Care. Retrieved from <https://www.uchicagomedicine.org/forefront/research-and-discoveries/articles/2017/may/uchicago-medicine-collaborates-with-google-to-use-machine-learning-for-better-health-care>
- Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, *2*, 719–731.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.