Using AI to Make Healthcare Decisions

Kevin Boylan

20252634

Review article for CS438 2021

Research Methods

Department of Computer Science

National University of Ireland, Maynooth

Co. Kildare

Ireland

Lecturers: Mr. Joseph Duffin and Dr. Stephen Brown

Contents

Abstract

The integration and use of artificial intelligence (AI) in healthcare decision making is drastically changing the healthcare industry. It is improving outcomes for patients, reducing costs, helping with staff shortages, and assisting medical practitioners in their work.

The benefits are demonstrable and numerous. The performance of AI relative to doctors, the improvements in diagnosis (particularly in medical imaging), the reduction in medical errors due to error flagging, the applications in surgery, AI's ability to assist patients, and the likely trajectory of the software are very encouraging.

However, there are many difficulties and ethical questions raised that are not yet answered for. There is a need for accountability and transparency for patients, a likely need for reduced societal expectations of medical data privacy as AIs need a large amount of this data to process and learn from, and a need to mitigate AI biases to avoid worsening healthcare disparities for marginalized groups. The legal and regulatory obstacles are also quite extensive and medical AI literacy needs to be improved and more integrated into education. These issues will need to be addressed for this technology to reach its potential and achieve widespread use in daily clinical practice.

Introduction

Machine Learning (ML) algorithms are transforming AI. In healthcare, ML algorithms process medical data and automatically improve by learning and recognizing patterns in this data, and then extrapolate from this to output medical decisions. Deep Learning (DL) is branch of ML that is based on artificial neural networks (these emulate the neural network of the human brain) (Bari et al., 2021) – DL is the basis for many of the advancements seen recently in healthcare AI.

The aim of this review is to produce a balanced summary of the available research on the use of AI in medicine to make healthcare decisions, noting the existing and future benefits of the technology, its pitfalls, and the barriers to progress and implementation that must be overcome.

The research was conducted using the PubMed and Mendeley databases. The search terms used for PubMed were “AI Literacy”, “Causal Machine Learning Accuracy”, “AiCure”, and I also found an article on PubMed while searching “AI In Healthcare Historical” on Google. The search terms used on Mendeley were “AI Healthcare”, “AI Medical Imaging”, “Causal Machine Learning”, “Deep Learning Causality”, “AI Medical Robotics”, “AI Surgical Robots”, and “AI Autonomous Surgery”.

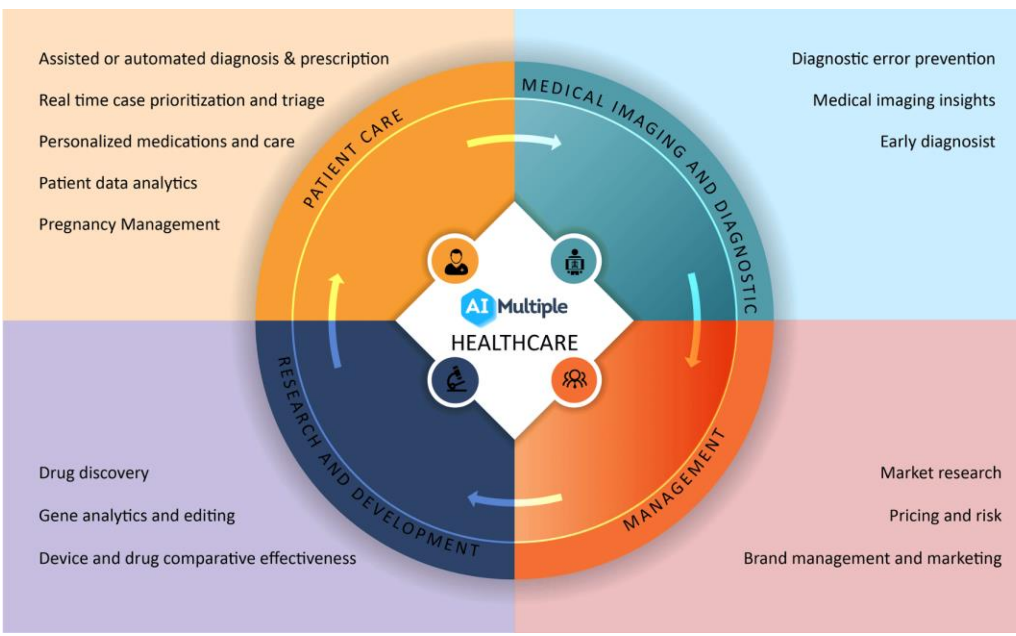

Fig1. Examples of AI Applications/Use Cases in Healthcare (Dilmegani, 2017)

Subsections

Historical Background

Medical Imaging

Medical Error Flagging

Patient Assistance

Robotics & Surgery

Ethical Issues – Data Ownership

Ethical Issues – Transparency & Accountability

Ethical Issues – Clinical Validation & Training Guidelines

AI Education

Ethical Issues – AI Biases

AI Performance Relative to Doctors

Historical Background

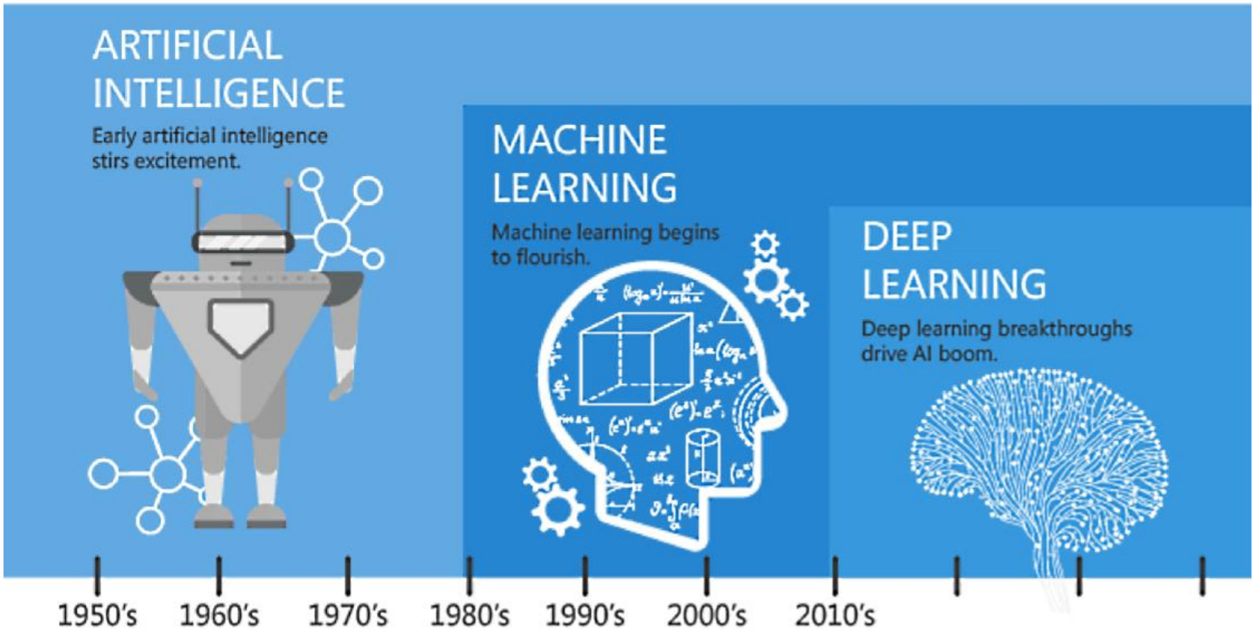

Fig2. Evolution of AI (Hee Lee & Yoon, 2021)

Following the conceptualization of AI in the 1950s, AI research experienced multiple periods of slow development due to a lack of funding between the 1970s-2000s that can be described as “AI Winters”. The 1970s were the point where AI research in healthcare began to proliferate. MYCIN was an early example of AI being applied to healthcare that was developed in this period. It was used to help diagnose patients with blood infections and treat them with appropriate antibiotics. In 1980, the American Association for Artificial Intelligence was founded with a subgroup dedicated to medical applications, and the development of ML algorithms began. The 2010s brought the most notable “AI Summer” where AI boomed and benefited the medical field, particularly with the adoption of Deep Learning (DL) algorithms (see Fig.2).

The use of real-time AI in medical apps started a shift in the 2010s towards more outcome-oriented treatment, and statistically, AI became more accurate on average at diagnosing disease than medical professionals. However, the technology had and still has hurdles preventing it from being fully utilized and advancing further in the medical field.

Two distinct branches of AI in healthcare developed. Physical AI involves devices used to physically help patients and assist medical staff. Virtual AI involves machine learning creating algorithms "through repetition and experience" (Wynants et al., 2020). Medical Imaging is a strong example of the uses of Virtual AI.

Medical Imaging

AI is being used to assist in diagnosis, reducing diagnostic error and identifying specific diseases with a high level of accuracy. For example, it is producing superior results in the Medical Imaging field (see Fig.1) through Deep Learning solutions. The diagnoses are more accurate and take less time. Examples of FDA-approved AI-based medical imaging systems include "icobrain" (used to interpret MRI and CT brain images) and "Transpara" (used to detect and assess likelihood of cancer) (Panayides et al., 2020). It also removes a large amount of subjectivity (Habuza et al., 2021) in the interpretation of results. These allow for more accurate predictions, reducing the percentage of false negatives and false positives which are detrimental to patients' physical and/or mental health. Using data from 15,000 patients, doctors at Moorfields Eye Hospital (London) created an AI that can operate on eye scans and recognize signs of eye disease with a "94% accuracy" (Hee Lee & Yoon, 2021) for over 50 eye diseases. This is operating at an expert level and the scan readings are being done in just seconds. This helps doctors to faster prioritize urgent cases that could result in a loss of sight and triage accordingly (see "Patient Care" section of Fig.1), and it reduces the time between the scan and treatment.

Medical Error Flagging

AI is also being used to flag medical errors and improve safety, e.g. MedAware is a platform that uses Machine Learning to reduce prescription errors. By retroactively looking back through data, a study discovered that 68.2% of its medical alerts would not have been produced by traditional Clinical Decision Support (CDS) support alert tools, and that MedAware's alerts had a high degree of clinical validity at 79.7% (Rozenblum et al., 2020).

Drug Adherence

AIs are also being used to monitor poor adherence and nonadherence more closely (patients not taking their medication or following the treatment plan that they've been prescribed) rather than relying on self-reporting and pill counts. The technology has clear benefits. A lack of drug adherence can cause an increase in hospitalization (e.g. via low adherence to treatments like antipsychotic and antiepileptic medication) and negatively affect the interpretability of clinical trials with 20-30% of clinical trials failing as a result.

AiCure utilizes deep learning neural networks and machine vision for patient and drug identification, and to confirm that a patient has swallowed the prescribed drug, with the added benefit of being able to document the time it has been ingested (Bain et al., 2017). It was compared in a 24-week study on subjects with schizophrenia to the effectiveness of the more traditional mDOT (modified directly observed therapy) where a portion of the patient's doses are taken under supervision. It positively modified behaviour, yielding a 25% increase in adherence over mDot which improves patient health. AiCure also allows medical professionals to intervene earlier if necessary and predict potential nonadherence.

Patient Assistance

AI-enabled nursing robots can retrieve drugs to reduce distractions for staff, and can be used to interact with patients, helping them with tasks such as walking and house chores. Their interactions with patients have become human-like and reliable enough that they are replacing human assistants (Habuza et al., 2021).

AI-enabled chatbots (Virtual Health Assistants) can use a more conversational style thanks to Machine Learning, Neural Networks, and Natural Languages Processing (Väänänen et al., 2021). Their current effectiveness from most studies appears to be positive or mixed (Milne-Ives et al., 2020) and more research is needed to quantify how effective they are and exactly how much they are reducing costs.

Robotics & Surgery

The process of training AI in robot-assisted surgery (RAS) to be used autonomously is underway. However, AI is not yet developed enough in environment perception and decision making to allow surgical robots used in clinical practice (e.g. the da Vinci surgical system) to perform surgeries autonomously. Instead, they are collaborative, requiring human control. They have a master-slave configuration where a surgeon directs the robotic tools, so the judgements are still being made by surgeons. However, they have already enabled increased surgical precision, more complex surgeries that were not previously possible, and are especially appropriate for surgeries involving repetition of the same movement.

An example of surgical AI training involved the use of imitation learning to get a da Vinci unit to mimic the placement of sutures (on an object, not a human) after watching a video of this action being performed. Imitation learning involves training an AI to recognize an individual surgical task as a series of subtasks to develop a model which can then be executed (Moglia et al., 2021).

AI's use in surgery is currently more developed for (1) preoperative planning which utilizes medical imaging and a patient's medical records to plan prior to the operation, and (2) computer-aided intraoperative guidance which allows critical patient information to be viewed in real-time during the operation. It is used for minimally invasive surgery (MIS) which means that minimal or no incisions are made during the operation, proving faster patient recovery times and less complications. The main areas of intraoperative guidance are 3D shape instantiation (a 3D reconstruction of a patient is compiled from 2D images), soft tissue tracking (learning strategies are used to differentiate organs in spite of issues like poor illumination), endoscopic navigation (techniques used to help guide endoscopes during endoscopic surgery, for example, using learning methods to estimate depth in endoscopic images), and Augmented Reality (projects "a semi-transparent overlay of preoperative image on the area of interest" on the patient) (Zhou et al., 2020).

Ethical Issues – Data Ownership

Data ownership is a critical issue and a potential bottleneck for the future of AI-assisted decision making in healthcare. Learning systems need to operate on huge amounts of data to be fully developed and to ensure that they are making reliable decisions with clinical utility. Otherwise, the possibilities will be much more limited if they are not permitted to infer from the patterns in these data sets (Panch et al., 2019). Therefore, large data infrastructures will need to be built that compile a huge amount of personal medical data. This could be an implementation barrier. There will need to be a discourse in broader society about data ownership and new laws and regulations will be required. It will be a contentious subject as some tolerance would need to be allowed for partnerships with the private sector, the population would need credible assurances that their private data will not be abused for commercial means, and they will need to trust and approve of these fundamental changes.

There are risks “of hacking if the platforms used for collecting, storing, and sharing of data have access to the Internet, like cloud platforms, or if the physical hard drives are not protected fully” (Moglia et al., 2021). Even if the societal benefits are clearly conveyed, some percentage of civilians may be uncomfortable with their data being shared if proportional government oversight to enforce a high standard for the security, privacy and integrity of their medical data being shared is not seen as being possible. A blockchain-style approach to data security is a possible solution but more research and development will be required.

Another privacy concern is the machine learning algorithms themselves - even when efforts are taken to de-identify data, ML algorithms can use a small number of data points to re-identify a patient. It seems that “traditional expectations for healthcare privacy might no longer be attainable” (Crigger & Khoury, 2019).

Ethical Issues – Transparency & Accountability

Transparency and accountability are also major issues (Lysaght et al., 2019). Machine Learning algorithms are being integrated into Clinical Decision Support Systems (CDSS), connecting into patients’ electronic health records (EHRs) and operating on their data to analyze, find patterns and make associations. Compared to traditional, rule-based systems like MYCIN which have a predictable input-output relationship, doctors are currently hesitant to rely on these systems and are arguably underutilizing them. It is key that doctors can maintain control of the diagnosis process. Therefore, these AI-dependent systems will need to be more understandable for doctors so they can decide whether to accept the conclusions made and provide reasonable explanations to patients. For example, it should be clear to doctors which data points are most heavily weighted. Unfortunately, as these ML systems are not hard-coded and as thoroughly verified as prior medical software, sometimes it is not possible to adequately explain the results generated by ML algorithms to patients (Lysaght et al., 2019), even with the assistance of software engineers.

The issues of transparency, attribution of liability (and data ownership) extend strongly to autonomous AI-driven robotic surgery. It is a particularly sensitive area due to the inherent risk of injury and death posed. “Hundreds if not thousands of fully annotated videos for each specific type of surgery” will be required to fully develop these AI (Gumbs et al., 2021), necessitating large data collection of these videos with patient permission. Mock operations on cadavers (human corpses) will need to be used to train these AI.

Ethical Issues – Clinical Validation & Training Guidelines

Robot-assisted surgery (RAS) is a good area to describe the difficulties of clinical validation, a process which will be necessary for AIs to be deployed in daily clinical practice. Existing studies on AI in RAS have major limitations with small datasets and little external validation on the models. This makes it difficult to determine how many samples are needed to train these models for proficiency in a specific task and how they perform on other datasets (Moglia et al., 2021). There are also no fully developed guidelines on how to train an AI for RAS. This will need to change.

Policy H-480.940 of The American Medical Association (AMA, 2018) provides some general guidelines for many of the issues highlighted in this review, indicating a strong intention to address these ethical issues. The most detailed focus of the policy is on clinically validating healthcare AIs, arguably due to the immense difficulty of this task.

AI Education

While AI systems need to be made more understandable and user-centered for doctors, the study of AI will also need to be rapidly integrated into medical school curricula as medical AI literacy is currently demonstrably low in education. A 2021 survey of 52 clinical faculty and 121 medical students revealed a poor awareness of medicine-oriented AI topics for both faculty (50% awareness) and students (30% awareness) (Wood et al., 2021) with most of their knowledge about AI being derived from pop culture. Faculty reported that they lacked a basic understanding of AI at a higher rate than students. This will need to be addressed but there does exist a desire among both groups to address their limited AI literacy with students being recommended to study this field for their PhDs or Masters.

Ethical Issues – AI biases

It is critical that diagnostic biases in healthcare are considered and minimized when building AI systems, or we could essentially see the automation of these biases. Bias is a systematic perpetrator of inequalities in healthcare, affecting diagnostic tools, medical assessment frameworks and medical education curricula being taught. There is a large risk that AI drawing from existing biased sources without adjustment will "exacerbate historic and existing medical mistreatment" (Straw, 2020) and health disparities, particularly if drawing from historical medical resources where, for example, homosexuality was categorized as a disorder. Rooting out biases when extracting from historical data is not a simple task.

While not medicine-based, a simple example of this issue in practice was Amazon's Machine Learning recruitment tool. It had to be discontinued in 2018 as it was evidently discriminating against women applying for technical roles, awarding resumes low ratings if they contained "women's" or "woman". This was a consequence of the historical data that the algorithm was being fed as training data. The tool flagged keywords of the candidates who were most successful in the past over a ten-year period (Nkonde, 2019).

The issue of AI inheriting bias extends beyond historical resources. Research has shown that current clinical bias and a lack of women symptomology in medical curricula causes higher rates of gender-based misdiagnosis with male and female patients with the same symptomatology being evaluated differently. As a result, women experience longer waiting times for a referral for cancer treatment, obtain more basic interventions and diagnostics on average, and suffer higher morbidity for conditions like heart disease.

Implicit biases must also be accounted for. Gender-based perception is evident in the treatment of personality disorders where despite exhibiting the same symptoms, women are more likely to be diagnosed with borderline personality disorder and seen as having hysterical personalities whereas men are more likely to be diagnosed with PTSD and seen as having more organic symptoms and antisocial personalities.

There is a strong relationship between sample bias and diagnostic bias. Many research trials have a disproportional amount of young white males, and women and racial/ethnic minorities are underrepresented and comparatively excluded in biomedical research. AIs must then operate on inadequate testing samples and training datasets and generalize from them which can lead the algorithms to essentially "define their own reality and use it to justify their results" (Straw, 2020).

Deploying ML algorithms into healthcare at this time without a comprehensive process to identify and control for bias is questionable. There is an insufficient understanding of how currently disseminated models are created and what specific data they have learned from. An empirical evaluation of 81 studies comparing AI with clinicians showed major bias, and a lack of justification and transparency with these AI models.

MINIMAR (MINimum Information for Medical AI Reporting) is a potential solution that proposes a minimum reporting standard on the data used to develop medical AI with 4 main standards: 1) how the final cohort in training data was obtained from the target population, 2) sensitive training set demographics and characteristics like race, 3) detailed explanation of model design and development, allowing the intent of the model to be interpreted and compared (replication should be permitted), 4) information on model evaluation (including optimization and validation) to help explain how local model optimization can be done (Hernandez-Boussard et al., 2020).

AI Performance Relative to Doctors

To put the usefulness of AI in this field into perspective, associative Machine Learning algorithms have an accuracy putting them within the top 48% of doctors. Therefore, they are roughly as accurate as the average doctor. The issue with these models is that causation does not always equal correlation, i.e. a disease may be positively correlated with a patient’s condition(s) but, for example, a doctor would rank the disease as being unlikely to be cause of the condition(s) while an associative algorithm cannot do this. They cannot make counterfactual inferences about the causal relations between elements in the data.

Causal machine learning algorithms introduce causal reasoning, i.e. they account for "the probability that the occurrence of" an effect "was in fact brought about the target cause" (Richens et al., 2020) and provide a ranking. As a result, they have an expert level of accuracy putting within the top 25% of doctors.

Discussion

The common theme among the papers studied in this review is that AI's integration into healthcare poses significant challenges but that these are worth confronting and addressing as the benefits of integrating AI are undeniable and wide-ranging. The way that an AI's results are arrived at is a key issue – the preparation of the training data is extremely important, and guidelines and frameworks need to be developed to ensure that these ML AIs can be clinically validated. Medical accountability, transparency, informed consent, privacy rights, AI biases, and AI education are also sticking points which will need to be fully addressed. Agencies like the FDA and the EMA (European Medicines Agency) have a responsibility to protect people and their approval for the continued integration of this technology will be critical.

Future Developments

Just like in the 2010s as seen Fig.2, AI-driven healthcare technologies are going to continue to boom, improve and evolve over the next decade. However, the pace of their integration into daily clinical work will progress much slower than the technology itself as there are legal and technical challenges which will need to be overcome. Stricter guidelines and frameworks for training AI will be developed with higher standards. Precision medicine (not medicine created for an individual but rather personalized treatment plans) will be primarily driven by machine learning and most imaging analysis will be done with the assistance of AI. AI will be used mostly as an aid and most medical practitioners will not be replaced by AI, but their role may shift slightly to the more human skills and aspects of healthcare which AIs are not as capable of like "empathy, persuasion and big-picture integration" (Davenport & Kalakota, 2019). Causal reasoning will become more utilized in Deep Learning solutions in and outside of medicine to overcome issues with predictive models (Vasudevan et al., 2021). For healthcare, AI will be used to provide increasingly more accurate and reliable diagnoses.

Word Count: 3499

References

AMA. (2018). *H-480.940 Augmented Intelligence in Health Care | AMA*. https://policysearch.ama-assn.org/policyfinder/detail/augmented%20intelligence?uri=%2FAMADoc%2FHOD.xml-H-480.940.xml

Bain, E. E., Shafner, L., Walling, D. P., Othman, A. A., Chuang-Stein, C., Hinkle, J., & Hanina, A. (2017). Use of a Novel Artificial Intelligence Platform on Mobile Devices to Assess Dosing Compliance in a Phase 2 Clinical Trial in Subjects With Schizophrenia. *JMIR MHealth and UHealth*, *5*(2). https://doi.org/10.2196/MHEALTH.7030

Bari, H., Wadhwani, S., & Dasari, B. V. M. (2021). Role of artificial intelligence in hepatobiliary and pancreatic surgery. *World Journal of Gastrointestinal Surgery*, *13*(1), 7–18. https://doi.org/10.4240/wjgs.v13.i1.7

Crigger, E., & Khoury, C. (2019). Making policy on augmented intelligence in health care. *AMA Journal of Ethics*, *21*(2). https://doi.org/10.1001/amajethics.2019.188

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, *6*(2). https://doi.org/10.7861/futurehosp.6-2-94

Dilmegani, C. (2017). *Examples of AI Applications/Use Cases in Healthcare*. https://research.aimultiple.com/healthcare-ai/

Gumbs, A. A., Frigerio, I., Spolverato, G., Croner, R., Illanes, A., Chouillard, E., & Elyan, E. (2021). Artificial intelligence surgery: How do we get to autonomous actions in surgery? *Sensors*, *21*(16). https://doi.org/10.3390/s21165526

Habuza, T., Navaz, A. N., Hashim, F., Alnajjar, F., Zaki, N., Serhani, M. A., & Statsenko, Y. (2021). AI applications in robotics, diagnostic image analysis and precision medicine: Current limitations, future trends, guidelines on CAD systems for medicine. *Informatics in Medicine Unlocked*, *24*, 100596. https://doi.org/10.1016/J.IMU.2021.100596

Hee Lee, D., & Yoon, S. N. (2021). Application of artificial intelligence-based technologies in the healthcare industry: Opportunities and challenges. *International Journal of Environmental Research and Public Health*, *18*(1). https://doi.org/10.3390/ijerph18010271

Hernandez-Boussard, T., Bozkurt, S., Ioannidis, J. P. A., & Shah, N. H. (2020). MINIMAR (MINimum information for medical AI reporting): Developing reporting standards for artificial intelligence in health care. In *Journal of the American Medical Informatics Association* (Vol. 27, Issue 12). https://doi.org/10.1093/jamia/ocaa088

Lysaght, T., Lim, H. Y., Xafis, V., & Ngiam, K. Y. (2019). AI-Assisted Decision-making in Healthcare. *Asian Bioethics Review*, *11*(3). https://doi.org/10.1007/s41649-019-00096-0

Milne-Ives, M., de Cock, C., Lim, E., Shehadeh, M. H., de Pennington, N., Mole, G., Normando, E., & Meinert, E. (2020). The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review. In *Journal of Medical Internet Research* (Vol. 22, Issue 10). https://doi.org/10.2196/20346

Moglia, A., Georgiou, K., Georgiou, E., Satava, R. M., & Cuschieri, A. (2021). A systematic review on artificial intelligence in robot-assisted surgery. *International Journal of Surgery*, *95*, 106151. https://doi.org/10.1016/J.IJSU.2021.106151

Nkonde, M. (2019). Is AI Bias a Corporate Social Responsibility Issue? *Harvard Business Review*, *1*, 1–5. https://hbr.org/2019/11/is-ai-bias-a-corporate-social-responsibility-issue?ab=hero-subleft-1

Panayides, A. S., Amini, A., Filipovic, N. D., Sharma, A., Tsaftaris, S. A., Young, A., Foran, D., Do, N., Golemati, S., Kurc, T., Huang, K., Nikita, K. S., Veasey, B. P., Zervakis, M., Saltz, J. H., & Pattichis, C. S. (2020). AI in Medical Imaging Informatics: Current Challenges and Future Directions. In *IEEE Journal of Biomedical and Health Informatics* (Vol. 24, Issue 7). https://doi.org/10.1109/JBHI.2020.2991043

Panch, T., Mattie, H., & Celi, L. A. (2019). The “inconvenient truth” about AI in healthcare. *Npj Digital Medicine*, *2*(1). https://doi.org/10.1038/s41746-019-0155-4

Richens, J. G., Lee, C. M., & Johri, S. (2020). Improving the accuracy of medical diagnosis with causal machine learning. *Nature Communications*, *11*(1). https://doi.org/10.1038/S41467-020-17419-7

Rozenblum, R., Rodriguez-Monguio, R., Volk, L. A., Forsythe, K. J., Myers, S., McGurrin, M., Williams, D. H., Bates, D. W., Schiff, G., & Seoane-Vazquez, E. (2020). Using a Machine Learning System to Identify and Prevent Medication Prescribing Errors: A Clinical and Cost Analysis Evaluation. *Joint Commission Journal on Quality and Patient Safety*, *46*(1). https://doi.org/10.1016/j.jcjq.2019.09.008

Straw, I. (2020). The automation of bias in medical Artificial Intelligence (AI): Decoding the past to create a better future. *Artificial Intelligence in Medicine*, *110*. https://doi.org/10.1016/j.artmed.2020.101965

Väänänen, A., Haataja, K., Vehviläinen-Julkunen, K., & Toivanen, P. (2021). AI in healthcare: A narrative review. *F1000Research*, *10*. https://doi.org/10.12688/f1000research.26997.2

Vasudevan, R. K., Ziatdinov, M., Vlcek, L., & Kalinin, S. v. (2021). Off-the-shelf deep learning is not enough, and requires parsimony, Bayesianity, and causality. In *npj Computational Materials* (Vol. 7, Issue 1). https://doi.org/10.1038/s41524-020-00487-0

Wood, E. A., Ange, B. L., & Miller, D. D. (2021). Are We Ready to Integrate Artificial Intelligence Literacy into Medical School Curriculum: Students and Faculty Survey. *Journal of Medical Education and Curricular Development*, *8*, 238212052110240. https://doi.org/10.1177/23821205211024078

Wynants, L., Smits, L. J. M., & van Calster, B. (2020). Demystifying AI in healthcare. *BMJ (Clinical Research Ed.)*, *370*. https://doi.org/10.1136/BMJ.M3505

Zhou, X. Y., Guo, Y., Shen, M., & Yang, G. Z. (2020). Application of artificial intelligence in surgery. In *Frontiers of Medicine* (Vol. 14, Issue 4). https://doi.org/10.1007/s11684-020-0770-0