Analyzing the Impacts of ML Image Recognition Systems on Physicians

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Valeria Pabon

Spring 2023

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Rider W. Foley, Department of Engineering and Society

Introduction

Artificial Intelligence (AI) systems have revolutionized many industries, and their recent implementation in the healthcare field has the potential to transform many aspects of medicine. A branch of computer science, AI is concerned with creating machines capable of executing tasks requiring human-level intelligence (Basu et al., 2020). AI systems can take in large amounts of data, employ mathematical algorithms that gather information from that data, and then the information that is collected generates an output that can help solve a problem, such as diagnosing cancer within radiological images (Basu et al., 2020). These systems have the potential to significantly improve patient diagnostics, drug discovery, and the personalization of treatments with greater accuracy and efficiency (Pietyra, 2022). One of the most promising areas of innovation is the implementation of AI systems in medical imaging for disease diagnostics.

Machine learning (ML) systems, a specific type of AI, are used for imaging processing and interpretation for medical imaging. These ML systems are especially useful in radiology because they can take in MRI and CT scans and act as a diagnostic tool through image recognition (Pesapane et al., 2018). Additionally, many studies have shown that these systems also have the potential to be equal to or even more efficient in disease diagnostics than experienced physicians. For example, one study showed that 7 of the 32 ML systems were able to diagnose lymph metastasis in women with breast cancer more accurately than the 11 experienced physicians (Ehteshami Bejnordi et al., 2017).

Although AI systems provide many benefits for medicine, some limitations need to be considered when understanding the extent to which AI systems can potentially transform the healthcare field. A specific limitation is that the creation of well-performing AI models is based on the availability of high-quality data. This type of data is needed because the AI systems are trained on a portion of the collected data, also called a training set data (Basu et al., 2020). If this data is biased, such as if it over-represents a particular race, gender, or age group, then the trained model will be biased. This overrepresentation of different groups would then lead to biased outputs that could affect whether a machine can detect specific diseases within the image (Basu et al., 2020). In order to prevent bias, the data collected and trained on must represent the demographics population for which it is being used. Additionally, AI systems are known to be "black-boxes" because of the complexity of the algorithms used to produce outputs. This complexity makes it hard for its users to identify patterns of inaccurate predictions or biases that are present until it is continuously observed (Codeacademy Team, 2023). With these limitations in mind, it is critical to understand how AI systems have altered the rights and responsibilities of physicians who rely on the systems to understand their impact on the medical field.

ML Systems for Clinical Use

Machine learning systems are an application of artificial intelligence where machines take in data and learn information that would be difficult for humans. ML incorporates computational models and algorithms that change over time as the system gets exposed to more data, allowing minimal human interference in the process (Burnham, 2020). The most complex form of machine learning involves deep learning, which consists of neural networks that attempt to mimic the human brain by the way it takes in data, allowing the system to learn from data without being explicitly programmed (Pesapane et al., 2018). The complexity of deep learning arises from the fact that this form of machine learning consists of layered algorithms that can take in and process unstructured data, such as text and images. These layered algorithms allow deep learning systems to adjust and fit themselves for accuracy, allowing the system to make predictions about a new image with increased precision (IBM, 2022). The ability of deep learning AI systems to learn about and make predictions of an image has allowed these technologies to be used as image recognition systems in many fields, specifically in the healthcare field for disease diagnostics (Awati, 2022).

This specific form of artificial intelligence is increasingly being applied in the healthcare field because of its ability to detect features in imaging data that exceed what humans, specifically physicians, can view (Davenport & Kalakota, 2019). These ML image recognition systems can recognize and classify patterns from imaging modalities such as X-rays, MRIs, and CTs. This classification is helpful for disease diagnostics because the images produced from the imaging modalities allow for the visualization of internal body structures, allowing the image recognition model to screen for illnesses (Zhang & Sejdić, 2019). ML image recognition systems can screen for illnesses because the algorithms are trained on an image dataset to extract specific features that allow them to then extract the same features from the image inputs (Huang et al., 2022). This ability to screen for illnesses is very beneficial because it allows for the early detection of diseases that are sometimes not visible to the physician in a scan, allowing the patient to be correctly diagnosed and treated earlier (Kumar et al., 2022). One of the primary uses of ML image recognition systems is for working in clinical decision support (CDS). CDS encompasses systems that can help to diagnose or treat human disease by automating clinical tasks that are normally done by physicians (Lyell et al., 2021). When working in CDS, ML systems can fall under assistive or autonomous support for physicians. Assistive systems can take measurements and act as a second opinion to help physicians make decisions. This support type is different from autonomous systems, which are capable of providing clinical and treatment options that will directly impact the patient (Clark, 2018). Image recognition systems that have been approved by the FDA and are starting to be implemented in the healthcare field encompass both of these support types.

Technological Citizenship and AI

The framework of technological citizenship will be used to evaluate the sociotechnical relationship between image recognition systems and their impact on physicians. This framework provides a way to analyze how technology can influence the rights and obligations of citizens within a democracy (Andrews, 2006). The rights of citizens are defined as the rights of access to knowledge, informed consent, and reasonable levels of risk exposure. Additionally, the duties of citizens include achieving technological literacy and protecting the civic good. In order to understand the impact of the implementation of image recognition systems, the rights and duties of physicians influenced by this technology must be analyzed. This focus is necessary because there currently needs to be more access to knowledge regarding the demographics of datasets used for algorithm training for ML image recognition systems. Companies that train the algorithms do not always report detailed information on the datasets that are used, which limits the ability of physicians to evaluate how well the technology will perform on their patients (PEW, 2021). This lack of transparency of the datasets may lead physicians to trust the results implicitly without questioning whether the results could be biased. Without knowing the demographics of the datasets, it is possible that the ML image recognition systems were trained on a dataset that overrepresents a specific gender, age, or racial group. This overrepresentation would lead to the image recognition system performing better on the patient scans within the overrepresented group and poorly on patients that fall within underrepresented groups. If the physician cannot access the training dataset information, it impacts the physician's ability to give informed consent. This impact of informed consent is because informed consent requires a complete understanding of the possible risks and consequences (*Informed Consent*, 2022). Suppose the companies producing the image recognition system do not provide information regarding the dataset. In that case, the physician is unable to know the risks of using the machine on scans of their patients that fall into different demographic groups.

Additionally, due to its "black-box" nature, the algorithm's complexity has impacted physicians' abilities to achieve technological literacy. This complexity makes it difficult to determine how exactly the machine produced the output that it did using the data, making it even harder for physicians to detect unwanted behaviors, such as biased outputs (Rahman & Scali, 2022). Deep learning systems used for image recognition of diseases rely on interconnected nodes in a layered structure that work together to produce an output for a given input. This artificial neural network is difficult to understand because each layer requires a specific threshold to be surpassed before proceeding to another layer for further calculations (Savage, 2022). The specific thresholds and layers are unknown since the deep learning network is able to derive features from image inputs itself and then determine the relevant features that it will analyze and use for differentiation (AWS, 2023). Since the deep learning network learns independently, it is difficult for physicians to understand what decisions go into classifying whether a specific disease is present in the image the AI system is scanning. Without this transparency in the algorithm being used, or the mathematical calculations being performed in the AI system, the ability to achieve technological literacy is hindered. Technological literacy requires the user to understand what the technology is and how it works. Physicians' ability to achieve technological literacy is impacted without understanding how the image recognition systems can diagnose diseases (ITEEA, 2023). This lack of understanding can greatly impact not only the physicians who use these AI systems, but also the patients who are having their scans analyzed. By utilizing this framework, the impact of image recognition systems can be analyzed through the assessment of the system's influence on the rights and duties of physicians.

Research Question and Methods

The research question that is the focus of this research is: how have ML image recognition systems affected the rights and duties of physicians in the U.S. when diagnosing patients? This understanding is essential for physicians who rely on this technology for disease diagnostics because the technology's results will ultimately affect their patient's treatment. Therefore, analyzing how this technology impacts physicians is necessary to ensure that the patients receive the most effective treatment possible.

The methods utilized to investigate the research question included document analysis of current image recognition systems that have received FDA approval. A repository of all 521 FDA-approved AI/ML-enabled medical devices currently marketed in the U.S. was used to identify the image recognition systems being implemented in the healthcare field (FDA, 2022). Each device in the repository was analyzed to see whether it met the criteria of being an ML-enabled device for disease diagnosis via image recognition. Using the device's submission number to find the device's FDA Indications for Use report and 510(k) Premarket Notification, information within these documents was used to determine whether the technology met the criteria. If all three criteria points: ML device, image recognition, and disease diagnostic, could not be determined using those two documents, then the device's company webpage was analyzed for further information. After all 521 devices were analyzed, the devices that met the criteria were placed in a separate list, and four were selected for further analysis.

For each device selected, industry reports, diagnostic training manuals, or clinical studies were searched. Keyword searches were conducted across PubMed, Web of Science (WoS), and Google searches to find relevant articles and other studies. These databases and search engines were used to provide the necessary documents to perform analyses that provided insight into how ML systems impact physicians' ability to give informed consent and achieve technological literacy.

Information regarding the ML system's training dataset's demographics for each device was compiled to illustrate whether the data was accessible by users and if it was representative of the patient population. The patient population was determined by finding the relevant demographic statistics for the disease the device was diagnosing. A table was used to provide the relevant demographics, including age, sex, and race, of the dataset to determine if any bias was present through the overrepresentation of certain groups by comparing them to the patient population. This information also provided insight into what type of dataset demographics devices train on that are approved by the FDA and whether this information was easily provided to users, which is needed by physicians who use the devices to ensure informed consent. Document analysis also provided insight into how the black-box algorithm for each device was presented to its users. This was done by determining whether or not enough information about the algorithm's process was presented for the user to achieve technological literacy. Looking at the training manuals, clinical studies, and FDA-approval reports also allowed for the identification of relevant trends seen in medical devices that ultimately receive FDA approval.

Results

The use of ML image recognition systems for disease diagnostics has impacted the rights and duties of physicians because of the lack of transparency by companies creating the technology. Without companies providing information regarding the algorithms used and the dataset their image recognition technology was trained on, this has prevented physicians from achieving technological literacy and hindered their ability to give informed consent. Although these systems have grown in the healthcare field, specific information about the technology, including clinical studies used to prove their effectiveness and manuals describing the algorithm processes, are still not readily available. For the clinical studies that were found, key pieces of the dataset's demographic information, such as race, were not reported. Additionally, manuals that included information about aspects of the algorithm processes were presented with language that would be difficult to understand if the reader was unfamiliar with the subject. Although all of the technologies analyzed underwent FDA approval, making either clinical studies or manuals readily available is optional. The findings of this research indicate a universal lack of transparency for clinically-used image recognition systems that impact the physicians who use them.

Clinical Studies

In order to receive FDA approval, many of the technologies underwent clinical studies to prove their effectiveness in diagnosing diseases. However, most of the studies used to prove the technology's effectiveness are not readily available nor required to be publicly available by the FDA. After performing many Google searches and going through multiple company websites, a clinical study with dataset demographics for each technology was found. All studies provided demographic information regarding the age and sex of the patient population, but none of them indicated any information regarding the racial makeup of the dataset (see Table 1).

Table 1: Demographics of Algorithm Dataset (Faiella et al., 2022; Jones et al., 2020; Kim et al., 2020b; Sim et al., 2020)

Technology	Diagnosis	Age (yrs)	Sex	Race
OsteoDetect	Fractures	Range: 22-90 Mean: 55	43% Male 57% Female	N/A
Quantib Prostate	Prostate Cancer	Range: 52.4-84 Mean: 66.1	100% Male	N/A
Auto Lung Nodule Detection	Lung Cancer	Range: 19-78 Mean: 52.4	57% Male 43% Female	N/A
Lunit INSIGHT MMG	Breast Cancer	Range: N/A Mean: 53	100% Female	N/A

The lack of racial demographics is very important for technologies detecting cancers that affect racial and ethnic groups disproportionately, such as breast, prostate, and lung cancer (see Figure 1). For example, prostate cancer has a higher incidence rate in black non-Hispanic men than in white non-Hispanic men. In order to ensure there is no bias in the Quantib Prostate algorithm, it would have to be trained on a dataset that closely matched the demographics seen in the population of individuals with prostate cancer. However, because that information was not reported, it is unknown whether the dataset accurately represents those racial and ethnic groups or potentially overrepresents some other groups. This could lead to the algorithm being able to detect prostate cancer better in patients who are overrepresented in the study compared to ones that are not.

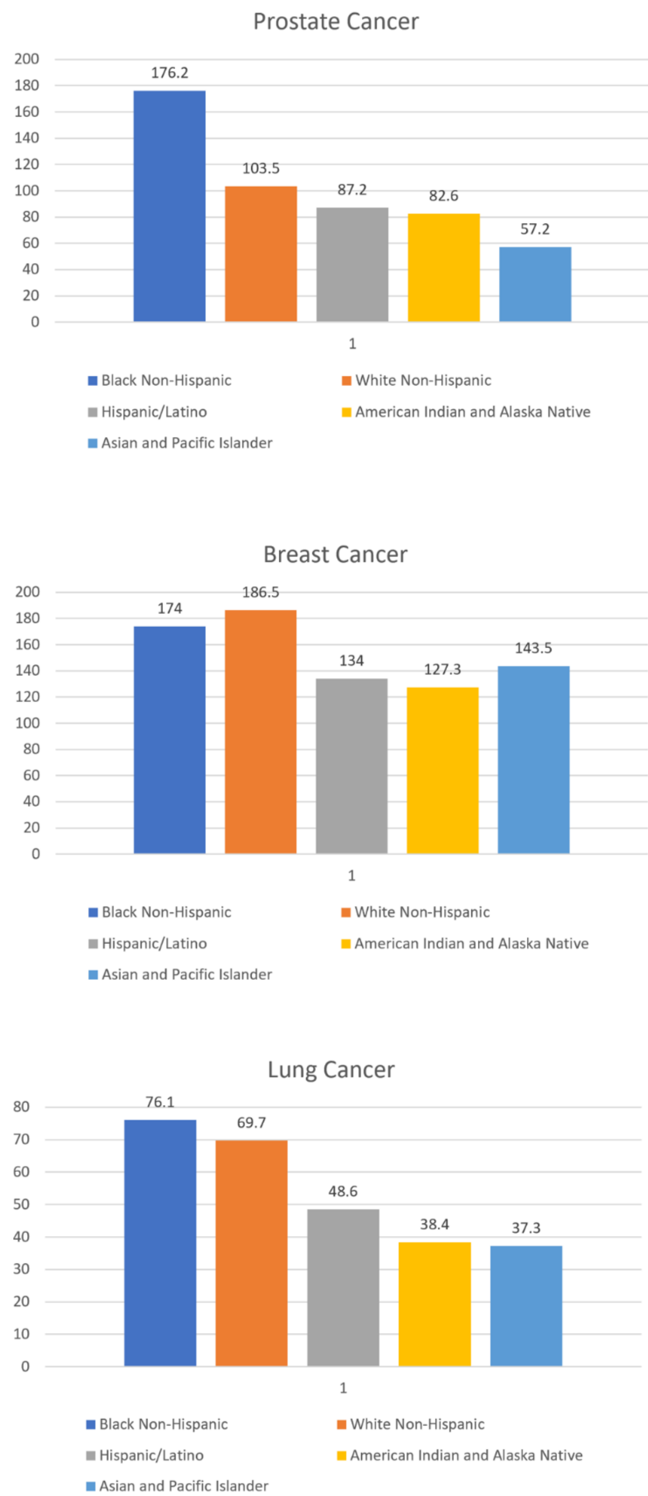

Figure 1: Demographics of incident rates (per 100,000 individuals) of prostate, breast, and lung cancer in the United States (Elflein, 2023; Ellington, 2022; Schabath et al., 2016)

These findings suggest that the lack of transparency in providing racial demographics of the datasets can significantly impact the physician's ability to give informed consent when using the technology on their patients. Since physicians cannot access knowledge that will allow them to know whether the technology is biased, they cannot give informed consent because of the concern about the efficacy of the technology on patients of specific racial groups.

Algorithm Manuals

Two of the four technologies, Lunit INSIGHT MMG and OsteoDetect, were the only technologies analyzed that gave public access to manuals that provided information regarding their respective algorithm processes. However, access to these manuals involved many visits to news articles and company websites, eventually leading to appendices in documents that provided this information. Quotes of the information presented in the AI algorithm manuals were placed in a readability calculator to determine whether technological literacy was achievable for its users (WordCalc, 2020). The Gunning Fog Index (FOG) value was calculated for each quote, and the index value that resulted estimated the years of formal education required to understand the text. Technological literacy would be achievable if the information presented had a reading level of 7th grade, which is a FOG value of 8 (Readable, 2011). For the information provided in the Lunit INSIGHT MMG manual, the following quote was placed in a readability calculator:

"The algorithm was trained in a semi-supervised manner in terms of existence of pixel-level labels, since only a portion of the exams were annotated. Training procedure consists of two sequential stages: patch-level training from scratch (stage-1), followed by image-level fine-tuning (stage-2)" (Kim et al., 2020a)

This quote received a FOG value of 18.8, indicating that achieving technological literacy would be difficult because this text requires a graduate level of understanding. A similar value indicating graduate-level education was needed to understand the text was also seen for information provided in the OsteoDetect manual. For example, the following quote in the OsteoDetect manual had a FOG value of 19:

"Upon receiving a DICOM object, the Input DICOM File Processing & Filtering component applies a number of filtering rules based on image characteristics and DICOM tags (age, modality, anatomy, contrast) to ensure that only eligible images are analyzed by the algorithm." (FDA, 2018)

In both these examples, the information about the algorithm processes for both technologies were written in a way where the reader could have difficulty understanding what the text was describing. Although physicians undergo a formal education that could be more than 18 years, it would be incorrect to assume that they would fully understand what the manual was describing since their later years of formal education are more medically based. This suggests that companies are not focused on ensuring that their users can achieve technological literacy when using their devices because they do not present their information in a way that can be easily understandable to the public.

Trends within FDA-Approved Technologies

Based on the document analyses performed, all image recognition technologies must provide the FDA with information, such as a 510(k) Premarket Notification, to prove that their device effectively detects diseases in patient scans. Although the FDA publicly posts its approval for the technologies, it does not give the public any information they use to base their decision.

For example, in the FDA news release of its approval for OsteoDetect, the FDA mentions that the study of 1,000 radiograph images submitted to assess its performance “demonstrated that the readers’ performance in detecting wrist fractures was improved using the software…when aided by OsteoDetect” (FDA, 2020). However, that news release contained no information about the study, and further searches for that study on databases and Google produced no results. This trend, as seen with other image recognition technologies, suggests that the FDA does not mandate information about the technologies to be provided by the companies. This ultimately shows that there is an accepted lack of transparency in the FDA approval process which can negatively impact the individuals who use these FDA-approved technologies. Physicians who use these FDA-approved technologies have the potential of using a device that does not have easily accessible training manuals or clinical studies, which hinders the physician from fulfilling their duty to ensure the best patient care since they wouldn’t know exactly how these FDA-approved devices work.

Discussion

The lack of transparency from the image recognition systems has negatively impacted the physicians who use it from achieving technological literacy and giving informed consent. Furthermore, these findings can provide insight into how different social groups and society may give meaning to these technologies. The results from the analysis of the clinical studies provided for each technology showed that there could be biases present in the algorithm that could affect how it works on different groups of people. Using the framework of Technological Citizenship, these biases could affect the results which could lead to the technology not satisfying the needs of either the patient or the physician.

Limitations to this research involve a potential human error in the selection process for the AI systems that fit the criteria for image recognition systems. It is possible that some of the technologies on the list of FDA-approved AI/ML-enabled medical devices fit the criteria and were missed. The selection process involved analyzing many documents and web pages which may have led to missed keywords that would have shown that the technology met all criteria but resulted in them being excluded from the list. Another limitation occurred during the data compilation for the image recognition systems because most of the documents that resulted from the keyword searches did not provide information about their algorithm processes. Variations of the keywords were searched, but this lack of data did not allow for the analysis of the description of the algorithm process for two of the technologies.

Future studies on these image recognition systems' impact in the healthcare field could be further elevated by analyzing more than four image recognition systems. This could provide more data to support conclusions about the common lack of transparency present. Additionally, testimonies from physicians who use these systems could be helpful when analyzing how those technologies have personally impacted them. This information is not readily available on the internet, so providing those testimonies will provide further evidence to the claim that these technologies are affecting their rights and duties.

Analyzing the social and human dimensions of technology is necessary for all engineers since the technologies that we create and work with have the potential to impact our society and the people in it. This research will advance my engineering practice by understanding the importance of being aware of the social ramifications of the technologies I am creating. In biomedical engineering, this is really important because this understanding will help me elicit empathy and compassion when coming up with technological solutions that address different medical problems. This research will also increase my awareness of actions in the technological design process that can later lead to bias being embedded in the technology and the importance of transparency.

Conclusion

Analyzing the image recognition system's impact on its users when used in a clinical setting can inform others starting to implement these systems in their workplace. As the use of AI technologies in the healthcare field continues to grow, it is necessary to understand this technology's limitations and the need for greater transparency within these systems. Without this transparency, it can lead to bias being embedded in these systems and perpetuated without realization. Physicians who use these technologies must be able to understand how these technologies work in order to ensure that the results are benefiting and not hurting their patients. In order to work towards a solution that helps to solve this problem of transparency, it is critical to consider the effects these systems have on various social groups and society. Understanding how the rights and duties of physicians are being impacted by these technological systems can help others to understand how these technologies have the power to change the healthcare field and the people in it. This knowledge will encourage others to push for more transparency in these systems across all fields, helping to decrease the potential of biases in these systems. Ultimately, this research can increase awareness of the effects of implementing these specific artificial intelligence systems and increase support for better education and transparency for its users.

References

Andrews, C. (2006). Practicing Technological Citizenship. *IEEE Technology and Society Magazine*, *25*(1), 4–5.

Awati, R. (2022, September). *What are convolutional neural networks?* Enterprise AI. https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network

AWS. (2023). *What is a Neural Network?* Amazon Web Services, Inc. https://aws.amazon.com/what-is/neural-network/

Basu, K., Sinha, R., Ong, A., & Basu, T. (2020). Artificial Intelligence: How is It Changing Medical Sciences and Its Future? *Indian Journal of Dermatology*, *65*(5), 365–370. https://doi.org/10.4103/ijd.IJD_421_20

Burnham, K. (2020, May 6). Artificial Intelligence vs. Machine Learning: What's the Difference? *Northeastern University Graduate Programs*. https://www.northeastern.edu/graduate/blog/artificial-intelligence-vs-machine-learning-whats-the-difference/

Clark, B. (2018, June 18). *Autonomous vs. Assistive AI: How to unlock real value in healthcare*. Digital Diagnostics. https://www.digitaldiagnostics.com/newsroom/autonomous-vs-assistive-ai-how-to-unlock-real-value-in-healthcare/

Codeacademy Team. (2023). *Dangers of the Black Box*. Codecademy. https://www.codecademy.com/article/dangers-of-the-black-box

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, *6*(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94

Ehteshami Bejnordi, B., Veta, M., Johannes van Diest, P., van Ginneken, B., Karssemeijer, N., Litjens, G., van der Laak, J. A. W. M., & and the CAMELYON16 Consortium. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA*, *318*(22), 2199–2210. https://doi.org/10.1001/jama.2017.14585

Elflein, J. (2023, February 2). *Prostate cancer incidence rate by ethnicity U.S. 2019*. Statista. https://www.statista.com/statistics/672946/prostate-cancer-incidence-rate-us-by-ethnicity/

Ellington, T. D. (2022). Trends in Breast Cancer Incidence, by Race, Ethnicity, and Age Among Women Aged ≥20 Years—United States, 1999–2018. *MMWR. Morbidity and Mortality Weekly Report*, *71*. https://doi.org/10.15585/mmwr.mm7102a2

Faiella, E., Vertulli, D., Esperto, F., Cordelli, E., Soda, P., Muraca, R. M., Moramarco, L. P., Grasso, R. F., Beomonte Zobel, B., & Santucci, D. (2022). Quantib Prostate Compared to an Expert Radiologist for the Diagnosis of Prostate Cancer on mpMRI: A Single-Center Preliminary Study. *Tomography*, *8*(4), Article 4. https://doi.org/10.3390/tomography8040168

FDA. (2018). *Evaluation oF Automatic Class III Designation For OsteoDetect* (No. DEN180005; p. 25). United States Food and Drug Administration. https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN180005.pdf

FDA. (2020, March 24). *FDA permits marketing of artificial intelligence algorithm for aiding providers in detecting wrist fractures*. US Food & Drug; FDA. https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-algorithm-aiding-providers-detecting-wrist-fractures

FDA. (2022, October 5). *Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices*. FDA; FDA. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

Huang, J., Li, J., Li, Z., Zhu, Z., Shen, C., Qi, G., & Yu, G. (2022). Detection of Diseases Using Machine Learning Image Recognition Technology in Artificial Intelligence. *Computational Intelligence and Neuroscience*, *2022*(1), 5658641. https://doi.org/10.1155/2022/5658641

IBM. (2022). *What is Deep Learning?* IBM. https://www.ibm.com/topics/deep-learning

Informed Consent. (2022). Oxford Learner's Dictionaries. https://www.oxfordlearnersdictionaries.com/us/definition/english/informed-consent

ITEEA. (2023). *Technological Literacy Standards*. Internation Technology and Engineering Educators Association. https://www.iteea.org/48897.aspx

Jones, R. M., Sharma, A., Hotchkiss, R., Sperling, J. W., Hamburger, J., Ledig, C., O'Toole, R., Gardner, M., Venkatesh, S., Roberts, M. M., Sauvestre, R., Shatkhin, M., Gupta, A., Chopra, S., Kumaravel, M., Daluiski, A., Plogger, W., Nascone, J., Potter, H. G., & Lindsey, R. V. (2020). Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs. *Npj Digital Medicine*, *3*(1), Article 1. https://doi.org/10.1038/s41746-020-00352-w

Kim, H.-E., Kim, H. H., Han, B.-K., Kim, K. H., Han, K., Nam, H., Lee, E. H., & Kim, E.-K. (2020a). *Supplementary Appendix for Changes in cancer detection and false-positive recall in mammography using artificial intelligence*. The Lancet Digital Health. https://doi.org/10.1016/S2589-7500(20)30003-0

Kim, H.-E., Kim, H. H., Han, B.-K., Kim, K. H., Han, K., Nam, H., Lee, E. H., & Kim, E.-K. (2020b). Changes in cancer detection and false-positive recall in mammography using artificial intelligence: A retrospective, multireader study. *The Lancet Digital Health*, *2*(3). https://doi.org/10.1016/S2589-7500(20)30003-0

Kumar, Y., Koul, A., Singla, R., & Ijaz, M. F. (2022). Artificial intelligence in disease diagnosis: A systematic literature review, synthesizing framework and future research agenda. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-021-03612-z

Lyell, D., Coiera, E., Chen, J., Shah, P., & Magrabi, F. (2021). How machine learning is embedded to support clinician decision making: An analysis of FDA-approved medical devices. *BMJ Health & Care Informatics*, *28*(1). https://doi.org/10.1136/bmjhci-2020-100301

Pesapane, F., Codari, M., & Sardanelli, F. (2018). Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *European Radiology Experimental*, *2*, 35. https://doi.org/10.1186/s41747-018-0061-6

PEW. (2021, August 5). *How FDA Regulates Artificial Intelligence in Medical Products*. PEW. https://pew.org/3yglbCS

Pietyra, W. (2022, September 15). Artificial Intelligence in data analytics - How AI improve efficiency, quality and accuracy? | CUX. *CUX*. https://cux.io/blog/how-does-ai-improve-the-efficiency-quality-and-accuracy-of-data-analytics/

Rahman, A., & Scali, G. (2022, August 22). *Explainability of AI Models to Prevent Bias in Clinical Systems*. Journal of AHIMA. https://journal.ahima.org/page/explainability-of-ai-models-to-prevent-bias-in-clinical-systems

Readable. (2011). *The Gunning Fog Index*. Readable. https://readable.com/readability/gunning-fog-index/

Savage, N. (2022). Breaking into the black box of artificial intelligence. *Nature*. https://doi.org/10.1038/d41586-022-00858-1

Schabath, M. B., Cress, W. D., & Muñoz-Antonia, T. (2016). Racial and Ethnic Differences in the Epidemiology of Lung Cancer and the Lung Cancer Genome. *Cancer Control : Journal of the Moffitt Cancer Center*, *23*(4), 338–346.

Sim, Y., Chung, M. J., Kotter, E., Yune, S., Kim, M., Do, S., Han, K., Kim, H., Yang, S., Lee, D.-J., & Choi, B. W. (2020). Deep Convolutional Neural Network–based Software Improves Radiologist Detection of Malignant Lung Nodules on Chest Radiographs. *Radiology*, *294*(1). https://doi.org/10.1148/radiol.2019182465

WordCalc. (2020). *Readability Calculator*. Wordcalc. https://www.wordcalc.com/readability/

Zhang, Z., & Sejdić, E. (2019). Radiological images and machine learning: Trends, perspectives, and prospects. *Computers in Biology and Medicine*, *108*, 354–370. https://doi.org/10.1016/j.compbiomed.2019.02.017