The Response to Algorithmic Bias in Disease Diagnosis

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

Wenxuan (Sharon) Zheng

Spring 2021

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

S. Travis Elliott, Department of Engineering and Society

Abstract

As the United States' demography is changing at an unprecedented rate, inequity has been one of the biggest challenges in the current healthcare system. The implementation of Artificial Intelligence (AI) can be a solution to avoid the biased human decision-making process during disease diagnosis. However, despite the AI algorithms are not inherently biased, algorithmic bias can be created throughout its development and implementation. In order to ensure equity in the healthcare system using diagnostic AI algorithms, the STS thesis will discuss the algorithmic bias in disease diagnosis and provide insights into future policymaking by identifying the source of algorithmic bias and discussing ways to encourage diversity. Through the lens of the co-production of science and social order, the challenges and opportunities in using diagnostic AI algorithms will be explored to minimize inequity.

Introduction

In the United States, the widening economic inequality and increasing population diversity have been accompanied by more health disparities. Many efforts have been done to promote equity in healthcare, which ensures high quality of care to be independent of personal characteristics, such as ethnicity, gender, age, disability, socioeconomic status, and geographic location. However, the United States is still ranked as the last in measures of equity among the developed countries (Berchick, Hood, & Barnet, n.d.; Braithwaite et al., 2017; "Health Insurance Coverage in the United States: 2017," n.d.). Despite the scientific advancement, there are significant differences in quality of care between low and high-income adults and between different ethnic groups (Berchick et al., n.d.; Braithwaite et al., 2017; "Health Insurance Coverage in the United States: 2017," n.d.).

One solution to combat inherent human bias is the implementation of Artificial Intelligence (AI) in the healthcare industry to avoid biased human decision-making while maintaining a high level of efficiency and accuracy. In disease diagnosis, all minorities were more likely to have undiagnosed diabetes compared to Whites by physicians (Kim et al., 2018). At the same time, AI has been intensely studied to solve complicated medical problems, and scientists believe that AI algorithms can improve the accuracy of disease diagnosis without any human bias. Many studies have shown the increasing trend of AI implementation in the healthcare field as AI technology is especially advantageous in disease diagnosis using radiology, where a robust AI algorithm can be built based on a large number of medical images (Davenport & Kalakota, 2019; Rodriguez-Ruiz et al., 2019).

However, AI has a long documented record of low diversity, and previous applications of AI in other applications demonstrate the prevalence of algorithmic bias, which systematically creates unfair outcomes that cause unintentional harm, such as discrimination based on skin color. For example, many facial recognition devices at an airport security checkpoint generally take longer to process people with darker skin and are less accurate at identifying them (Cook, Howard, Sirotin, Tipton, & Vemury, 2019). Previous implementation in the healthcare industry to facilitate hospital efficiency at a management level also demonstrates algorithmic bias, including the exclusion of disadvantaged groups. In the case of predicting no-shows, AI predictions use models that consider many features of a patient. As a result, although unintentional, patients with lower income, pre-existing conditions, and addiction problems will be regarded as a low priority due to the high correlation with no-shows (McCullough, n.d.). Realizing the social dimensions of healthcare and their contribution to algorithmic bias is the first step to reduce disparities in AI healthcare.

The goal of equity in healthcare using AI is to benefit everyone fully and equitably from the phenomenal capacity of scientific research instead of only benefiting the majorities at the cost of minorities. Ultimately, a sense of trust cannot be established unless people understand that they are obtaining high-quality care from AI and being treated equally without any discrimination. Therefore, collective intelligence will help to make a reliable and robust diagnostic system. This thesis will explore the current causes of inequity due to the deterministic nature of AI and the uncertainties of regulations in general. Then, the current Food and Drug Administration (FDA) approved diagnostic AI algorithm will be studied to provide more insights on the potential source of algorithmic bias and workplace diversity. Finally, through the lens of co-production of science and social order, this paper will use the results to identify the challenges and opportunities to minimize AI bias. Both the scientific and social dimensions will lead to implications and recommendations for future efforts.

The Deterministic Nature of AI

Most AI algorithms are composed of convoluted structures and encrypted details. They also make decisions solely based on trained datasets and give results with no explainability. The deterministic nature of AI makes it particularly difficult to be implemented in a safe, empowering, and satisfying environment, as healthcare requires equal and high-quality care for physical, mental, and social wellbeing.

Transparency is one of the biggest challenges for AI in the healthcare industry. First of all, the AI algorithm is a "black-box" to the physicians who will be using it, and even the basic understanding of AI algorithms can be hard to learn, yet healthcare professionals will have to understand the limitations of the AI algorithm they use to avoid mistakes (Miller, 2019). The low transparency not only impedes the implementation of AI but also has profound negative effects on equity if the user does not understand the limitations in applicable populations. Furthermore, if the trained algorithm gives a false result, the physicians are unable to identify the root cause due to the high complexity of AI algorithms (Esteva et al., 2019).

The incomprehensiveness comes from the fact that AI algorithms are not knowledge-based but data-based. The nature of data-based algorithms can be demonstrated by a Russian neural network-powered app Artisto, which correctly predicted Donald Trump's victory in the election in 2016 against Hilary Clinton based on facial features of the past presidents ("What was reached by artificial intelligence in 2016 | Earth Chronicles News," n.d.). However, giving the correct predictive results for the presidential election is not equivalent to understanding the social and political context of the United States. In fact, the most plausible explanation for the accurate prediction is that there were no female presidents in the United States history. The AI algorithm can only simply make predictions based on correlations to the pre-existing datasets, so that Hillary Clinton is disadvantageous because she is an outlier to the dataset, giving her disadvantages. Similarly, in the context of disease diagnosis using radiology, the trained AI algorithm doesn't understand the human anatomy like the healthcare professionals do but just simply makes predictions based on correlations to the datasets. Therefore, an algorithm trained on one of the populations is unlikely to achieve the same accuracy to a different population, which can exacerbate pre-existing trends in the training dataset that resulted from inequity.

The disparities of care can be furthered by the fact that healthcare professionals using AI algorithms are unable to explain the results and give a range of options for the patients, which are basic patient expectations. Patients also expect that the physicians will be familiar with their records, which will be less likely with AI algorithms. If patients' expectations are unmet, they are less likely to return for ongoing and follow-up care (Lateef, 2011). Especially with the implementation of AI in healthcare, patient-centered care may shift away to achieve high efficiency at unnecessary costs, further increasing inequity.

Together, the non-transparency and incomprehensiveness result in poor trust and inefficient communication between healthcare professionals and patients. Physicians may spend even less time with low-income adults who are already disadvantaged because of the perception of the high quality of care given by the technology, ignoring the subjective and qualitative care that is also critical. Other disadvantaged patients might be less willing to go to the hospital because they think they might be treated differently. The deterministic nature of AI needs to be addressed in order to reduce inequity and provide more well-rounded care that is satisfactory.

The Uncertainties in Regulation

Currently, there is a lack of regulation on AI-based algorithms used in the healthcare system. The most influential social force for the regulation of new technology implementation in the healthcare system is the Food and Drug Administration (FDA), which is responsible for the regulation of anything related to public health, including drugs, biological, and medical devices. Any "software intended to be used for one or more medical purposes without being part of a hardware medical device" will be defined as a medical device by the FDA, including the software used for diagnosis, treatment, and prevention. However, current FDA regulations on medical devices do not apply to AI-based software. Therefore, the first FDA proposal in 2019 on AI-based software introduces a Total Product Life Cycle (TPLC) to access the quality and organizational quality of the company to have reasonable assurance of high-quality AI-based software development, testing, and performing monitoring.

The challenge of bringing AI into healthcare is unique. One contributing factor is that AI algorithms used for disease diagnosis will continue to evolve as the availability of datasets increases after FDA approval because they are designed to continuously learn. For example, the AI algorithm implemented in a hospital may be tailored more toward the needs of the physicians and patients in the local community. The evolving nature of AI requires different types of regulation because the AI algorithm itself will no longer be the end-product, which is usually regulated by FDA and other sister regulators. The pre-existing regulations cannot be applied to AI algorithms due to the dynamic nature (Gerke, Babic, Evgeniou, & Cohen, 2020). Moreover, AI algorithms can present large differences in performance between testing environments and in actual practices. AI algorithms can be more vulnerable to other factors involving humans and the environment. The implementation of AI can be highly affected by the pre-existing workflow, team composition, level of skills, and training (Brynjolfsson & Hitt, 2000). The environment can largely influence the way physicians interact with AI algorithms and how they apply them to patients. Therefore, it is unlikely that the application of AI algorithms will be as consistent as the usage of drugs or medical devices. The variances may potentially make people in the region with low income more disadvantageous due to the limitation on resources, skills, and training.

The current insufficient regulation and the unique nature of AI introduce many uncertainties, making the regulation more difficult and raising new questions on equity in healthcare using AI - How can we ensure the quality of care doesn't change as AI is implemented in hospitals at different demographic locations? How the quality of care can be maintained over time? And if we can solely rely on the FDA, given the dynamic nature of AI algorithms?

Coproduction of Science and Social Order

AI has the full potential to have a positive impact on the healthcare industry. However, it's not the science community's place to claim that one AI software is safe and effective, but many other institutions, which can be examined through the co-production of the science of society. The framework will explore how scientific knowledge can be embedded following social orders. In order to improve the excellence and equality of healthcare, both science and society have the responsibility to minimize algorithmic bias in healthcare. Without adequate social order that regulates changes, AI algorithms can inherit prejudice from prior works and reflect or even exacerbate current bias in society. The ways in which society seeks to organize and control AI will influence the ways AI is implemented in healthcare.

The interactions between science and society are essential to co-production. For example, "race" in science indicates a meaningful biological concept while also containing a social factor.

Both perspectives should be considered to avoid inequity in healthcare using AI. The AI algorithms are designed to ignore the social factors and treat all data equally, yet they create unintentional consequences that result in inequity when ignoring those social factors. Despite that the algorithm isn't intrinsically biased, inequity will be increased without social order.

Until recently, only 29 AI-based software was approved by the FDA for the application of disease diagnosis that is mostly using radiology (Benjamens, Dhunnoo, & Meskó, 2020). This software fulfills the requirements of quality required by the FDA, but trials that ensure equal treatments for the patients are not required. In order to promote equity of AI healthcare, the challenges will be demonstrated by identifying potential sources for algorithmic bias during preliminary research and recognizing the importance of workforce diversity in the social aspect.

Research Method

The representative cases are chosen from the 29 FDA-approved AI-based software, including Arterys Oncology DL, Critical Care Suite, Icobrain, Eko Analysis Software, and DreaMed. The 5 AI-powered software can be used for disease diagnosis in oncology, emergency care, neurology, cardiology, and endocrinology respectively in the United States hospitals. The potential sources of algorithmic bias in disease diagnosis at the current state will be addressed by looking at the datasets used for the algorithm training before the FDA approval. Most of the information is available in scientific journals. Current state workforce diversity in disease diagnosis using AI will be explored by analyzing the management and development team of the representative cases. The number of females and racial minorities was determined using a combination of images and surname search. Images and names of each individual can be accessed through the official websites of the company.

Potential Source of Algorithmic Bias

First of all, an essential factor that contributes to unequal healthcare is biological. Multiple studies have shown the genetic differences that result in an increase in risk factors for minorities. One report suggests that the similarity of traits linked to asthma in European Americans and African Americans is only 5%, which makes African American children 10 times more likely to die from asthma compared to non-Hispanic white children (White et al., 2016). Therefore, data should include a range of gender, ethnicity, races, ages, geographic regions, and prior health conditions.

Given that variances exist between individuals and groups, AI algorithms can create bias because they are based on pre-existing training data at an early development stage. If there is any bias in data availability or the data of the scientists' choice in general, the algorithms would reflect the bias when facing a patient. Especially when the scientific experiments are mostly concerned with isolated variables, the generalizability in a social setting is concerning due to the choice of scientists. The dataset used in the initial AI algorithm design for Arterys Oncology DL, Critical Care Suite, Icobrain, Eko Analysis Software, and DreaMed were studied as shown in Table 1 ("Abstract 12591: Artificial Intelligence Detects Pediatric Heart Murmurs With Cardiologist-Level Accuracy | Circulation," n.d.; Chelu et al., 2016; Nimri et al., 2014; Sauwen et al., 2017).

Table 1. Dataset used for preliminary AI algorithm before FDA approval

	Number of Patients	Age	Female	People of Color (%)
Arterys Oncology DL*	59	39±15 yrs	45.80%	/
Critical Care Suite	/	/	/	/
Icobrain**	35	/	/	/
Eko Analysis Software***	42	/	/	/
DreaMed	24	12-43 yrs	/	/

* Included patients with prior health conditions
**Acquired at the university hospitals of Ghent and Leuven in Belgium
*** Patients with innocent murmurs and other abnormal heart sounds were excluded

The data in Table 1 shows that much information about the dataset is not available in the scientific journal. The publications are not generally concerned with information that is directly relevant to the development of AI algorithms, such as people of color. Companies may also protect their data for trading purposes, which makes it more difficult to identify the potential sources of algorithmic bias. Another source of bias shown in Table 1 is the number of patients involved in the scientific research studies. The dataset being investigated involves only a small group of people ranging from 24 to 59 at the early stage of development. Some of the patients come from the same community. For example, Icobrain used data from patients that are from the university hospitals of Ghent and Leuven in Belgium, which can be problematic because the patients are concentrated in one geographic region. The generalizability of the AI algorithm using this data set is low toward different environments and patients. Finally, although some AI algorithms at the early development stage, such as Arterys Oncology DL, include patients with prior health conditions, many others don't. Eko Analysis Software, an AI-powered stethoscope with the patient and provider software, excluded patients with abnormal heart sound during early scientific research. In summary, science itself is unable to ensure equity in healthcare using AI, and other social orders have to be present to reduce the inequity inherent in scientific research.

Clinical trials can further exacerbate inequities when they disproportionately target the privileged group. One example is the gender imbalance in the trials for cardiovascular diseases, which comprises a population that is 85% male and mostly postmenopausal females (Dougherty, 2011). The underlying reason is the fear of the disruption of standardized results by female menstruation cycle, and the intentional uneven distribution of male and female has negative impacts on disease treatments for females, where some of them are more likely to receive lower levels of treatment for cardiovascular diseases (Bugiardini, Estrada, Nikus, Hall, & Manfrini, 2010). In general, poor, black, or female patients are less likely to receive the medicines they need (Rathore et al., 2000). If the algorithm is trained or tested on this dataset during clinical trials, it will further augment the disparities by decreasing the chance of minorities receiving proper treatments.

Workforce Diversity

AI algorithms need a diverse deep learning experience, which can be promoted in a few ways. First is the inclusion of a diverse group of people and increasing the awareness of diversity during the development stage. Nowadays, most people in AI-related fields are males - 80% of AI professors are males, and 80%-90% of the staff in big AI technology companies are also male - and very few are minorities ("Gender, Race, and Power in AI," n.d.). More diversities in the development team can bring a more comprehensive worldwide perspective to the algorithm to prevent any bias.

Secondly, when designing clinical trials, researchers should include a diverse population of study participants, recruit participants from different practice settings, and collect data on a broad range of health outcomes (Tunis, Stryer, & Clancy, 2003). The cutting-edge biomedical research should not happen in isolation, and more resources should be allocated to translational research to encourage collaboration to effectively decrease the disparity created by independent scientific research based on a homogenous population.

Finally, a diverse collection of algorithms and a group of physicians should be encouraged to work together during the implementation phase. The different results by multiple algorithms can help scientists identify aspects that have been previously overlooked and provide significant feedback for improvements. Additionally, Identifying the formation of algorithmic bias is not only important for creating equality, but also to provide insights into how physicians and AI algorithms can work together. There can be cases where AI algorithms make most of the decisions for the majority of the privileged population, where physicians can be more focused on interventions for the minority population.

Table 2. Workforce diversity of the representative FDA approved AI algorithm

	Total number	Female	People of Color	African/African Americans	Asian/Asian Americans
Arterys Oncology DL	11	1	2	0	1
Critical Care Suite	23	3	6	1	5
Icobrain	54	20	4	1	1
Eko Analysis Software	18	3	5	0	5
DreaMed	12	3	1	0	0
Total	118	30	18	2	12
Total (%)	100%	25.40%	15.30%	1.69%	10.20%
US Demography by 2044	100%	50.00%	50.30%	12.00%	7.90%

The current workforce lacks underrepresented minorities and the creativity in reducing inequities in healthcare using AI. Table 2 includes the information from the leadership teams in management and development from the representatives using AI algorithms approved by FDA. US demography by 2044 is used for comparison because of the rapid expansion of racial minorities and immigrants, and the prediction will be used for comparison. Among all 118 people, only 30 are female, which is composed of 25.40% of the leadership teams in management and development. Racial minorities are underrepresented, as only 15.30% of the team members are racial minorities. African/African Americans are especially underrepresented in the team as only 2 out of 118 people are African/African Americans, which is only 1.69%. On the contrary, the Asian/Asian Americans are not underrepresented, as 10.20% of the team members are Asian/Asian Americans, which is higher than the demography predicted by 2044. However, the team is still disproportionately white males, as 59.3% of the leadership team are white males. Therefore, the lack of underrepresented minorities is not only in the imaging informatics from a scientific perspective, but also in the social dimension.

Discussion

The analysis done on the current state of using AI algorithms for disease diagnosis shows that AI algorithms are learning through a view that is narrow in focus as the dataset used in early research is not representative of a diverse population. At the same time, the current leadership is not strong enough to address bias and diversity from the start. The co-production of science and social order is needed to ensure that the improvement of healthcare will not only benefit the high-income racial majorities but also minorities living in remote areas.

A scientific solution is to use data augmentation to enrich the dataset by artificially modifying the available dataset for disease diagnosis using radiology (Shorten & Khoshgoftaar, 2019). The dataset can be expanded in order to improve the performance of AI algorithms and their ability to generalize. Certain features can be added to accommodate the modalities adapted at certain hospitals. Other technical solutions are also able to help achieve generalization. For example, Grad-CAM provides “visual explanations” for the decision-making process of AI algorithms by highlighting the area with the most “attention” (Selvaraju et al., 2019). The increased explainability lends insights into how and why the AI algorithm does not work on certain cases, facilitating researchers to identify dataset bias.

The inequity created by diagnostic AI algorithms in healthcare also results from the big gap between the narrow focus of the current dataset and the worldview. The missing information from the early developmental stage of the 5 representative cases suggests the need for creating datasets that are more well-rounded for more equitable AI algorithms. Therefore, data collection should be improved to include personal characteristics, such as ethnicity, gender, age, disability, socioeconomic status, and geographic location, which can facilitate in documenting, tracking, and understanding the inequity created by diagnostic AI algorithms (Kilbourne, Switzer, Hyman, Crowley-Matoka, & Fine, 2006). The efforts for better data collection cannot solely rely on regulations, but it is a collective effort. Improved data collection techniques will be needed starting from the early scientific research and clinical trials to the implementation of diagnostic AI algorithms in the hospitals. Only a coordinated system can provide insights into algorithmic bias due to limited datasets and develop interventions to reduce disparities.

Another factor revealed by the 5 representative cases that increase inequity is the current low workforce diversity in the leadership of the management and development teams. Healthcare workforce diversity is important to increase access to care to disadvantaged populations by providing better opportunities for minorities to see providers of their own race and to improve adherence to equity through better leadership and policies aimed to serve vulnerable populations (Williams, Walker, & Egede, 2016). Healthcare professionals and scientists should be required to be trained in a more diverse environment that not only increases knowledge and skills to improve the usage of diagnostic AI tools, but also more equitable care and access for patients. It has been shown that a more culturally diverse training environment has a positive relationship with improved patient outcomes, which is achieved primarily through greater access for minorities (Lie, Lee-Rey, Gomez, Bereknyei, & Braddock, 2011; US Department of Health and Human Services, n.d.). In summary, workforce diversity needs to be increased through implementing a more culturally diverse environment by introducing more minorities in the leadership or providing training, policies, and programs that support the vulnerable populations.

The complementary strategy to reduce inequity requires diagnostic AI algorithms to be embedded in society following the technical and cultural expectations. The collaborative intelligence between scientists and leaders in the healthcare industry is essential.

Conclusion

Using diagnostic AI algorithms can extend disparities already existing in the current healthcare system. This thesis incorporates the co-production of science and social order and aims to contribute to the successful implementation of diagnostic AI algorithms that ensure equity in the healthcare industry. New technical developments can offer solutions to current challenges by enriching the available dataset and provide more insight into the potential sources of inequity, but the social significance should also be recognized and improved. The social dimensions of technological advancement can also facilitate development that improves equity, such as determining methods to promote diversities. In conclusion, diagnostic AI algorithms are the potential solution to current challenges of inequity in the healthcare industry, and social accommodation is part of the solution for successful implementation.

Bibliography

Abstract 12591: Artificial Intelligence Detects Pediatric Heart Murmurs With Cardiologist-Level Accuracy | Circulation. (n.d.). *Circulation*.

Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *npj Digital Medicine*, *3*, 118.

Berchick, E., Hood, E., & Barnet, J. (n.d.). Current Population Reports.

Braithwaite, J., Hibbert, P., Blakely, B., Plumb, J., Hannaford, N., Long, J. C., & Marks, D. (2017). Health system frameworks and performance indicators in eight countries: A comparative international analysis. *SAGE open medicine*, *5*, 2050312116686516.

Brynjolfsson, E., & Hitt, L. M. (2000). Beyond computation: information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, *14*(4), 23–48.

Bugiardini, R., Estrada, J. L. N., Nikus, K., Hall, A. S., & Manfrini, O. (2010). Gender bias in acute coronary syndromes. *Current vascular pharmacology*, *8*(2), 276–284.

Chelu, R. G., van den Bosch, A. E., van Kranenburg, M., Hsiao, A., van den Hoven, A. T., Ouhlous, M., Budde, R. P. J., et al. (2016). Qualitative grading of aortic regurgitation: a pilot study comparing CMR 4D flow and echocardiography. *The International Journal of Cardiovascular Imaging*, *32*(2), 301–307.

Cook, C. M., Howard, J. J., Sirotin, Y. B., Tipton, J. L., & Vemury, A. R. (2019). Demographic effects in facial recognition and their dependence on image acquisition: an evaluation of eleven commercial systems. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, *1*(1), 32–41.

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare.

Future healthcare journal, *6*(2), 94–98.

Dougherty, A. H. (2011). Gender balance in cardiovascular research: importance to women's health. *Texas Heart Institute Journal*, *38*(2), 148–150.

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., et al. (2019). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24–29.

Gender, Race, and Power in AI. (n.d.). .

Gerke, S., Babic, B., Evgeniou, T., & Cohen, I. G. (2020). The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *npj Digital Medicine*, *3*, 53.

Health Insurance Coverage in the United States: 2017. (n.d.). . Retrieved March 11, 2021, from https://www.census.gov/library/publications/2018/demo/p60-264.html

Kilbourne, A. M., Switzer, G., Hyman, K., Crowley-Matoka, M., & Fine, M. J. (2006). Advancing health disparities research within the health care system: a conceptual framework. *American Journal of Public Health*, *96*(12), 2113–2121.

Kim, E. J., Kim, T., Conigliaro, J., Liebschutz, J. M., Paasche-Orlow, M. K., & Hanchate, A. D. (2018). Racial and ethnic disparities in diagnosis of chronic medical conditions in the USA. *Journal of General Internal Medicine*, *33*(7), 1116–1123.

Lateef, F. (2011). Patient expectations and the paradigm shift of care in emergency medicine. *Journal of Emergencies, Trauma and Shock*, *4*(2), 163–167.

Lie, D. A., Lee-Rey, E., Gomez, A., Bereknyei, S., & Braddock, C. H. (2011). Does cultural competency training of health professionals improve patient outcomes? A systematic review and proposed algorithm for future research. *Journal of General Internal Medicine*, *26*(3), 317–325.

McCullough, C. (n.d.). Who is least likely to attend? An analysis of outpatient appointment DNA data in NHS Greater Glasgow & Clyde.

Miller, D. D. (2019). The medical AI insurgency: what physicians must know about data to practice with intelligent machines. *npj Digital Medicine*, *2*, 62.

Nimri, R., Muller, I., Atlas, E., Miller, S., Fogel, A., Bratina, N., Kordonouri, O., et al. (2014). MD-Logic overnight control for 6 weeks of home use in patients with type 1 diabetes: randomized crossover trial. *Diabetes Care*, *37*(11), 3025–3032.

Rathore, S. S., Lenert, L. A., Weinfurt, K. P., Tinoco, A., Taleghani, C. K., Harless, W., & Schulman, K. A. (2000). The effects of patient sex and race on medical students' ratings of quality of life. *The American Journal of Medicine*, *108*(7), 561–566.

Rodriguez-Ruiz, A., Lång, K., Gubern-Merida, A., Broeders, M., Gennaro, G., Clauser, P., Helbich, T. H., et al. (2019). Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. *Journal of the National Cancer Institute*, *111*(9), 916–922.

Sauwen, N., Acou, M., Bharath, H. N., Sima, D. M., Veraart, J., Maes, F., Himmelreich, U., et al. (2017). The successive projection algorithm as an initialization method for brain tumor segmentation using non-negative matrix factorization. *Plos One*, *12*(8), e0180268.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2019). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International journal of computer vision*.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of big data*, *6*(1), 60.

Tunis, S. R., Stryer, D. B., & Clancy, C. M. (2003). Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *The Journal of the American Medical Association*, *290*(12), 1624–1632.

US Department of Health and Human Services. (n.d.). The rationale for diversity in the health professions: A review of the evidence.

What was reached by artificial intelligence in 2016 | Earth Chronicles News. (n.d.). . Retrieved March 11, 2021, from https://earth-chronicles.com/science/what-was-reached-by-artificial-intelligence-in-2016.html

White, M. J., Risse-Adams, O., Goddard, P., Contreras, M. G., Adams, J., Hu, D., Eng, C., et al. (2016). Novel genetic risk factors for asthma in African American children: Precision Medicine and the SAGE II Study. *Immunogenetics*, *68*(6–7), 391–400.

Williams, J. S., Walker, R. J., & Egede, L. E. (2016). Achieving equity in an evolving healthcare system: opportunities and challenges. *The American Journal of the Medical Sciences*, *351*(1), 33–43.