Improving Public Perception of Artificial Intelligence in Medicine

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

Xinyuan Zhu

Spring 2022

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

Kathryn A. Neeley, Associate Professor of STS, Department of Engineering and Society

Introduction

Healthcare in the United States is an extremely complex system that is plagued by inefficiency and economic inaccessibility (UCLA 2014, p. 9). Wasteful spending in administration, treatment, and poor coordination is the primary cause of the problem. As a result, the United States is ranked lowest among other countries for healthcare spending efficiency. Figure 1 depicts this by showing that the United States has significantly more average spending per person on healthcare but has the lowest life expectancy compared to other countries. Figure 1 also shows that healthcare inefficiency is not a recent problem, it has been prevalent for over the past 20 years.

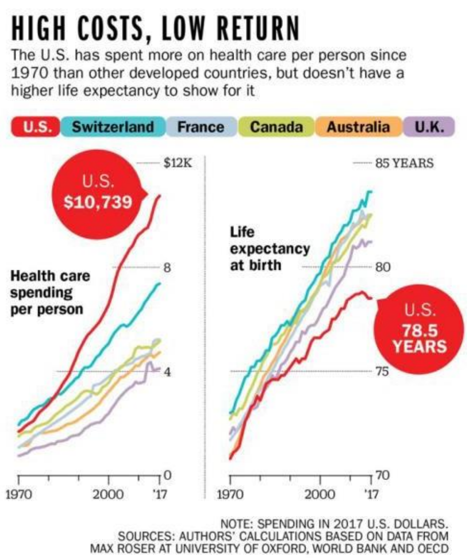

Figure 1: Life Expectancy and Average Spending Graph (Case & Deaton, 2020)

To help mitigate this problem, a significant amount of resources have been allocated towards the research and development of new technologies to provide lower treatment costs and improve workflow (Carayon, 2009). Of these, artificial intelligence (AI) is one of the technologies with the highest potential to improve the overall healthcare system (Bohr, 2020).

Eric Topol, a cardiologist, acknowledges the potential of AI systems to revolutionize healthcare by stating

> The promise of artificial intelligence in medicine is to provide composite, panoramic views of individuals' medical data; to improve decision making; to avoid errors such as misdiagnosis and unnecessary procedures; to help in the ordering and interpretation of appropriate tests; and to recommend treatment (Topol, 2019, p. 18).

Despite AI having been used in medical applications since the early 2000s, there is still poor public reception. If the population stays on the current trend of increasing average spending per person on healthcare, more people will be alienated from being able to receive any adequate treatment. As a result, only the interests of large corporations will be served with higher profit margins and continuous upcharge on medical products.

This paper will go over why the general perception of AI systems is poor, despite the evidence that shows AI is sophisticated enough to replace doctors and reduce healthcare spending overall. The principal-agent problem will be used to answer the question on the leading cause of why AI is poorly received. From this, a resolution will be proposed.

AI Systems Have Poor Public Perception

Technological and Economic Efficacy

AI systems in medicine present a much faster and cheaper way to perform healthcare related tasks to the benefit of all parties involved, with results comparable to or exceeding human capability. As previously stated in the introduction, extensive studies have been conducted recently that demonstrate a high accuracy level of AI systems. A preliminary systematic review and meta-analysis found that AI systems had the same level of accuracy as humans based on medical imaging (Sandiou, 2019). Some studies have even shown that AI systems can outperform humans. In a peer-reviewed study from Babylon Health and the University College London, an improved AI model had an accuracy of 77.26%, ranking it in the top 25% of doctors (Leibowitz, 2020). The key advantage that AI systems provided over humans was humans being able to compute many different possibilities much faster and access databases more efficiently.

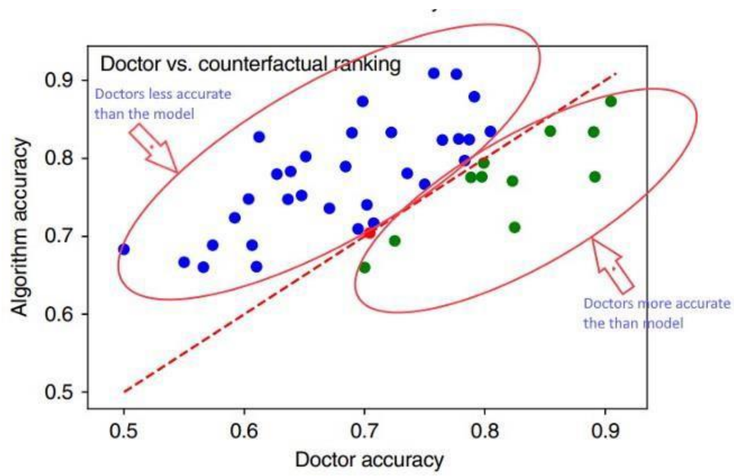

Figure 2: The performance of AI vs Doctors (Leibowitz, 2020).

Extensive research and data analysis has also shown economic benefits to using AI systems. Costs are improved for both developments and patient costs. For pharmaceuticals, use of AI is projects to save up to $54 billion annually (Paul, Sanap, Shenoy, Kalyane, Kalia & Tekade 2020). In general healthcare, AI is projected to save up to $150 billion annually by 2026 (Paul, Sanap, Shenoy, Kalyane, Kalia & Tekade 2020).

The Social Potential of AI Systems

In the book *Deep Medicine: How Artificial Intelligence Can Make Health Care Human*

Again (Topol, 2019), cardiologist Eric Topol argued that medicine currently lacks compassion and empathy for the patients because doctors are too overwhelmed by work to create any meaningful connections. The workload of doctors also makes them prone to medical errors. By introducing AI to decrease the medical expertise required by all doctors, they should have more time to focus on a more personalized experience for the patients. Topol sees restoring a social connection between doctors and patients as the greatest potential of AI by stating:

> The greatest opportunity offered by AI is not reducing errors or workloads, or even curing cancer: it is the opportunity to restore the precious time-honored connection and trust – the human touch – between patients and doctors. Not only would we have more time to come together, enabling far deeper communication and compassion, but also, we would be able to revamp how we select and train doctors… This leveling of the medical knowledge landscape will ultimately lead to a new premium: to find and train doctors who have the highest level of emotional intelligence (Topol, p. 26)

Lapse in Technology and Acceptance

Despite the research that shows AI can revolutionize healthcare, public acceptance of having healthcare handled by AI shows general trends of distrust and unacceptance. In a study of over 200 students from Boston University and New York University, students presented with the opportunity to take a free assessment of stress levels showed that 40% signed up when told a doctor would be administering the test while 26% signed up when told an artificial intelligence would be performing the diagnosis. Additionally, when given the choice between having a more expensive option of a doctor performing the diagnosis and a cheaper option of having an artificial intelligence do it, participants were more likely to choose the doctor. In both studies, participants were informed that the accuracy rate of the doctor and the artificial intelligence were equal (Longoni, 2019). In another study involving 1909 participants of varying statuses of age, education, and occupation, a series of statements were evaluated by the participants, who would then assign a score of 1-5 based on how strongly they agreed with a statement, with 5 being the strongest agreement. The statements were skewed towards favoring AI, such as “Even if computers are better at evaluating scans, I still prefer a doctor” and “The sooner I get the results, even when this is from a computer, the more I am at ease.” (Yakar, Ongena, Kwee, and Haan 2022, p. 3). The individual scores were then combined into an aggregate, represented by a general score of 1-5 for overall trust in AI, with 5 being the highest level of trust. Radiology and dermatology had an average score of 2.82 for trust while robotic surgery scored lower in 2.75 (Yakar, Ongena, Kwee, and Haan 2022). Even with statements skewed in favor of AI, participants exhibited general distrust. The study also demonstrates that simply telling an individual that AI is more effective than humans is insufficient in gaining their trust.

Filling The Knowledge Gap

Understanding the dissonance between the sophistication of AI and its adoption for public use is not achievable by solely performing a literature review; it requires analysis through a framework to synthesize an argument. Given the proven effectiveness of the technology, the current state of technology does not contribute to the poor public reception. Therefore, the purpose of this research is to determine primary factors outside of technology that inhibit more widespread acceptance using the Principal-Agent Problem. Through this, I will cover better ways to alleviate public concerns about placing their trust in AI for healthcare since previous literature shows that simply presenting evidence is not sufficient.

Methods

The principal-agent problem is a useful analysis tool in this context because it focuses on the difference of interests within groups of people and can provide further insight into the consequences of such. There are two main groups involved, one is known as the “principals” which delegates authority. The other group, “agents,” are the ones to act and make the decisions with the resources given by principals. A problem happens when both parties have conflicting interests where an agent can act at the expense of the principal (Pureza, n.d.). A prime exampleof this would be in politics, where the voters are the principals, who elects a candidate to the office as the agent. There is a conflicting interest since the voters want what is good for themselves while the candidate must act in the interests of the party. Figure 3 depicts the relationship between the two parties.

Figure 3: There will almost always be a conflict of interest between different parties, potentially causing a lapse in decisions that would negatively affect the agent (Pureza, n.d.).

Applying the Principal-Agent Problem to Engineering

While the Principal-Agent problem is primarily used in the world of finance, its concepts have a wide variety of applications to engineering. A paper written by Professor Anita Ceric Titled "Communication Risk in Construction Projects: Application of principal-agent theory" (Ceric, 2012) demonstrates an active connection between the Principal-Agent problem and civil engineering. In this journal, the main cause of the conflict of interest is caused by asymmetric information; defined by separate parties having different information at the same time. Parties facilitate asymmetric information by purposefully not sharing the hidden information because of self-interest. The asymmetric information between parties is split up into different types of risk, known as hidden characteristics, hidden information, and hidden intentions. "Hidden characteristics are associated with selection; hidden action and/or hidden information are associated with moral hazard; and hidden intentions are associated with hold-up." (Ceric 2012, p.769) Each hazard occurs at different stages of a construction project. Hidden characteristics occur when two parties do not have full information about each other prior to contracting or hiring.

Hidden action occurs after a working relationship between two parties is established and hidden action creates a moral hazard risk. A concrete example is given by Ceric when she states "For instance, the client cannot be sure that firms, once hired, will fully mobilize their capabilities on the client's behalf or on behalf of other clients of theirs. (Ceric, 2012 p.769)" Finally, hidden intentions can occur at any stage of the project, which causes delays and hold-ups. Hidden intention can happen when either the contractor or client give more resources to the other party with the intention that they will use them to benefit the project, but the party given the resources decides to act differently.

Based on the Delphi method, Ceric concludes the most important parties within the principal-agent problem are two agents that do not have a direct contract with each other. Minimizing the asymmetric information between those two parties will reduce the risks and hazards involved in the principal agent problem. There are two main ways to reduce asymmetric information between parties, which are screening and monitoring. Screening gathers more information about an agent before a contract is made while monitoring is to make sure that agents are acting in accordance with the contract after a contract has been made.

From this analysis, Ceric provides an expanded model of the principal agent problem, shown in figure 4 that encompasses more parties and shows how one party can act as both the principal and the agent.

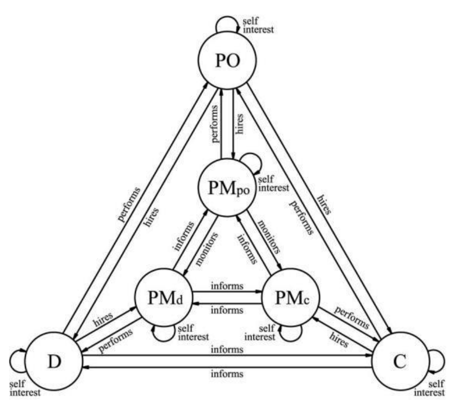

Figure 4: Principal-agent framework applied to construction projects, where the parties involved are PO: Project Owner, C: Contractor, D: Designer, PM_{PO}: Project owner's project manager, PM_d : Designer's project manager, PM_C: Contractor's project manager (Ceric 2012)

Using the Principal-Agent Problem in Medicine and Artificial Intelligence

Ceric's conclusion of needing to minimize the asymmetric information between two parties of the principal agent problem will be the basis for gathering results. The parties she identified for construction managers, project owners, and contractors can be substituted with developers, shareholders, and patients to view asymmetric information overlaps. Ceric initially only used the Principal Agent problem for analyzing construction staff, but her argument about the most important relationship discussed in the previous paragraph pairs well with patients and developers. Analyzing this and applying screening and monitoring will be used to propose a solution.

To apply the Principal-Agent problem to the use of AI in medicine, the relevant populations will need to first be identified in order to assign each a “principal” or an “agent” status based on their relationship to other groups. The paper Bench to bedside: The technology adoption pathway in healthcare (Clark, Dean, Bolton, Beeson 2019) will be used since it provides a clear and concise timeline that explicitly states the different groups used.

After each population related to AI in medicine has been identified, the relationship between each population will be analyzed with a one-on-one basis to determine the principal and agent pairing. Within this analysis, the conflicts of interests and asymmetric information will be identified. Once the connections between populations have been found, the two parties without any direct contract can be found. Finally, ways of decreasing the asymmetric information will be proposed to create a resolution.

Results

Identification of Parties

Figure 5 depicts a broad overview of the path that medical technology takes before widespread implementation to the public. This is split up into seven stages and viewing the people affected by each stage would show the best populations to analyze through the principal agent problem.

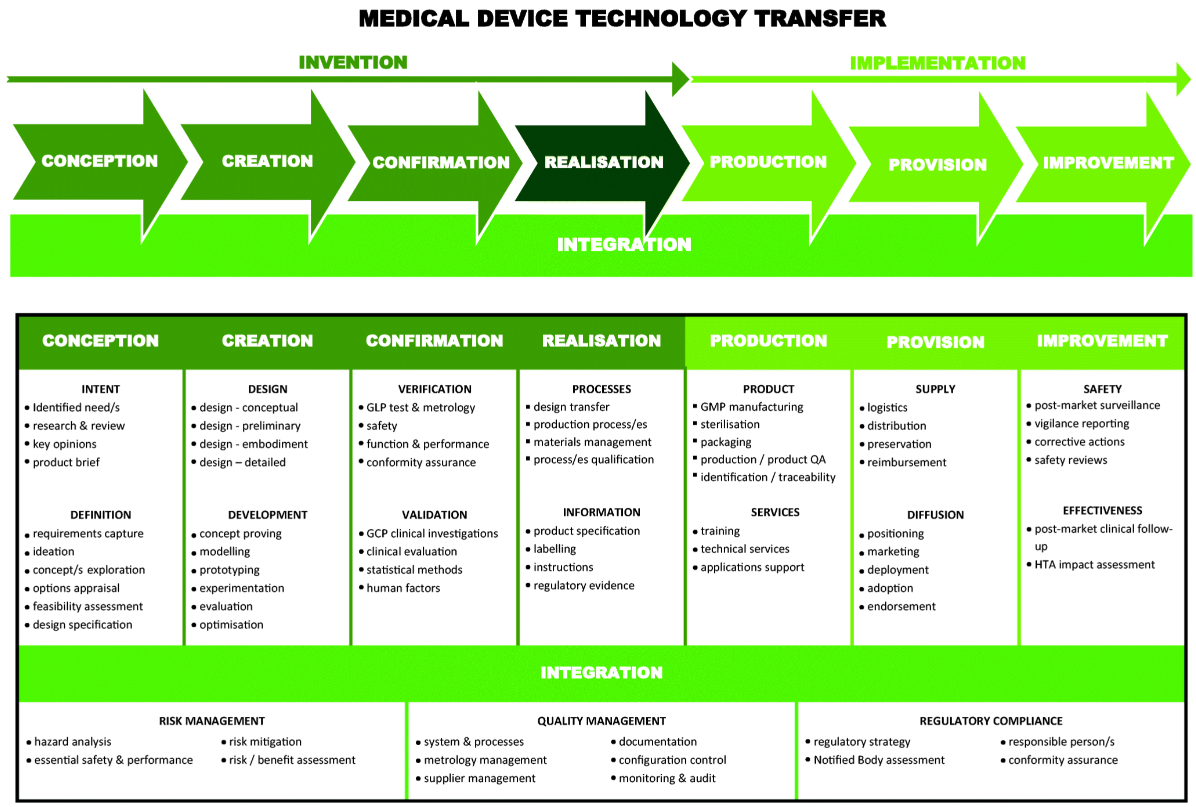

Figure 5: The timeline of medical technology from inception to implementation (Clark, Dean, Bolton, and Beeson 2019)

Based on the report made, the following populations were identified to be used within the framework.

Developers – This is the population that creates, maintains, and updates artificial intelligence systems within hospitals. Their significance in this process is the prevalence throughout all of the stages in the timeline, being responsible for creation and refinement of technology.

Patients – This is the population that receives treatment from new technologies. Their significance in the process is being involved in the validation of a technology and vital for measuring effectiveness.

Hospital Staff – The population that uses AI for treatment options. They are the facilitators of technology that will reach patients and require specialized training.

Government – The government is the authority figure that authorizes new technology to be used for patients and conducts independent analysis to determine efficacy and compliance. However, their influence on public perception is minor compared to developers and hospital staff.

Patients as a Principal and Hospital Staff as an Agent

In the interaction between patients and hospital staff, the patients are the principals because they give hospitals money. In return, the hospital staff should act in the best interest of the patient by providing adequate treatment options. Asymmetric information is prevalent here, which creates the risks previously defined by Ceric. As Topol observed throughout his time working as a cardiologist, hospital staff often are overwhelmed and will prioritize speed over accuracy. Patients are often unaware of this risk as they believe that doctors have their best interest in mind, which is a hidden action problem.

Hospital Staff and Developers as principals and agents

In these two groups, the hospital staff are the ones that become the principals because the creation of new medical technology are based on their needs. This forms a consumer supplier relationship between the hospital staff and the developers. Hidden action is the most prevalent risk here because of deceptive practices from developers in recent years. For example, IBM introduced Watson in 2014, an artificial intelligence system focused on cancer treatments. Watson promised to provide an accurate list of diagnoses given patient symptoms. Each diagnosis from Watson would be given a confidence score to help doctors make more accurate judgements. However, the results from real-world applications have been poor. Watson demonstrates a hidden action from when hospitals purchased licenses for Watson, expecting IBM to be acting in making healthcare better, but the reality was they were in the market looking for a cash-grab (Ross, 2016).

This relationship can also be reversed, with developers acting as principals and hospital staff acting as agents. When a technology is developed and given to hospitals, the developers are giving hospitals the authority to make sure that their products are used in an ethical manner and follows proper operating protocol to ensure the safety of all parties.

Overall Principal-Agent Network

The previous sub-sections identified the pairs of parties that make up the principal agent problem in the context of AI in medicine. When combined together, the populations form Figure 6 to provide the network used for analysis.

Figure 6: Major parties involved in the principal-agent problem

Most Important Relationship

As Ceric stated, the most important relationship is the one between two agents that have no contract with each other. Figure 6 shows that patients and developers do not have a direct contract with each other. Their contract is facilitated through hospitals. Therefore, minimizing asymmetric information between patients and developers is paramount to increasing public approval.

Within many engineering firms, a non-disclosure agreement must be signed for confidentiality due to competition that rises with other firms. Due to this, it is nearly impossible for a patient to understand the full technical details of an AI system. Likewise, there are barriers for developers to gain more information about patients due to medical confidentiality, which restricts the flow of information to developers. The lack of information for either side creates a hidden characteristic problem.

The most important risk present here is the hidden intention. As discussed previously with the IBM Watson case, IBM had the intention of making a larger profit over the functionality of the system, which was unknown to the hospital staff. The resulting inefficacy of this system further promoted distrust within patients and AI systems.

To lessen the negativity regarding AI systems to the patients, more information on the use of AI systems must be more readily available.

Limitations

Although the principal-agent problem can be used to improve the public perception of AI, it only plays a small role in the widespread adoption of technology. The framework is mainly useful when looking at small groups of tangible populations. Only human populations are considered, and the specific technology is not considered an equal. Technology on exists to be part of the conflict of interest between parties. Therefore, examining social factors such as human interaction would be more effective using a different framework, such as actor-network theory.

Conclusion

The solutions to resolve asymmetric information from Ceric can be applied here to create a resolution that will help improve the public perception of AI systems. As the curriculum in most engineering classes show, most of the skills and knowledge that are learned are done through a process of self-discovery through challenging assignments and projects. A similar method of approaching screening can be done for better effectiveness. Rather than telling someone to just trust AI systems and that they are accurate, a better way would be to give them more resources to find out the information by themselves to help bolster the knowledge that they gain. While it is impossible to get every person to agree with a topic, this strategy seeks to be a more effective way to improve general perception.

Monitoring is more complicated and can be done in multiple ways. A concern that was found throughout this research is that there was no easy way for patients to get information easily and readily, having to rely mainly on third-party sources that were prone to bias. Therefore, a continuously updated AI tracking system maintained by an official, such as the government would be monumental in providing better information to the public about AI, with data such as success rate, time to perform, and type of treatment. Another way to achieve monitoring would be for a doctor to present patients with their results and an AI's results on an individual basis. Generally, people like to think of themselves as the "main character" so making the revelation of knowledge more personalized would help.

The main point shared between monitoring and screening is to decrease the amount of asymmetric information by making information more readily available. If the public opinion of AI in healthcare were to improve, then demand would grow and this technology would be adopted widespread more quickly, providing greater access to cheaper and quicker healthcare.

References

Ahuja A. S. (2019). The impact of artificial intelligence in medicine on the future role of the physician. PeerJ, 7, e7702. https://doi.org/10.7717/peerj.7702

Anonymous (2019, July 29). Future AI Opportunities for Improving Care Delivery, Cost, and Efficacy. Health IT Analytics. https://healthitanalytics.com/news/future-ai-opportunities- for-improving-care-delivery-cost-and-efficacy

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. Journal of medical Internet research, 22(6), e15154. https://doi.org/10.2196/15154

Bohr, A., & Memarzadeh, K. (2020). The rise of artificial intelligence in healthcare applications. Artificial Intelligence in Healthcare, 25–60. https://doi.org/10.1016/B978-0-12-818438-7.00002-2

Bord, J., Bruke, W., Dudzinski, D. Confidentiality. University of Washington Department of Bioethics and Humanities. https://depts.washington.edu/bhdept/ethics-medicine/bioethics-topics/detail/58#:~:text=Confidentiality%20is%20one%20of%20the,is%20provided%20by%20the%20patient

Carayon P, Karsh B-T, Cartmill RS, et al. Incorporating health information technology into workflow redesign - summary report (10-0098-EF) Agency for Healthcare Research and Quality. October 2010. https://digital.ahrq.gov/ahrq-funded-projects/incorporating-health-information-technology-workflow-redesign#nav-publications

Case, A., & Deaton, A. (2020 February 20). How Healthcare Costs Hurt American Workers and Benefit the Wealthy. *Time.* https://time.com/5765845/health-care-problems-in-america/

Ceric, A. (2012). Communication risk in construction projects: Application of principal-agent theory. Organization, Technology, and Management in Construction, 4(2). 10.5592/otmcj.2012.2.8

Clark, D., Dean, G., Bolton, S. et al. Bench to bedside: The technology adoption pathway in healthcare. Health Technol. 10, 537–545 (2020). https://doi.org/10.1007/s12553-019- 00370-z

Landi, H. (2019, April 25). Nearly half of U.S. doctors say they are anxious about using AI-powered software: Survey. Fierce Healthcare. https://www.fiercehealthcare.com/practices/nearly-half-u-s-doctors-say-they-are-anxious- about-using-ai-powered-software-survey

Leibowitz, D. (2020, September 29) AI Now Diagnoses Disease Better Than Your Doctor, Study Finds. Towards Data Science. https://towardsdatascience.com/ai-diagnoses-disease-better- than-your-doctor-study-finds-a5cc0ffbf32

Longoni, C. & Morewedge, C. (2019, October 30) AI Can Outperform Doctors. So Why Don't Patients Trust It? Harvard Business Review. https://hbr.org/2019/10/ai-can-outperform- doctors-so-why-dont-patients-trust-it

Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., & Tekade, R. K. (2021). Artificial intelligence in drug discovery and development. Drug discovery today, 26(1), 80–93. https://doi.org/10.1016/j.drudis.2020.10.010

Pureza, G. M. (2022, January 28). Principal-agent problem. WallStreetMojo. https://www.wallstreetmojo.com/principal-agent-problem/

Robbins, R., & Brodwin, E. (2020, July 31). An invisible hand: Patients aren't being told about the AI systems advising their care. STAT. https://www.statnews.com/2020/07/15/artificial-intelligence-patient-consent-hospitals/

Roy, K. (2019, September 16) Healthcare in the U.S. is inefficient because of wasteful spending. The Diamondback. https://dbknews.com/2019/09/16/healthcare-industry-costs-waste-spending-drugs/#:~:text=Wasteful%20spending%20is%20made%20up,in%20treatment%20and%20clinical%20waste.&text=Because%20prices%20aren't%20standard,services%20greater%20than%20their%20value

Ross, C. (2018, February 16). IBM pitched its Watson supercomputer as a revolution in cancer care. it's nowhere close. Stat News. Retrieved March 25, 2022, https://www.statnews.com/2017/09/05/watson-ibm-cancer/

Topol, E. (2019) Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again

University of Los Angeles Health (2014, Spring). U.S. Ranks Near Bottom in Efficiency of Healthcare Spending. U Magazine, 34(2), 9.

Yakar, D., Ongena, Y. P., Kwee, T. C., & Haan, M. (2022). Do people favor artificial intelligence over physicians? A survey among the general population and their view on Artificial Intelligence in medicine. Value in Health, 25(3), 374–381. https://doi.org/10.1016/j.jval.2021.09.00