Algorithmic decision-making in healthcare:

Exploring the ethical considerations of using algorithms for healthcare decisions, such as patient diagnosis, treatment recommendations, or resource allocation.

Sunil Chomal, Jyoti Vasudev, Brendon Gory

August 3, 2023

CSCI E-184 Data Science and Artificial Intelligence Ethics, Governance, and Laws (Summer 2023)

Harvard University Extension School

Prof. Bruce Huang

Table of Contents

Abstract ... 3
Introduction ... 3
Analysis-Benefits ... 4
Promoting Equity in Healthcare ... 4
AI tools can lessen racial bias in pain treatment ... 5
Telemedicines / Telehealth using AI ... 6
Analysis-Drawbacks ... 7
Difficulty with Public Trust in AI ... 7
How AI Models Could Fail to Detect Skin Cancer in Diverse Populations ... 8
"Work with Me, Don't Just Talk at Me - When "Explainable" Is Not Enough" ... 9
Case Study 1 – UK[15] ... 10
Case Study 2 – USA[16] ... 11
Public Perception ... 14
Introduction ... 14
Data source ... 14
Potential Biases ... 15
Data Collection ... 15
Data Analysis ... 17
Key findings ... 18
Safeguards ... 18
Recommendations ... 19
References ... 20

Table 1: AI chatbot applications 7
Table 2: Case study one (key) responses 8
Table 3: Correlations of openness, concern, and benefit scores with all study variables (Case study 2) 11
Table 4: Data source for reddit study 12
Table 5: Sentiment Analysis 15

Figure 1: Typical AI lifecycle 5
Figure 2: A schematic actor network based on patient clinician interactions for clinical care 9
Figure 3: Complete advanced technology-enabled healthcare socio-technical system 9
Figure 4: Mean scenario openness scores with 95% CIs. Large vertical lines indicate grand mean with 95% CI. (Case study 2) 10
Figure 5: EDA for case study 2 11
Figure 6: Reddit word cloud 17

Abstract

In the modern era, AI has become more prevalent in healthcare. Reasons stem from better efficiency, farther reach, and overcoming staffing shortages, et al. AI has a use in healthcare just like any industry. If governance, fairness and equity remain unchecked alongside the rise of AI in healthcare, the benefits will unethically serve the privileged few.

The arguments in this paper explore benefits for and drawbacks with AI in healthcare. Recommendations will also be outlined that address the shortcomings and faults of the problems currently faced using AI in healthcare. Part of developing trust is to promote the fairness and equity AI can offer. As opinions of AI in general continue to develop, healthcare's AI protocol must be subject to the same scrutiny. The research of others corroborates the paper's discussions.

Introduction

There still exists distrust in how most people think of AI in healthcare[1]. Despite the growing accuracy and how much healthcare professionals rely on AI, some of the public do not feel comfortable with the results given by an algorithm; the trust remains with the medical professional.

Not all individuals receive the same treatment for pain or illness. Disparities will be explored as well as solutions AI can offer. Telemedicine can be a great way for better efficiencies in healthcare, and sometimes, necessary as seen with the COVID-19 pandemic. Drawbacks do exist with the technology, such as failure to detect illness or disease based on a person's race. Understanding the challenges now with balancing the data and improving algorithms can quickly eradicate these issues. It is important to remain vigilant with the new technology so that inappropriate treatment does not settle in as a price to pay for using AI. Equity can, and must, be achieved.

A main concern is the equity of all those who seek medical care. For the first time in a long while, we have the chance to reshape bias and equity within healthcare. However, without proper jurisprudence, these biases and unfairness that exist in medical treatment may continue. Arguments for better governance will be raised. Other emerging technologies receive proper oversight. Healthcare's adoption of AI should be no exception. In fact, it has been proposed that integrating patient and public involvement (PPI) in healthcare AI may help in adoption and acceptance of these technologies[6].

Outside of the opinions written by academics and journalists, research was conducted using Reddit comments for specified Reddit groups (known as subReddits). Natural Language Processing (NLP) produced sentiment analysis as well as a word cloud highlighting the most used words in the forums. The analysis gives context to the thoughts compiled surrounding the ethics and bias researched from other contributors. Though the owners of the content have their own bias (favorable view of technology on a technological forum, views gather from people who know how to use Reddit, people with Internet access, etc.), the results shed light on the public perception of this nascent, growing field.

Analysis-Benefits

Promoting Equity in Healthcare

Bias in algorithms and data applies to the healthcare industry. Data can mislead or misrepresent populations while the algorithms perpetuate the bias. The bias can lead to unfairness in how patients are treated. Establishing equity by using AI benefits the currently underprivileged when it comes to healthcare. This does not mean just poor, aging populations, minorities, and gender bias play a part in unintended bias by healthcare providers. Understanding the existing biases is the beginning of the solution.

Dr. You Chen, PhD argues using human-centered AI (HCAI) will help mitigate existing biases[2]. A "diverse group of stakeholders" oversees the development of AI in healthcare. This includes measuring where biases occur and if algorithms perpetuate existing biases. The source of unfairness does not lie specifically in an algorithm or the data on which it operates. Proper management of the AI, from levels of administration to development, will better serve the patients. Data features such as race, gender or income could obfuscate the results. In particular, race was found to give incorrect results for chronic kidney disease (CKD)[3]. Adjusting for this metric, or removing it all together, allowed for better treatment and diagnosis of Black patients; White patients unfairly received better care because of the data and the algorithms in CKD machine learning.

The goal of machine learning, and AI as a whole, is for correct prediction of Y for P(Y|X). For any value(s) of X, the results for Y should be consistent. Once this is achieved, AI can offer better care for the underprivileged. A marked improvement from the current state of healthcare. Having insurance or not also plays a role in the algorithms reading MRI scans[3]. Once these inequities and unfairness of the covariates used in machine learning have been addressed, patients can expect equal, fair and unbiased care. Long-standing, unintentional biases can be greatly reduced by incorporating AI in healthcare practice.

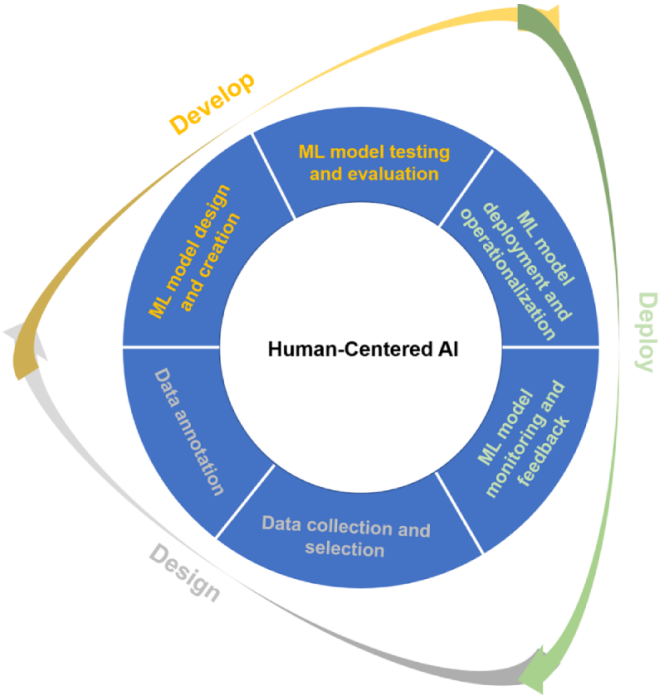


Figure 1: Typical AI lifecycle[2]

The lifestyle figure highlights some of the possible pitfalls in making AI equitable. The “design” portion must be made fair and inclusive of all who make up a population of interest. The “develop” stage must adhere to the same standards. Once the technology has been deployed, equity and fairness can be an expected outcome. This is a huge step forward from the current, unintentional bias that exists in healthcare. By adhering to each stage of AI development: **Data Collection and Selection; Data Annotation; Machine Learning Design, Creation, and Evaluation; Machine Learning Model Deployment, Operationalization, Monitoring, and Maintenance**[2], healthcare AI can yield great benefits.

AI tools can lessen racial bias in pain treatment

According to the U.S. Department of Health and Human Services, Black and Hispanic people receive worse care on 40% of the department’s care quality measures and this is particularly striking in pain treatment.

Research shows how Black patients are systematically under-treated for pain compared to white patients, due to stereotypes, misinformation, and lack of empathy from health care providers. The historical roots of this structural issue may be in the legacy of slavery and medical experimentation, but it extends beyond individual prejudice and reflects the racial disparities and injustices in the US healthcare system[9].

As a glimpse into the purpose of AI development work that practitioners like us undertake, Pierson et al published an article in Nature Digital Medicine[10] showcasing the use of a deep learning model that can help reduce racial disparities in pain management.

The authors developed a deep neural network (DNN) that was trained on over 50,000 knee radiographs from the Osteoarthritis Initiative (OAI) and Multicenter Osteoarthritis Study (MOST) cohorts. The DNN was able to predict patients’ pain scores with an accuracy of 0.67, which is comparable to the inter-rater reliability of radiologists. The DNN also discovered image features that were associated with pain but were not captured by the conventional radiographic scoring system, such as Kellgren-Lawrence grade (KL-grade). These features included bone shape, texture, and density.

The article suggests that the DNN model could help address racial disparities in pain management. They showed that Black patients reported higher pain levels than White patients with the same KL-grade, suggesting that KL-grade alone is not sufficient to capture the complexity of pain experience. However, when the DNN model was used to predict pain, the racial gap was reduced by 47%, indicating that the DNN model could better account for the factors that contribute to pain beyond KL-grade.

The model learned to identify subtle image patterns that were associated with pain but **were not detectable by radiologists or conventional image analysis methods**. The paper demonstrates the potential of ML to improve the quality and equity of pain care, by providing a more objective and personalized measure of pain that can complement the subjective and biased self-reports.

Telemedicines / Telehealth using AI

The US is the biggest economy in the world, and notwithstanding that, it is an example of what is wrong with how we deliver health services. A 2023 paper by Kenneth Antonio Colón[14] lists why we need to look beyond the conventional.

1. High cost of healthcare and increasing by the day
2. The disparity between the rural and urban world
3. The cookie-cutter approach in which there is little to no personalization for individuals.
4. Lastly, the burnout of physicians and the need for scalability. The recent pandemic is a prime example of that.

What is needed is a fundamentally different health system. We need a health system that leverages new payment and care delivery models, coupled with innovative technologies, to truly put patients at the center of their care, thus creating an improved healthcare experience and better patient outcomes at radically lower costs.[14].

Telemedicine/Telehealth distributes health-related services and information via electronic and telecommunication technologies. It allows long-distance patient and clinician contact, care, advice, reminders, education, intervention, monitoring, and remote admissions[13].

We can significantly enhance the effectiveness of Telemedicine by introducing AI-assisted automation using natural language processing (NLP). AI assistants can automate clinician workflows, enabling clinicians to focus their time on patient care, not repetitive tasks, and enabling a deeper level of personalized treatment. Kenneth provides some compelling use cases listed in the table below of use of AI to deliver tailored services, freeing up physicians to focus on patients.

Application	Action
Symptom triaging	• Connect with patients and guide them through a series of questions regarding their histories and current symptoms • Deliver a report of presented symptoms and possible conditions to the patient's physician, which the physician can consult during a follow-up appointment with the patient
Reminders	• Specific use cases, including to take their prescribed medications, reminders about upcoming appointments, etc.
Automated follow-up	• Review status of symptoms presented during a previous consultation with the physician (e.g., if a physician marks that a patient presented symptoms of cough or sore throat and was prescribed an appropriate treatment regimen, a digital assistant automatically follows up to inquire about the success of the regimen [e.g., "how has your cough been since beginning your medication?"])
Assessments and questionnaires	• Automate common assessments, for example, a PHQ-9 for depression to gain information for intake forms and other documents in a more user-friendly manner
24-hour hotlines or emergency services	• Direct patients during crises or times when care teams are unavailable • Instantly notify care teams and/or a patient's dedicated emergency contact as needed
Automate preventive health measures	• If data points deviate significantly from the patient's average (e.g., a decrease in step count, limited movement, and reported symptoms of sluggishness), message the patient in real time, suggesting consultation with his/her care team.

Table 1: AI chatbot applications

Analysis-Drawbacks

Difficulty with Public Trust in AI

Algorithms and proper datasets can advance accuracy, equity and fairness using AI in healthcare. However, the public perception and their trust of AI in healthcare must improve. People do not feel AI can be accountable for the decisions it makes like a human would be accountable[5]. AI is difficult to understand; algorithms with many features are uninterpretable. In fact, transparency with the models would be unfair to those with difficulty understanding mathematical models. Transparency does need to be developed for the public to understand the AI healthcare models. It needs to be described in a lucid manner so that people can understand what the features mean and how they are used.

If the public perceive the algorithms as a "black box"[5], they will not understand the decisions made. When people do not understand how something works and it seems magical, there will be little trust in that thing. With important decisions made on someone's health, people will find comfort in the person making those

decisions. It's the personal touch where people find comfort, especially with a troubling diagnosis. People want to know more information; the interpretability of the model leads to mistrust of the AI that made the decisions. Patients want the accountability of someone deciding and do not feel comfortable with the lack of accountability that an algorithm has, or the dataset chosen to train the model. Moreover, there are concerns that AI could further distrust in the medical field[7]. If medical professionals rely more on AI to make medical decisions, the public will believe less in those decisions. Without addressing the immediate concerns the public has on the explainability of AI, it could further erode the perception of all medical decisions - people will fear AI concluded and not the medical professional relaying their findings.

It has been argued that we need not trust the AI used in healthcare[7]. Instead, we continue to instill our trust in the medical professionals who use AI as a tool, just as we were to trust an MRI machine or scalpel. Trust in medicine must remain with the individuals involved without waiting for AI to develop empathy or interpersonal relationships. Once the public shifts the trust from AI to the people using AI, the technology might be more conducive to acceptance among the public. It might be too much to ask for the public's knowledge of AI and understanding; keeping the trust with healthcare providers allows AI to be used as a tool without the implications of being understood. Many do not understand the complexities of a commonly used item like a smartphone, but it is still used with the belief that it will work. DeCamp and Tilburt caution not to draw a parallel between healthcare AI and everyday technology used[7]. Public's trust in AI differs, especially in healthcare. The professionals using the tools must remain as a conduit for understanding, just like with the tools they use in healthcare today.

Banerjee, Alsop, Jones and Cardinal showed that informing the patients (public) of the AI, in all phases, helped mitigate the mistrust they had with the treatment[6]. They started with a simple linear regression model that could be understood. From there, the complexity was increased with the diligence of updating the patients. They were informed of the data collection used in the study and their data used in the algorithms. Each step of the process was dutifully explained to the patients within the study. Concerns of ethics and fairness could be discussed and addressed. The dialog between the healthcare providers, patients and engineers allowed the AI to become more accurate and where the public understood motives and decisions made along the way. Based on how better the public felt with the involvement of the program, this approach should be explored more with future healthcare AI products. The simplicity of involving patients in the process improved trust.

How AI Models Could Fail to Detect Skin Cancer in Diverse Populations

Availability of advanced medical equipment in developed countries at a greater scale than rest of the world also means that data collection from such equipment is predominantly corresponding to Caucasian population, and it is structurally difficult to acquire representative data from other ancestries.

A study by Haenssle et al[11], which tested a machine-learning algorithm for detecting skin cancer based on images from the International Skin Imaging Collaboration (ISIC) database, illustrates the issue. The algorithm performed better than 58 dermatologists in diagnosing skin lesions, but the database had a limited representation of dark skin types and ethnicities. This could mean that the algorithm might not be as accurate or reliable for people with darker skin tones, who have different patterns and features of skin cancer than lighter-skinned people.

Another study by Adamson et al[12], which analyzed the diversity of skin images in 23 publicly available datasets that could be used for training or testing AI models for skin cancer diagnosis, surfaces the aspect of thoroughly

preparing medical data before feeding it to a model. The study found that only four datasets had information on skin type, and only two datasets had information on race or ethnicity. Most of the images were from fair-skinned individuals, and very few images were from people with brown or black skin. The study also found that the labels and annotations of the images were inconsistent and incomplete across the datasets. These issues could lead to AI models that are biased, inaccurate, or unsafe for diagnosing skin cancer in diverse populations.

Sufficient and quality data collection is the basic premise of a reliable AI solution and any tool built without this foundation is likely to falter. An AI solution is quite fallible in this regard as opposed to a human who relies on experience and subjective observations and can make up for the lack of available training data. Short of sufficient examples, AI models can potentially worsen the existing disparities instead of bridging the gap.

"Work with Me, Don't Just Talk at Me - When "Explainable" Is Not Enough"

The above title is very thought-provoking, and we have taken this from a fascinating article by Brian Pickering[15]. Healthcare is more than just understanding symptoms and explaining treatments. To a large extent, it involves emotion and empathy, which today, AI chatbots are incapable of and often are many layers below the surface, even to be visible.

The diagram below shows the workflow as of today involving the patient and the medical community.

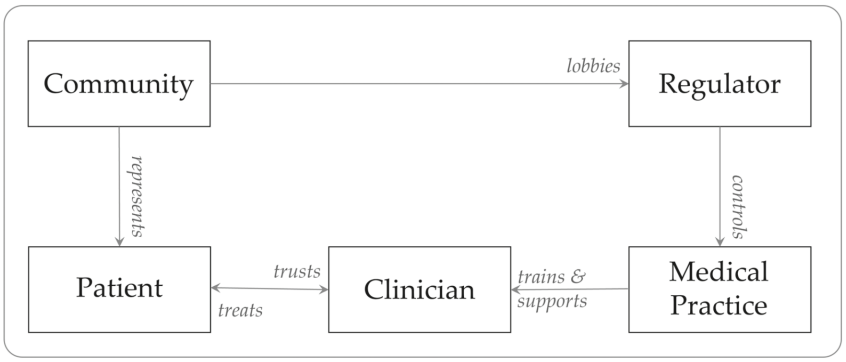


Figure 2: A schematic actor network based on patient clinician interactions for clinical care

Now with AI involved in the process, it would become more complicated and look like this.

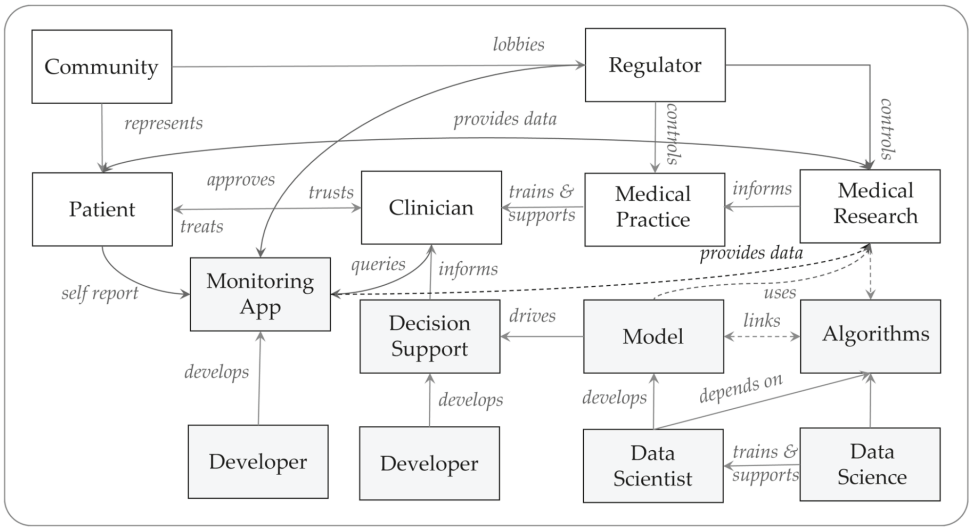


Figure 3: Complete advanced technology-enabled healthcare socio-technical system

While we debate the role of AI and data science in healthcare, its function is layered with a myriad of actors, making it very complicated to control what we build into AI and what the patient experiences.

Case Study 1 – UK[15]

Dataset: 400 UK private citizens were asked about their perceptions of using healthcare apps. The survey was anonymous, and according to the authors, the participants were balanced across relevant demographic categories for the UK.

Responses: Responses to Some Statements about Using Healthcare Apps are as below.

	Agree	Disagree
I don't trust healthcare apps will get it right for me	186	214
I don't feel the health service always has enough time for me	250	150
Using a healthcare app on my own means I'm being fobbed off	84	316
Using healthcare apps means I'm not so alone	203	197

Table 2: Case study one (key) responses (from Case Study 1)

Study Analysis:

- Regarding app accuracy, a substantial minority (186/ 400 46.5%) are suspicious of technology in this setting.
- There is also a feeling that access to healthcare is not always straight.
- forward: 250/ 400 (62.5 %) of respondents felt that the health service did not have enough time for them as individuals.

- The overwhelming majority (316/ 400 or 79 percent) did not think introducing a self-monitoring app indicated that the health service wished to avoid caring for them.
- Finally, when asked whether using an app would help mitigate feelings of isolation, respondents were split almost 50/ 50 (203/ 400 or 50.8 % agree with the statement).

Case Study 2 – USA[16]

Although the results are interesting, the biases in the data are difficult to judge as the dataset is not made public. A similar study in the US by Alison et. al. [16] shows similar results but also shows the biases we have come to see in data. The study shows interesting results, correlating with the earlier study, i.e., patients are more comfortable using AI when it comes to aspects which today are more done by labs, and tend to be less open to situations where they expect humans to be involved.

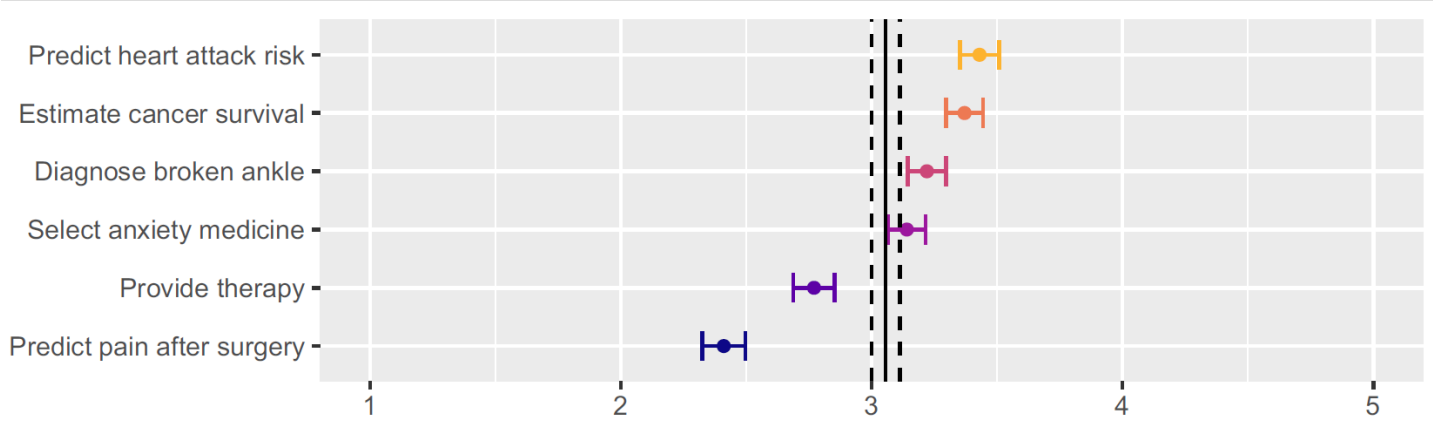
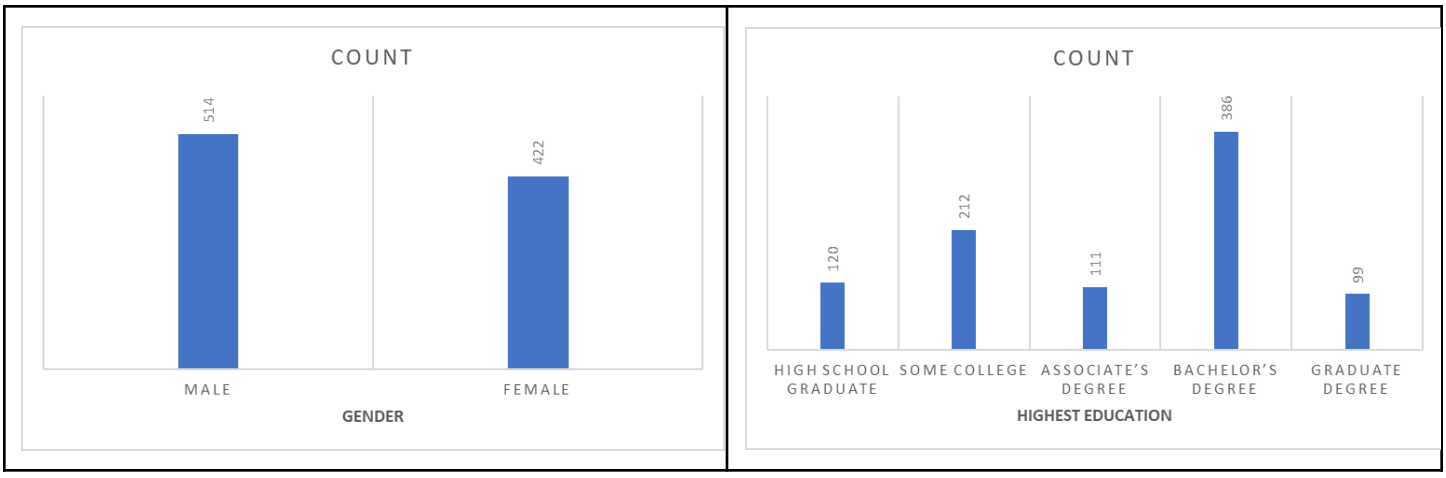


Figure 4: Mean scenario openness scores with 95% CIs. Large vertical lines indicate grand mean with 95% CI. (from Case study 2)

Exploratory Data Analysis: The following histograms show various views of the demographics of the survey data.



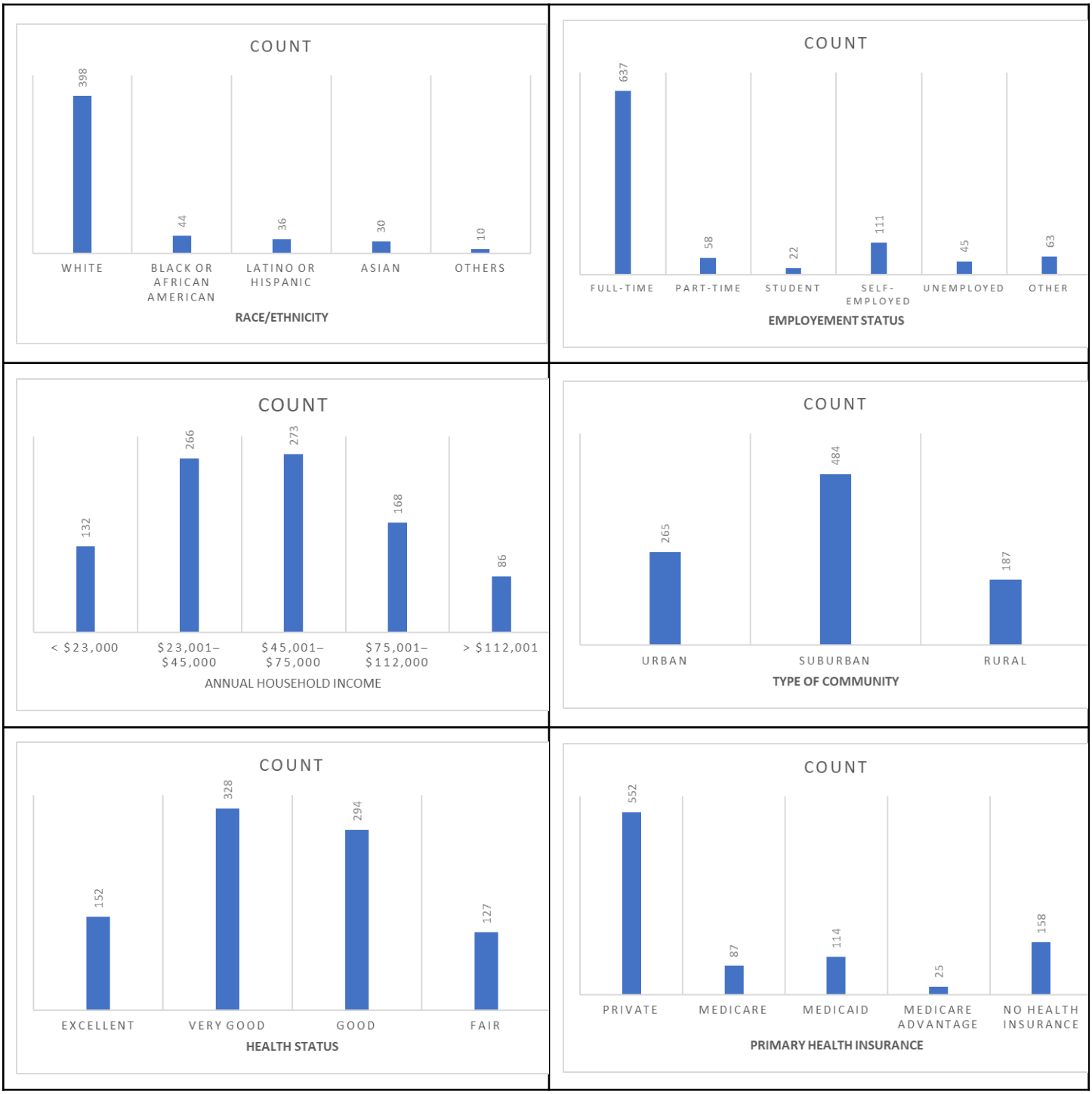


Figure 5: EDA for case study 2

Biases: The data clearly has a bias towards white, male, educated participants who are in good health and have private insurance. This clearly biases the results to a specific demographic. However, to the credit of the authors, they have gone a step ahead and did a correlation analysis of openness, concern and benefit response to demographics.

	Openness		Concern		Benefit	
	r	95% CI	r	95% CI	r	95% CI
Socio-demographics						
Age	−.12	[−.18, −.06]	.06	[.00, .12]	−.03	[−.09, .03]
Sex (1 = Male, 0 = Female)	.10	[.04, .16]	−.20	[−.26, −.14]	−.03	[−.09, .04]
Race (1 = White, 0 = Non-White)[a]	−.05	[−.11, .01]	.01	[−.05, .07]	−.08	[−.14, −.02]
Ethnicity (1 = Latino, 0 = non-Latino)	.06	[.00, .12]	−.09	[−.15, −.03]	−.02	[−.08, .04]
Household income	.07	[.01, .13]	−.08	[−.14, −.02]	.07	[.01, .13]
Community type	.06	[.00, .12]	−.06	[−.12, .00]	.01	[−.05, .07]
Employment status[b]	.17	[.11, .23]	−.18	[−.24, −.12]	.05	[−.01, .11]
Education	.04	[−.02, .10]	.03	[−.03, .09]	.01	[−.05, .07]
Health status and access						
Health status	.08	[.02, .14]	−.12	[−.18, −.06]	−.02	[−.08, .04]
Healthcare location[c]	.03	[−.03, .09]	−.01	[−.07, .05]	.02	[−.04, .08]
Healthcare choice[d]	.08	[.02, .14]	−.06	[−.12, .00]	.11	[.05, 17]
Health insurance (1 = Yes, 0 = No)	.09	[.03, .15]	−.10	[−.16, −.04]	.05	[−.01, .11]
Healthcare satisfaction (n = 735)	.11	[.04, .18]	−.07	[−.14, .00]	.14	[.07, .21]
Psychosocial variables						
Health System Trust Index	.27	[.21, .33]	−.27	[−.33, −.21]	.21	[.15, .27]
Trust in technology	.41	[.36, 46]	−.21	[−.27, −.15]	.41	[.36, .46]
Faith in technology	.38	[.32, .43]	−.10	[−.16, −.04]	.46	[.41, .51]
Conscientiousness	.02	[−.04, .08]	.11	[.05, .17]	.15	[.09, .21]
Agreeableness	.08	[.02, .14]	.11	[.05, .17]	.20	[.14, .26]
Extraversion	.08	[.02, .14]	−.12	[−.18, −.06]	.04	[−.02, .10]
Emotional stability	.08	[.02, .14]	−.06	[−.12, .00]	.07	[.01, .13]
Openness (trait-based)	.07	[.01, .13]	.07	[.01, .13]	.05	[−.01, .11]
Social conservatism	−.01	[−.07, .05]	−.10	[−.16, −.04]	.05	[−.01, .11]
Economic conservatism	−.06	[−.12, .00]	−.06	[−.12, .00]	.02	[−.04, .08]

N = 936 (except as noted for specific variables in Tables 1 and 2)

[a] Participants who selected any race other than White, or in addition to White, were classified as Non-White for purposes of this analysis

[b] 1 = full-time employment, 0 = all other options

[c] 1 = doctor office or private clinic, 0 = all other options

[d] 1 = great or some choice; 0 = little to no choice

Table 3: Correlations of openness, concern, and benefit scores with all study variables (from Case study 2)

This exploratory analysis of variables associated with openness, concern, and perceived benefit indicated that socio-demographic and health variables were largely unrelated. There were modest relationships of age and sex to openness: older participants were less open, and males were more open than females. Females also responded more negatively when presented with concerns. Full-time employment status was associated with greater openness and lower concern. People with greater healthcare choice and healthcare satisfaction perceived more significant benefits, and lower health status was associated with greater concern.

Introducing advanced technologies into clinical practice involves many different challenges. Meeting them all requires a concerted effort by all parties to understand one another and continuously co-create acceptable

innovation. It is essential to understand where technology fits and its introduction's effects on existing healthcare processes and especially healthcare relationships. Clinicians need to be able to understand but also engage with technologists to ensure that AI-generated results are appropriate and do not compromise the clinician's ethical duties toward their patients. Patients and clinicians should engage with data scientists and other technologists to monitor and understand how advanced technologies work to make informed decisions about their use.

Public Perception

Introduction

While we spent significant time evaluating various advantages and disadvantages of AI in healthcare, and discussed the same among us, it was important that we looked beyond this team to see how a large set of people perceived this topic.

Data source

There are various platforms on which healthcare & AI is being actively discussed such as Facebook, Instagram, Twitter etc. However, as of today, most of these platforms disallow data scraping. These platforms "might" have developer API's that allow you to capture data, but access to those seems beyond the time available for this project. We have chosen *reddit* as the source of data for this exercise. To access reddit data, we used "PRAW: The Python Reddit API Wrapper"[8].

We looked at the following subreddits that are active on the topic of AI.

subreddit	**users**
r/artificial	397k
r/artificialinteligence	233k
r/machinelearning	2.7m
r/chatGPT	2.6m
r/openAI	476k
r/GPT3	486k
r/datascience	967k

Table 4: Data source for reddit study

In these subreddits we query for submissions with the following keywords in the title of the submissions [*medicine, healthcare, health, diagnosis, patient, diagnosis, doctor, pharma*].

Potential Biases

1. Biases introduced by our selection of subreddits. We have tried to use multiple subreddits, instead of focusing on just one to alleviate this.
2. Bias introduced by our choice of search query. We have carefully chosen a wide selection of keywords to get submissions from a wide range of audiences.
3. Using reddit itself is a source of bias. It excludes demographics that don't use reddit.
4. People who feel comfortable and know how to use a social media site, like Reddit, might be more inclined to be positive or at least neutral on AI in healthcare.

Data Collection

Data collection process involved the following steps:

1. Create a reddit account.
2. Register as a developer app.
3. Use PRAW APIs to get access to the comments for the discussion.

```
# create a reddit instance
# the specific credential have been removed for security purpose

reddit = praw.Reddit(
    client_id="XXX",
    client_secret="XXX",
    user_agent="my user agent",
    username="XXX",
    password="XXX"
)

reddit.read_only = True
```

```
# this approach retrieves posts from each subreddit

# add subreddits to get posts from
subreddits = ['artificial', 'artificialinteligence', 'machinelearning', 'chatGPT',
'openAI', 'GPT3', "datascience"]
comments = []
topics = 0

# loop through subreddits
for subreddit in subreddits:
```

```
    # get posts for a search query related to healthcare
    # ignoring posts with 'taste' in title as reddit users seem obsessed with posts in which they try to get AI bots a 'taste of their own medicine'
    submissions = reddit.subreddit(subreddit).\
        search(query='(title:medicine OR title:healthcare OR title:health OR title:diagnosis OR title:patient OR title:diagnosis OR title:doctor OR title:pharma NOT title:taste)', \
               sort="hot", limit=None)
    # for each submission, get title, main body, comments and replies
    for submission in submissions:
        comments.append(submission.title)
        comments.append(submission.selftext)

        submission.comments.replace_more(limit=None)
        tmp_comments = []
        comment = submission.comments.list()
        tmp_comments = [c.body for c in comment]
        comments.extend(tmp_comments)
        topics += 1
        print(str(topics) + " ] " + submission.subreddit.display_name + " - " + submission.title + " - " + str(len(tmp_comments)))
    time.sleep(5) # delay is required to prevent error 429

len(comments)
```

```
# join all comments in the list into a single string
all_comments = " ".join(comment for comment in comments)
```

```
# Create and generate a word cloud image:
wordcloud = WordCloud(collocation_threshold=30).generate(all_comments)

# Display the generated image:
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.show()
```

```
# do sentiment analysis

sentiment = TextBlob(all_comments).sentiment
print(f'Polarity: {sentiment.polarity}, Subjectivity: {sentiment.subjectivity}')
```

The final dataset is a set of 981 submissions, across 7 subreddits and a total of 8964 comments with a word count of 438,920. We have not collected any personal unique identifiers for redditors while collecting this data.

Data Analysis

We decided to do a word cloud generation to see what is trending, and also do sentiment analysis on the dataset. For both these we need a *textblob* from the comments, which turns out to be roughly 2.7mb in size (word count of 438,920 and 15,374 sentences).

Figure 6: Reddit word cloud

Word cloud does not throw up surprises, as ***AI***, ***data***, ***people***, ***healthcare***, and ***doctor*** are the most actively used terms. What is interesting to note is that ***will***, and ***think*** are also trending showing that people are looking at AI in healthcare beyond just a passive tool, and the community at large is aware of its implications.

A sentiment analysis on the same dataset yields the following results:

Sentiment Polarity	0.12	**range**: -1 to +1
Sentiment Subjectivity	0.48	**range**: 0 to 1

Table 5: Sentiment Analysis

The general sentiment is better than neutral / slightly positive, but people seem to be biased / subjective in their opinions on this topic.

Key findings

1. The comments are from subreddits related to AI, Machine Learning, GPT and Data Science, with query terms related to healthcare and medicine, which are dedicated to discussing and sharing news, articles, and opinions on advancements in AI in healthcare. Comments have different formats and purposes, ranging from asking questions, making statements, to sharing links to articles or videos.

2. AI fueled personalized medicine is a popular topic among the participants. There is also a debate and discussion on the role and value of AI versus human intelligence in medicine.

3. Ethical and social implications of AI for health are a prominent topic of discussion, as well as concerns about the trustworthiness, fairness, accountability, and transparency of AI in healthcare.

4. There is a high level of awareness and interest in the current state and future direction of AI in healthcare. Many comments are recommendations for courses or resources to learn AI skills specific to healthcare indicating a high demand and interest for education and training in this field.

5. Most comments are positive and optimistic, highlighting the potential benefits and opportunities of AI for improving diagnosis, treatment, research, and innovation in medicine.

6. Some comments are neutral and curious, asking questions or seeking information or advice on AI and healthcare. A comment from a user who claims to have built the closest to an A.I. Bayesian Brain with Human-like logic in healthcare was met with awe and skepticism.

7. Specific applications or examples of AI in healthcare are discussed, indicating the diversity and innovation in this field.

Safeguards

1. We chose a discussion from a community of 7.8 million members. We also chose keywords which are trending in various discussions (about 900 comments).
2. PRAW API allows us to get the reddit username of people posting the comments. We have ensured that we do not have these details in the data we have collected using this API.
3. To be in line with the terms & conditions of reddit developer apps API, we are not uploading the downloaded comments as part of this submission.

Recommendations

Better communication and understanding surrounding AI need to be delivered to the public. We have found numerous examples where the public do not trust AI in healthcare, mostly for its accountability. Proposing a framework where data scientists, clinicians and patients all working together can improve perceptions. Positive and negative aspects of AI in healthcare have been explored. If people can trust AI more, the benefits will outweigh the detractors of incorporating the latest AI technology in healthcare. These improvements include:

1. Lessened bias in datasets. All data used to train the models must be deemed equitable and fair. This includes considerations of gender, race, age, socioeconomic status, et al. If group fairness takes priority over individual fairness, some individuals or underrepresented minority groups may not experience the same treatment. With healthcare, these oversights can be terminal.
2. Reduced bias in algorithms. Like with data, algorithms can unfairly discriminate against individuals - or if trained on individuals, unfairly discriminate against groups. It cannot be assumed that with unbiased and fair data, the algorithms will maintain that fairness. The slightest decisions in grouping, weighting or even random values can alter the results. The altered results can further be exacerbated and continue until the results become unfair and biased.
3. True governance over AI in healthcare. Committees on AI can better trust the public's opinion. These committees should consist of healthcare professionals, law professionals, non-profit consultants, academics, et al. - not just data scientists. Varying angles of expertise advising and reviewing the AI can ensure unbiasedness and fairness through the life of the technology.
4. Complete transparency. Though AI can be hard to understand for those outside of the technology field, it does not exclude it from being explained. The articles cited in this paper consistently detail the mistrust in healthcare AI because of the lack of interpretability of the models. People do not trust what they do not understand. Being forthcoming with the features, how the algorithm works, and sharing the data (anonymously encoded) would allow individuals to understand how decisions are made. This might even encourage the public to be more inquisitive, setting a standard for all AI models. Transparency creates understanding and trust.
5. Standards developed to regulate AI. Like data have regulations such as HIPAA, COPPA, GDPR, AI needs the same kind of governance. The EU AI Act is a great start, but the regulation needs to apply to AI globally. The U.S. Algorithmic Accountability Act of 2022 also promotes transparency and accountability to AI technologies, which encompass healthcare. Most importantly, transparency and traceability will help the public understand the underpinnings of AI, and hopefully, increase trust.
6. Disclosing results generated from AI. By explaining to the public which of their diagnoses have been given by AI, people can question and investigate the results. They can challenge the fairness of the data on which the algorithms were trained, and the possible bias built into the algorithms.

References

[1] Greater public trust in AI would help NHS detect cancers early, experts say | E&T Magazine

[2] Human-Centered Design to Address Biases in Artificial Intelligence

[3] Algorithmic fairness in artificial intelligence for medicine and healthcare | Nature Biomedical Engineering

[4] The Pros and Cons of Healthcare Chatbots

[5] For Patients to Trust Medical AI, They Need to Understand It

[6] Patient and public involvement to build trust in artificial intelligence: A framework, tools, and case studies - ScienceDirect

[7] Why we cannot trust artificial intelligence in medicine - The Lancet Digital Health

[8] PRAW 7.7.2.dev0 documentation

[9] Black Americans are Systematically Under-Treated for Pain. Why?

[10] An algorithmic approach to reducing unexplained pain disparities in underserved populations | Nature Medicine

[11] AI Skin Cancer Diagnoses Risk Being Less Accurate For Dark Skin - Slashdot

[12] Characteristics of publicly available skin cancer image datasets: a systematic review

[13] Telehealth - Wikipedia

[14] Creating a Patient-Centered, Global, Decentralized Health System

[15] Future Health Scenarios AI and Digital Technologies in Global Healthcare Systems – CRC Press

[16] Exploring perceptions of healthcare technologies enabled by Exploring perceptions of healthcare technologies enabled by artificial intelligence: An online, scenario-based survey artificial intelligence: An online, scenario-based survey