**AI Chatbot for Management of Depressed African Americans: A Pilot Study**

**Proposal Summary (300 words)**

We propose a research project that enhances an existing evidence-based artificial intelligence tool to address needs of depressed African Americans. In alignment with the North Stars, the project trains an African American Doctoral/Masters student in AI, expanding the available workforce and building the community’s capacity to address AI. The project increases awareness about AI by a national-scale social media campaign to the alumni, and members, of Alpha Phi Alpha fraternity, the first fraternity established for African American men. The existing AI tool recommends antidepressants for 16,775 general-population patient subgroups, each representing a unique combination of medical history. For each of these subgroups, the project will analyze the appropriateness of the recommendations for African Americans, using the All of Us database and the published literature. All of Us is a collection of ~400,000 individuals who have agreed to share their EHR data. African Americans are over-sampled. The proposed project has developed novel methods for analysis of data within All of US. We plan to use LASSO regressions to identify variables that predict African American’s response to antidepressants within these data. We then plan to use factorial combination of non-zero variables in these regressions to generate the needed prompts and completions. To reduce hallucination/fabrication in the AI model, for completion we plan to use text found in the original source document. We do not plan to probabilistically generate this text. The recommendation is a single word (name of antidepressant with highest remission rate). This word is wrapped in previously prepared fixed (not generated) text.

# Specific Aims

This pilot is responding to NS III of the AIM-AHEAD program, by studying artificial intelligence (AI)-based technology applied to mental health problems. Depression is a growing health problem, with the proportion of US adults with depression increasing, especially for African Americans. Antidepressant medications are a first-line treatment for depression; however, a majority of depressed patients do not experience improvement with their first antidepressant. The current consensus guidelines for prescribing antidepressants have been criticized as (i) not adequately responding to patients’ comorbidity and history of illness, (ii) relying on outdated assumptions of mechanism of depression, and (iii) lacking empirical evidence that following the guidelines provides better care. Additionally, adherence to these guidelines is low. Minority populations, including African-Americans, are not well represented in antidepressant studies, contributing to reduced antidepressant effectiveness in these populations. Identifying the most appropriate antidepressant can be challenging, due to the large number of depression treatment options and detailed demographic and medical history information needed to identify the best choice. There is a significant need to synthesize available evidence regarding antidepressant effectiveness and provide personalized treatment recommendations.

Recent advancements in machine learning (ML) and natural language processing (NLP) through large language models (LLM), like ChatGPT, are changing how people seek information, including health guidance. In this new paradigm, users engage in natural language conversations with a dialogue system. We propose developing and evaluating an LLM-based dialogue system for engaging patients in conversations to identify antidepressant recommendations. The dialogue system will engage patients in back-and-forth communication, to solicit needed patient information, make medication recommendations, and provide a detailed explanation for recommendations. The dialogue system will integrate the expanding body of antidepressant research into a single tool and allow patients to directly access this clinical knowledge through natural language. We hypothesize this natural language dialogue system will offer benefits over structured, survey-based medication selection tools, in terms of the user experience and data collection completeness.

**Aim 1: Perform a comprehensive retrospective of antidepressant effectiveness among African American patients.** We will identify and aggregate available information regarding antidepressant effectiveness among African Americans, including published literature and Food and Drug Administration (FDA) applications, to create the Consolidated Antidepressant Guidelines (CAG). CAG will include the findings from approximately 552 antidepressant studies. In addition, we will summarize the data from All of Us on the response of African Americans to antidepressants.

**Aim 2: Develop an AI dialog system for individualized, evidence-based, African American-centric depression management.** We will develop the Knowledge-enhanced Antidepressant Recommendation Dialogue System (KARDS) that leverages state-of-the-art LLM to guide patients through a semi-structured interview to identify the best treatment options. KARDS will utilize knowledge grounding approaches incorporate the medical knowledge of CAG from Aim 1, to provide personalized, explainable, evidence-supported antidepressant recommendations. KARDS will use an empathetic tone and allow users to describe themselves and their experiences in their own words. It will utilize a dialogue agent to guide the conversation flow, restrict conversation topics, and identify safety concerns to protect patients. The code and models will be made available, and the system will be publicly hosted through the George Mason University web site.

**Aim 3: Perform a user study with African American patients to evaluate system functionality and user preferences.** We will prospectively recruit 30 dyads consisting of clinicians and African American patients with depression to evaluate the functionality and user experience of KARDS. As a baseline, we will create a structured, survey-based tool to guide patients in identifying the CAG-supported recommendation(s). Patients will use both the survey-based tool and KARDS to identify CAG-supported recommendations, to understand modality preferences (structured vs. natural language). Clinicians and patients will provide feedback regarding dialogue system acceptability. We will compare the CAG-supported treatment with consensus guidelines, to understand protocol changes.

**Aim 4: Increase awareness and develop capacity in African American community**. In this pilot, we plan to develop a diverse, equitable, and inclusive AI/ML workforce by recruiting a doctoral/master student research assistant from minority communities. Once trained, this individual will contribute to AI/ML community capacity building.

# Significance

This project is significant because it focusses on antidepressant selection, a challenging problem in which clinicians and patients are facing difficulties. Majority of depressed patients (60%) do not benefit from their first prescribed antidepressant [[[1]](#endnote-2)]. Prescribing antidepressants is difficult, due to the quantity of antidepressants, volume of available research, represented populations, publication bias, and other factors. There are many possible choices (> 20 antidepressants on the market), and clinicians can combine these medications with other medications. There are more than 500 studies of antidepressant effectiveness, each relying on a small patient sample [[[2]](#endnote-3)], and most failing to provide subgroup analysis in African Americans [[[3]](#endnote-4)]. Negative studies are not published and are not available to clinicians, despite FDA reporting, resulting in a skewed understanding of drug efficacy [[[4]](#endnote-5)]. Genetic profiling has not helped anticipate treatment response [[[5]](#endnote-6)], and few studies have clarified how patients’ comorbidities affect antidepressant response [[[6]](#endnote-7)]. Published data often focuses on negligible differences among antidepressants [[[7]](#endnote-8)], even though large differences exist in patient subgroups [[[8]](#endnote-9)]. AI Guided depression management can address these shortcomings

This project is significant because it addresses the gap in management of depressed African Americans, this is an underserved community-defined priority. This population has notably poorer response to antidepressants than any other group [[[9]](#endnote-10)]. Only 37.1% of African Americans with any mental health diagnosis are in treatment. African Americans have been historically excluded from studies of effectiveness of antidepressants [[[10]](#endnote-11)]. As such, it is plausible that current guidelines, especially AI algorithms, may be perpetuating a pattern of poor care for African Americans. In general, three issues have been raised regarding AI-guided­ disease management: (i) AI systems can “insidiously magnify” current care that has failed to adjust to the needs of marginalized groups [[[11]](#endnote-12)]; (ii) digital interventions are not tested on patients with the highest unmet health care needs and therefore may fail in pragmatic, real world use [[[12]](#endnote-13)]; and (iii) there may be stronger social bias among minorities for reporting mental health problems, and AI systems may have an incomplete view of depression problems confronting African Americans [[[13]](#endnote-14)]. This project addresses the needs of African Americans through meta-analysis of findings reported in the literature and through separate analysis of data available on African Americans in All of Us database [[[14]](#endnote-15)]

Patient-facing dialogue systems hold transformative potential for the healthcare sector and are increasingly prominent in psychiatric applications, predominantly through *therapy-bot* implementations [[[15]](#endnote-16), [[16]](#endnote-17)]. Likewise, we are proposing a patient-facing application. The system engages patients in detailed, time-consuming, medical history intake and provides point-of-care summary and prescription recommendations to the patients’ clinicians. The proposed system is acceptable to busy clinicians and not disruptive to clinic processes. Patients are also comfortable with it because the natural language modality provides an intuitive, empathetic, stigma-free interface [15, 16]. We are not aware of prior research focused on developing and evaluating an antidepressant recommendation system. The proposed antidepressant recommender dialogue system will address this gap and explore open research questions related to dialogue system effectiveness and acceptability in this clinical application. The proposed dialogue system will also explore approaches for jointly incorporating the natural language understanding and generation capabilities of LLM with clinical knowledge in the form of formula and tabular results.

# Approach and Timeline

**Aim 1: Analyze effectiveness of antidepressants among African American patients.**

We will gather, and consolidate, existing data on the efficacy of antidepressants using well established procedures for meta-analysis [[[17]](#endnote-18)], emphasizing the African American population. This comprehensive effort involves collecting information from more than 500 published papers and 20 FDA applications. The published literature address minority treatment as secondary subgroup analysis although 7 studies address it as primary analysis [[[18]](#endnote-19)]. The studies are typically small in scope, in that they are limited to comparison of one or two medications in small sample data [e.g., [[19]](#endnote-20)]. These studies are typically not pragmatic with the exception of 6 studies. Typically, the response to treatment is discussed in broad terms, as if minorities are homogenous. They pay little attention to the unique patterns in medical history of minorities [[[20]](#endnote-21)]. A recent review concluded that response to treatment among minorities “remains largely unknown” [[[21]](#endnote-22)]. To remedy the shortcomings of published literature, we plan to use existing data from 6 studies conducted by Patient-Centered Outcome Research Institute (PCORI) and the All of Us database. The All of Us database over samples African Americans and therefore is ideal to supplement published literature.

We will follow a published protocol for evaluating antidepressants post-FDA release using EHR data [[[22]](#endnote-23)]. Because EHR data are observational, we control for confounding through stratification. The primary outcome of interest in this analysis is the patient-reported remission, which is not available in claims or EHR data. A surrogate measure is needed. One could rely on 100-day continuation (no augmentation or switching) at recommended dosage. We have shown that this surrogate measure accurately measures patient-reported remission (c-statistic = 0.93) [[[23]](#endnote-24)].

This proposal develops a novel method for addressing health disparities in AI systems developed for general population but used by African Americans. An AI system is organized around prompts and completions. Prompts constitute patients’ relevant medical history. Completions indicate the recommendation. We define an AI system as racially biased if it recommends non-optimal medications to African Americans but optimal ones to the general population. Our 2021 study of effectiveness of 15 antidepressants in the US population had created 16,775 subgroups of patients, each with unique medical history [22]. Our study provided data for optimal prescriptions in each of the 16,775 subgroups. For these subgroups, using the All of Us data, we will report if the current algorithm recommends optimal treatment for African Americans. This method of analysis examines AI recommendations in covariate balanced, stratified, homogenous subgroups. If a recommendation is not optimal for African Americans, the advice is changed to correspond to their experiences within All of Us data or in published data. To make sure that a recommendation is available for all subgroups, first factors that affect response of African Americans to antidepressants are developed in All of Us Data through LASSO regressions. Within the general population subgroups, factorial combination of these variables is used to create prompts; and the optimal antidepressant constitutes the completion. These prompts and completions are used to correct bias in the general population recommendations. This method of analysis is preferred to a random prospective clinical trial because it is more pragmatic and takes into account the comorbidities of patients. The corrected Consolidated Antidepressant Guideline (CAG) will reflect both the meta-analysis of findings relevant to African Americans in the published literature and the experiences of African Americans in publicly available databases, such as All of Us (primary) and PCORI data (secondary).

**Aim 2: Develop a dialogue system for individualized, evidence-based antidepressants recommendations.**

A picture containing text, screenshot, font

Description automatically generated

Figure 1. Dialogue system architecture

We will develop the Knowledge-enhanced Antidepressant Recommendation Dialogue System (KARDS), which will engage users in a back-and-forth conversation (multi-turn dialogue) to acquire the patient information needed to identify appropriate antidepressant medication. Figure 1 presents the KARDS architecture, which consists of three primary, interconnected components: dialogue agent, LLM, and CAG. The dialogue agent controls the patient interview, including the conversation flow, tracks the dialogue state, and uses the LLM to interpret and generate text. The dialogue agent queries the CAG from Aim 1 to determine the needed patient information and identify specific study results relevant to the patient. While it may be possible to provide sufficient prompt-based instructions to an LLM to conduct the patient interview without the use of an external dialogue agent, we elect to use a deterministic, rule-based dialogue agent in this pilot study to enable greater control of the patient interview and improve repeatability.

In KARDS, all LLM inputs and outputs flow through the dialogue agent. The LLM serves two primary functions: 1) interpret the natural language patient inputs, including extracting semantic representations of patient provided information and 2) generate natural language responses for the patient, including questions, based on the needed patient information identified by the dialogue agent. The LLM will generate text using an empathetic tone, to improve patient comfort. For the LLM, we will use a pre-trained language model that is instruction-tuned, so it can follow natural language instructions from the dialogue agent. Candidate LLM include proprietary models, like ChatGPT [[[24]](#endnote-25)], and publicly available models, like medAlpaca [[[25]](#endnote-26)].

Table 1 presents an example patient interview. The dialogue state consists of a structured representation of the acquired *patient profile*. At the start of the conversation, the patient profile is empty (no patient information), and all antidepressant studies are deemed relevant to the patient. As the conversation progresses, additional information about the patient is acquired (*Incremental patient profile* in Table 1), which grows the patient profile and reduces the number of relevant studies. During the patient interview, the dialogue agent directs the LLM to inquire about specific formation and instructs the LLM to output a structured representation of the patient provided information that can be incorporated into the growing patient profile. The dialogue agent will follow a high-level interview structure based on typical medical practice (e.g., demographics, previous antidepressants, current medications, etc.) and will identify specific questions for the patient based on the acquired patient profile and CAG. The conversation concludes once the patient profile is sufficient to make antidepressant recommendations. KARDS will provide a summary of the collected information to the patient. After the patient information is confirmed, KARDS will output a description of the recommended medications, list of the relevant studies, and an explanation for the medication decisions. The KARDS output will automatically send the patient’s clinician a brief point-of-care recommendation and explanation, with option to examine a complete record of the conversation and the supporting evidence.

**Table 1. Example Patient Interview**

|  |  |  |
| --- | --- | --- |
| **Patient Interview** | **Incremental**  **Patient Profile** | **Relevant Studies** |
| KARDS: Have you been diagnosed as having major depression?  PATIENT: I thought I had bipolar but my doctor said no. | {EligiBLE: yes} | s1, s2, s3, … sm |
| KARDS: What is your age and biological sex?  PATIENT: I am 32 and female | {Age: 18 to 40,  Sex: female} | s1, s2, s3, … sm |
| KARDS: Which antidepressant medications have you taken in the last 12 months?  PATIENT: I took citalopram but it did not work, then I took Bupropion and that did not work either. Now I am not sure. | {Last AD: Bupropion – no response,  # AD Hx: 2} | s1, s2, s3, … sm |
| … | … | … |
| KARDS: We found 258 patients with similar backgrounds. Among them, fluoxetine had the best outcome with 38% of patients responding to it. We shared this information with your clinician and encourage you to discuss the next step with your clinician. It’s important that you do not alter your medication without consulting your clinician, as sudden changes in antidepressants can have severe consequences, including increased risk of suicide. Would you like more information about our recommendation and next steps?  PATIENT: I have an appointment next week but I like to know more about the side effects of fluoxetine before then. | {HEALTH ed: more on fluoxetine,  Alert: sent to clinician,  NEXT: pending visit} | s2 |

**Ethical, legal, and privacy considerations:** The development of the AI system for the general population was approved by the George Mason University’s Institutional Review Board and we expect that they will also approve the modifications to the system to adjust for African Americans’ needs. To improve safety, we will only enroll patients who have a mental or primary clinician. The dialogue agent determines if the patient is in mental health crisis (e.g., indication of self-harm or misplaced anger). If a critical interaction is identified, the patient’s clinician will be automatically notified (as per signed consent form) and the patient is referred to a local phone crisis line. To reduce risk of accidental release of information, patient identifiers are encrypted and kept under passwords.

LLMs have been reported to fabricate wrong inferences. Several steps will be taken to reduce this. First, recommendations are based on text in the original source document/data. The recommendation is not generated probabilistically so there is no room for fabrication. Second, the recommendation is a single word, name of one in 15 antidepressants. This recommendation will be wrapped in one of 15 fixed (not generated) text. Third, the dialogue agent, will monitor the conversation flow by asking the LLM if the client is on topic. If there is a “drift in the topic”, the LLM is instructed to bring the conversation back to allowed medical history topics.

|  |
| --- |
| This pilot is planned for a proposal to NIMH |

**Aim 3: Impact of dialogue on depressed African Americans and their clinicians’ reaction to point-of-care alerts.**

We will recruit 30 depressed African Americans online, and through them, their clinicians. Our efforts to recruit patients online through Google and Facebook have been successful. In a preliminary study, we advertised on Google to 29,636 individuals. In one week, 426 individuals clicked through the advertisement to reach our survey-based decision aid. The average cost of the advertisement was $3.84 per person expressing an interest in the AI guideline. The daily rate of use was 15 individuals per day. On Facebook, we advertised to 50,501 individuals. In one week, 1,110 individuals clicked through the advertisement. The average advertisement cost was $1.74 per individual expressing an interest in the AI-guideline. The daily rate of use of the decision aid was 39 individuals per day. Patients will be asked to complete both a natural language KARDS interview and a structured closed-ended medical history survey (revised to include findings from Aim 1). The current version of closed ended survey is available at <https://hi.gmu.edu/ad/intro>. The order of taking these two surveys will be set randomly to reduce learning effects. Across these two methods of data collection, we will examine (1) ease of use, (2) time burden, (3) comprehensiveness of the collected information, (4) and the patient’s intent to discuss their medication with their clinician.

While the patient is engaged in a lengthy dialogue with the AI system, the clinician experiences the AI as a brief point-of- care alert. We will contact the participating clinicians and present the KARDS’ recommendation and the accompanying explanation. Our experience shows that we can effectively recruit the clinicians to provide us with an hour of their time in exchange for $250 honorarium. In a preliminary unpublished study, Goldberg and Alemiconducted IRB-approved interviews with 12 physicians and 7 mental health nurse practitioners using these procedures. We will get clinician reactions, regarding the quality of the antidepressant recommendations, sufficiency of the supporting information, and acceptability of the dialogue-based approach. We will ask the clinician about intent to change the patient’s medication. This pilot is limited to intent-to-treat because we expect that 32% of clinicians will be AI guideline concordant [[[26]](#endnote-27)] and thus not likely to prescribe any change. This pilot is not powered to detect actual changes in prescriptions or subsequent patient outcomes.

**Aim 4: The study will increase awareness of AI and develop capacity in African American community**.

The study plans to recruit a Ph.D./Masters student and train the person in use of AI systems for management of depression in African Americans. When the student graduates, he/she will add to the available capacity for AI in minority communities. The attorney general of Virginia has informed educational institutions not to follow race-specific recruitments [[[27]](#endnote-28)]. To comply and succeed in attracting a qualified candidate, the position is open to all races; but we advertise heavily to African American students. To encourage African Americans to apply for the position, we will use “Red Balloon” searches within social media. The first Red Balloon search was sponsored by the Defense Advanced Research Projects Agency (DARPA) of the U.S. Department of Defense. In this search, social media was used to identify the coordinates for 10 red balloons, suspended at fixed locations across the country. Surprisingly, the search was able to identify the location of balloons within a day. We plan to contact members, and alumni, of the Alpha Phi Alpha fraternity, the first intercollegiate Greek-letter fraternity established for African American men. Members and Alumni will be identified through LinkedIn® but we have already contacted some, including Walter J Smiley Jr., 1st Vice-Polemarch of Alexandria-Fairfax (VA) Alumni Chapter of Kappa Alpha Psi. Alumni and members will be asked to refer us to their contacts; and likewise, the contacts will be asked to refer us further, creating a chain of referrals. The project will pay $50/person to the last three individuals in the referral chain that ends with an African American applicant to the university. We expect this method of recruitment to meet the requirements of Virginia Attorney General and increase the likelihood of receiving competitive applications from African Americans.

**Project Timeline**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tasks** | **Q1** | **Q2** | **Q3** | **Q4** |
| IRB |  |  |  |  |
| Recruitment of student |  |  |  |  |
| Aim 1 – Literature & Data Analysis |  |  |  |  |
| Aim 2 – Dialogue System |  |  |  |  |
| Aim 3 – User study |  |  |  |  |
| Public distribution of AI system & Code |  |  |  |  |
| Publication of findings in Peer Reviewed Journals |  |  |  |  |

# Investigative Team

Farrokh Alemi, PhD, is a Professor of Health Informatics at George Mason University Department of Health Administration and Policy. He has worked in both academia and the health industry. He maintains patents on (1) sentiment analysis, (2) measurement of episodes of illness and (3) personalized medicine. He has more than 125 peer reviewed publications and is the author of 3 books, including a widely used book on statistical methods for big data in healthcare. His research focuses on causal analysis of massive data available in electronic health records. This research has required balancing of data to remove confounding prior to estimating causal impact of interventions. His publications have contributed to predictive medicine, precision medicine, comparative effectiveness of medications, sentiment analysis, NLP, risk adjusted statistical analysis, causal networked models, identifying trajectories of diseases, and predicting prognosis of patients with multiple morbidities. Dr. Alemi maintains a decision aid for selection of antidepressants at http://MeAgainMeds.com. Alemi is the author of multi-morbidity index, used in management of polypharmacy patients. He has worked with diverse groups of patients including children, nursing home residents and patients with diabetes, major depression, heart failure, anemia, hypertension, trauma, drug abuse, and other diseases. In addition, Alemi was a pioneer in online management of patients and has provided Congressional testimony on role of Internet in health delivery.

Janusz Wojtusiak, PhD, is a Professor of Health Informatics at George Mason University Department of Health Administration and Policy. He is also the Director of the Machine Learning and Inference Laboratory. He has expertise that spans ML, AI, and health informatics in clinical decision support and knowledge discovery in medical data, and a wide range of applications of these fields in health care. His particular area of interest is in developing algorithms that derive simple, transparent and usable models from complex health data to predict patient and population outcomes. He studies how to create and evaluate reproducible, unbiased and trustworthy algorithms and models. He has authored, or co-authored, over 100 research publications and continues to collaborate with multiple national and international institutes. He has numerous joint publications with Dr. Alemi, the PI of this project. These joint publications include a paper on development of a decision aid for management of major depression, the focus of this pilot.

Kevin Lybarger, PhD, is an Assistant Professor in the Department of Information Sciences and Technology at George Mason University. His research interests combine data-driven ML and NLP in the context of important real-world problems. His research explores the intersection of NLP and clinical informatics, including clinical documentation and clinician-patient communications. His experience includes the use of LLM to interpret, analyze, and transform clinician and patient-generated natural language. He has extensive expertise extracting structured semantic representations of patient social, behavioral, and medical history information from free-text clinical documentation in the EHR, to enable real-time and large-scale use of this text-encoded information. Dr. Lybarger’s research explores natural language conversations, including the automatic summarization of clinician-patient dialogue from office visits and the identification of distorted patterns of thinking in therapist-patient text message conversations. Dr. Lybarger and Dr. Alemi are currently working on methods of restricting hallucinations in LLM, through restricting generated text to the text found in the original source and through statistically generating comprehensive set of prompts and completions.

All three members of the team have worked collaboratively, in the past 9 months, to bring AI and LLM to (1) diagnosis of COVID at home from presenting symptoms and (2) selection of antidepressants based on patient’s medical history. The current project is a natural outgrowth of their collaboration together.

**Budget**

See attached

# References

1. Pies R. Are antidepressants effective in the acute and long-term treatment of depression? Sic et Non. Innov Clin Neurosci. 2012 May;9(5-6):31-40. PMID: 22808447; PMCID: PMC3398684. [↑](#endnote-ref-2)
2. Penn E, Tracy DK. The drugs don't work? Therapeutic Advances in Psychopharmacology. 2012 Oct;2(5):179-88. [↑](#endnote-ref-3)
3. Walsh BT, Seidman SN, Sysko R, Gould M. Review Placebo response in studies of major depression: variable, substantial, and growing. JAMA. 2002 Apr 10; 287(14): 1840-7. [↑](#endnote-ref-4)
4. Turner EH, Matthews AM, et al. Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med. 2008;358:252–60. [↑](#endnote-ref-5)
5. Uher R, et al. Genetic predictors of response to antidepressants in the GENDEP project. Pharmacogenomics J. 2009 Aug;9(4):225-33. [↑](#endnote-ref-6)
6. Penn E, Tracy DK. The drugs don't work? antidepressants and the current and future pharmacological management of depression. Therapeutic Advances in Psychopharmacology. 2012 Oct;2(5):179-88. [↑](#endnote-ref-7)
7. Kirsch I., Deacon B., et al. Initial severity and antidepressant benefits: A meta-analysis of data submitted to the food and drug administration. PLoS Med 2008; 5: 260–268. [↑](#endnote-ref-8)
8. Alemi F, Min H, Yousefi M, Becker LK, Hane CA, Nori VS, Wojtusiak J. Effectiveness of common antidepressants: a post market release study. EClinicalMedicine. 2021 Oct 25;41:101171. [↑](#endnote-ref-9)
9. Murphy E, Hou L, et al. Race, genetic ancestry and response to antidepressant treatment for major depression. Neuropsychopharmacology. 2013 Dec;38(13):2598-606. [↑](#endnote-ref-10)
10. Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? Philos Ethics Humanit Med. 2008 May 27;3:14. [↑](#endnote-ref-11)
11. Torous J, Myrick K, Aguilera A. The need for a new generation of digital mental health tools to support more accessible, effective and equitable care. World Psychiatry. 2023 Feb;22(1):1-2. [↑](#endnote-ref-12)
12. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. NPJ Digit Med. 2019 May 13;2:38. [↑](#endnote-ref-13)
13. Mitchell S, Potash E, Barocas S, D'Amour A, Lum K. Algorithmic fairness: Choices, assumptions, and definitions. Annual Review of Statistics and Its Application. 2021 Mar 7;8:141-63. [↑](#endnote-ref-14)
14. Ramirez AH, Gebo KA, Harris PA. Progress with the All of Us Research Program: Opening Access for Researchers. JAMA. 2021 Jun 22;325(24):2441-2442. doi: 10.1001/jama.2021.7702. PMID: 34115137. [↑](#endnote-ref-15)
15. Budler LC, Gosak L, Stiglic G. Review of artificial intelligence‐based question‐answering systems in healthcare. Data Min Knowl Discov . 2023:e1487. [↑](#endnote-ref-16)
16. Viduani A, Cosenza V, Araújo RM, Kieling C. Chatbots in the Field of Mental Health: Challenges and Opportunities. Digital Mental Health. 2023 Jan 1:133-48. [↑](#endnote-ref-17)
17. Gosmann NP, Costa MA, et al. A systematic review and network meta-analysis. Psychol Med. 2023 Jun 6:1-10. [↑](#endnote-ref-18)
18. Fuentes D, Aranda MP. Depression interventions among racial and ethnic minority older adults. Am J Geriatr Psychiatry. 2012 Nov;20(11):915-31. [↑](#endnote-ref-19)
19. Varner RV, Ruiz P, Small DR. Black and white patients response to antidepressant treatment for major depression. Psychiatr Q. 1998 Summer;69(2):117-25. [↑](#endnote-ref-20)
20. Maj M, Stein DJ, et al. The clinical characterization of the adult patient with depression aimed at personalization of management. World Psychiatry. 2020 Oct;19(3):269-293. [↑](#endnote-ref-21)
21. Fuentes D, Aranda MP. Depression interventions among racial and ethnic minority older adults: a systematic review across 20 years. Am J Geriatr Psychiatry. 2012 Nov;20(11):915-31. [↑](#endnote-ref-22)
22. Alemi F, Min H, et al. Procedure for Organizing a Post-FDA-approval Evaluation of Antidepressants. Cureus. 2022 Oct 3;14(10):e29884. [↑](#endnote-ref-23)
23. Alemi F, Aljuaid M, et al. A surrogate measure for patient reported symptom remission in administrative data. BMC Psychiatry. 2021 Mar 4;21(1):121. [↑](#endnote-ref-24)
24. OpenAI. GPT-4 Technical Report. Preprint. 2023. doi: 10.48550/arXiv.2303.08774. [↑](#endnote-ref-25)
25. Han T, Adams LC, et al. MedAlpaca--An Open-Source Collection of Medical Conversational AI Models and Training Data. Preprint. 2023. [↑](#endnote-ref-26)
26. Kessler RC, Berglund P, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA. 2003 Jun 18;289(23):3095-105. [↑](#endnote-ref-27)
27. See <https://www.oag.state.va.us/media-center/news-releases/2548-march-9th-2022-attorney-general-miyares-demands-fairfax-county-middle-school-stop-racially-discriminating-against-children> accessed on 6/8/2023. [↑](#endnote-ref-28)