**Proposal to Train ChatGPT to Give Advice on Antidepressant for African Americans**

Version of March 22, 2023

For More Information Contact falemi@gmu.edu

FUTURE OF HEALTHCARE: Within the next five years, patients with behavioral health problems will first consult an AI system for advice on optimal treatment; and, then go to clinicians for check on the advice and prescriptions.  Our optimism about this timeline is based on integration of our existing data mining research and capabilities of ChatGPT. Based on analysis of more than 10 million patients' experiences with 15 antidepressants, our research has identified optimal antidepressants for 16,700 patient subgroups.  ChatGPT contributes to this research in two ways. First, with its vast knowledge of more than 1,000 published articles on the efficacy of antidepressants, it provides a broad context and depth that goes beyond our narrowly defined data mining models. Second, ChatGPT's conversational model provides a practical method to gather patients’ medical history and other information needed by the AI system. Much like a doctor visit, the ChatGPT enhancements will ask a series of questions, in English, about missing information in patient’s initial description of their medical history. This combination of data mining and ChatGPT trumpets the coming of AI and demonstrates fundamental ways in which direct-to-consumers, online, information services can change health care.

ADVANCE IN HEALTH EQUITY: While many technical issues have been resolved, one important concern remains. Both data mining and ChatGPT summarize current practices. Only 37.1% of African Americans with any mental health diagnosis are in treatment. Current practices serve predominantly White population and may not reflect their response to treatment. African Americans have also been historically excluded from studies of effectiveness of antidepressants. As such, it is plausible that current guidelines, especially AI guidelines, may be perpetuating a pattern of poor care for this population. Critics points out that AI systems “insidiously magnify” current care that has failed to adjust to the needs of marginalized groups. **This project evaluates the impact of AI-guided selection of antidepressants for African Americans** and addresses this health equity concern.

NEED FOR AI: Prescribing antidepressants is difficult because (a) there are more than 20 antidepressants in the market, (b) there are thousands of studies on the effectiveness of antidepressants, each relying on a small non-minority sample of patients, (c) negative studies are not published, despite being reported to the Food and Drug Administration, (d) genetic profiling has not helped anticipate patient response to antidepressants, (e) few studies have clarified how patients’ comorbidities affect their response to antidepressants, and (f) published data reports negligible differences among antidepressants, despite large differences in subgroups of patients. According to our data, AI-guided care can facilitate recovery of 17.5% more patients from depression than prescription of the average clinician.

OUR PLAN: This project examines the effectiveness of the proposed AI system through two Aims. In Aim 1, we plan to use the "All of Us" database, organized by the National Institutes of Health (NIH). This database includes Electronic Health Records of 412,000+ volunteers—but perhaps more important to this study—the database over-samples minorities. In this pragmatic, retrospective data, we will examine the effectiveness of 15 common antidepressants among African Americans and compare the recommendations of the Texas consensus panel and our AI system. Since observational studies can be biased, we plan to use a matched case-control study with inverse propensity weights. Thus, we make sure that effectiveness of antidepressants are compared in the same, or similar, subgroups of patients. We hypothesize that neither consensus, nor AI guidelines, will be optimal for African Americans, and that adjustments need to be made to the AI system to reflect the experiences of African Americans.

SAMPLE SIZE FOR AIM 1: There are no well-established procedures for sample size determination in weighted regression studies. The usual practice is to include more than 10 times the number of variables in the model. The number of variables in each of the 15 LASSO regressions is expected to be less than 30. Thus, data on 300 patients are needed. As of February 2023, 412,000 individuals have contributed their data to All of Us, of which 74,160 (18%) are African Americans. Among these African Americans, the 15 most common antidepressants were used frequently enough to provide sufficient samples: 9,121 used sertraline, the most common, and 667 used Nortriptyline, the least common antidepressant.

In Aim 2, we plan a randomized clinical trial among African Americans recruited from 20 participating clinics across the nation and through online advertisement. Since many depressed African Americans are not in treatment, it is important to supplement clinic recruitment with online advertisement. Patients will be randomly assigned to either ChatGPT enhanced AI guidelines or Texas guidelines. The study will follow-up with patients in 100 days to see how the advice was used and whether the patient experienced symptom remission, as measured by the Hamilton index. These data will be analyzed using covariate-balanced ANOVA. We will distinguish between patients receiving antidepressants as-assigned by randomization and as-treated by the clinician. In situations where the clinicians have not prescribed our recommended medication, the calculation of counterfactuals is done in two steps. First, a network model is constructed using a chain of time-constrained regressions. In the second step, the arc between antidepressant treatment and remission of symptoms is removed from the network; and a counterfactual is calculated from remaining components of the network.

SAMPLE SIZE FOR AIM 2: For comparison of patients as-assigned, a one-tailed Z test of proportions will be performed. A sample size of 200 patients is needed for a proportion difference of 0.05 to 0.10, to be captured with a power of 80% to 95%. We expect that 32% of clinicians will be guideline-concordant; thus, for as-treated comparisons, we need 629 patients. Our preliminary data indicates that through advertisement we can have 39 depressed patients per day visit our online AI site; thus, in 3 years, we have access to a sufficient number of cases.

AFRICAN AMERICAN PERSPECTIVE: The project has organized a national, paid, study board, composed of leading African American or minority psychiatrists and scientists. This board will have a say on all aspects of the project, including project aims, budget, and personnel. The board will also be tasked to examine the project’s compliance with direction provided by the Partnership on Artificial Intelligence Incident Collection Efforts, the USA Digital Services Act, and the European Union Artificial Intelligence Act.

FUNDING HISTORY: The data mining component of this study was funded previously by the Robert Wood Johnson Foundation and the State of Virginia. A pilot project to combine data mining with ChatGPT was funded internally by George Mason's College of Public Health.

DISTRIBUTION OF FINDINGS: All products from this research, including the AI system, will be in the public domain and available freely on the web.