Title: Facilitating Community-Informed Opioid Prescription Guidelines

Team: Raja Safiullah (rsafiull), Manzoor Mirza (mmmirza), Charet Bolton (crbolton), Yi Zhou (yizhou6), Yixi Qiu (yixiq)

Section I: Background
The Opioid Crisis was declared in 2017 and its economic burden is estimated at $78.5 billion a year in the U.S.[1] Since its widespread acknowledgment, the U.S. government has taken a number of measures to prevent opioid-related abuse and overdose. The Centers for Disease Control's (CDC) latest measure to support these efforts is updating the CDC *Clinical Practice Guideline for Prescribing Opioids* from 2016. The 2022 draft of the guideline was revised to include things like new evidence-based practices and greater shared decision-making for more individualized treatment. After concerns for patient harm precipitated the 2016 *Clinical Guideline for Prescribing Opioids*, the CDC recognizes the paramounce of social and community considerations for safe prescribing and provision of opioids and has opened up public comments for the draft of the 2022 guideline.[2] Public comments on the proposed draft are accepted from February 4, 2022 to April 11, 2022.

These comments will illuminate the stories, concerns, and experiences from the patient, provider, and community side to delineate the impact of the change in guidelines. Our team's analysis on this data will prove invaluable to decision makers within the CDC. Through our insights, policymakers can better understand the values and preferences of both the patient and the provider regarding opioid use and pain management. Our analysis can also highlight issues of inequity or inaccessibility that the 2016 and/or drafted guideline elicits. The comments may also identify positive aspects of the 2016 guideline that patients and practitioners find effective. The collection of these insights will support policy makers and decision makers in adjusting the guideline to better reflect the needs and priorities of those impacted. These insights can promote guideline design that improves accessibility, supplements or modifies pain management solutions that patients found distressing or ineffective, and adjusts or clarifies protocols to avoid the misapplication of the guideline and its impact.

Section II: Policy Questions of Interest
Our policy questions will reveal the priorities and preferences of those impacted by the clinical guideline. The answers to the following questions will allow the policymakers responsible for the clinical guideline to supplement and modify the drafted guidelines to make them more effective for all stakeholders.

- What concerns do people have regarding the drafted *Clinical Practice Guidelines*? Do these differ by patient, practitioner, or organization entity types?

[1] Florence CS, Zhou C, Luo F, Xu L. The Economic Burden of Prescription Opioid Overdose, Abuse, and Dependence in the United States, 2013. *Med Care*. 2016;54(10):901-906. doi:10.1097/MLR.0000000000000625.

[2] Centers for Disease Control and Prevention. Proposed 2022 CDC Clinical Practice Guideline for Prescribing Opioids. *Regulations.gov*. 2022. https://www.regulations.gov/document/CDC-2022-0024-0001

- What is the general sentiment towards the guidelines? Negative or positive? What kind of stakeholders are inclined to be negative or positive?

Section III: Dataset Description

We pulled our data from the Regulations.gov API using Python code.[3] At the time of extraction, we had 3,620 comments. As comments are being collected until April 11th, more comments will be uploaded following the submission of this proposal. We expect to have approximately 4,300 comments by submission close. All comments are public submissions. These comments contain a range of stakeholders such as pharmacy technicians, patients and their families, physicians, organizations combating opioid addiction, and pharmaceutical companies and associations. Our dataset currently has the unique identification of the comment as one attribute, the number of duplicates for the comment, and the text of the comment itself. The comments are unstructured as they are plain text and do not have a uniform format or structure.

Section IV: Methods (needs help, feel free to make any and all edits you like)

We intend to perform our analysis on comments in three parts. First, we would perform parsing and cleaning of data through tokenization and removal of stop words (partially identified through TF-IDF). Using this corpus, we would perform text analysis and visualizations to identify the most frequently occurring words and entities using Spacy's named entity recognition on the dataset. This would be instrumental in understanding the most common ideas across documents. We will do this analysis in addition to topic modeling to identify possible key ideas and themes across comments. In addition, we also intend to perform co-occurrence analysis to identify the common types of words that most frequently co-occur for patients, providers, and individuals which could help us understand the values, concerns, or priorities which are most likely to coexist across each entity type.

Second, we'll apply clustering methods like GMM and LDA to explore whether there appear certain topics people share concerns with. This helps to provide a whole picture of stakeholders' attitudes and concerns for policymakers to improve this guideline design. The biggest obstacle to this part will be the unsupervised nature of analysis per se, which makes it difficult to determine the "successful" criteria. But we'll address it by checking its visualization presentation. Finally, we'll use sentiment analysis to inform policymakers of whether individuals are mainly for or against the drafted guideline to determine whether major guideline revision may be needed.

Section V: Expected Findings

We are performing our analysis with the hypothesis that we will get a varying degree of perspectives due to our large sample of comments and that opinions will vary between patients, providers, and organizations. This will allow us to gain insights on what is currently working with the 2016 *Clinical Practice Guideline for Prescribing Opioids* and areas of improvement or concern. Secondly, through our unsupervised classification of comments, we hope to see sentiments being clustered through similar words and phrases that help uncover the patterns of concerns, ideas, or feedback regarding the guidelines.

---

[3] Regulations.gov API. Retrieved on April 4, 2022.
https://api.regulations.gov/v4/comments?filter[commentOnId]=0900006484f67412 (requires API Key)

Section VI: Limitations and Considerations

We recognize some limits to our use of public comments. The data skews towards patients sharing their experience with opiate prescriptions, so the data may be unbalanced when considering other submitter types such as organizations. Also, people commented on both the impact of the 2016 guideline and their opinion on the 2022 drafted guideline, so we will have to associate which guideline they are expressing their opinion on- if they are not just sharing their experience. We still deem this data invaluable due to its ability to capture multiple sentiments across various stakeholders. Our data cleaning and analysis will also address the possibility of duplicate comments (such as those created by advocacy groups) to ensure balanced representation.