Exploring the Digital Landscape of Antidepressant Use: Demographic Insights, Emotional Analysis, and Topic Modeling

Jianfeng Zhu, Xinyu Zhang, Ruoming Jin, Hailong Jiang, Deric R Kenne

Submitted to: Journal of Medical Internet Research
on: May 28, 2024

Table of Contents

Original Manuscript..... 5

Exploring the Digital Landscape of Antidepressant Use: Demographic Insights, Emotional Analysis, and Topic Modeling

Jianfeng Zhu[1] PhD; Xinyu Zhang[2] PhD; Ruoming Jin[1] Prof Dr; Hailong Jiang[1] PhD; Deric R Kenne[3] Prof Dr

[1]Department of Computer Science, Kent State University Kent US
[2]Department of Communication, Michigan State University East Lansing US
[3]Center for Public Policy and Health, College of Public Health, Kent State University Kent US

Corresponding Author:
Jianfeng Zhu PhD
Department of Computer Science, Kent State University
800 E Summit St
Kent
US

Abstract

Background: Antidepressant medications are widely used to manage depression and anxiety disorders. Given that these medications have been around for many years, their side effects are well documented. Apart from common side effects, there are new emerging data reporting unexpected side effects or stress related challenges that clinicians should be aware of.

Objective: The goal of this study to analyze the demographic profiles, symptomatology, and platform-specific differences in online discussions about antidepressant side effects. Exploring the antidepressant side effects in different gender and age, then the clinical with detailed information to improve the documentation and management of drug side effects.

Methods: We collected dataset from Ask a patient and Reddit platform. A mixed methods approach including quantitative analysis and qualitative analysis. Which we leverage Natural Language Processing method to examine the demographic and symptoms comments, and exploring the language linguistic patterns on word, topic, and theme of posts.

Results: The analysis highlighted significant gender-based differences in antidepressant side effects. Females reported more gastrointestinal, neurological, sleep, general physical, and emotional symptoms, while males reported more sexual health issues. Users frequently discussed withdrawal symptoms and challenges with switching medications, revealing anxiety about the withdrawal process. Unique symptoms such as toe infections, visual changes, and pulsatile tinnitus were also noted. The strong correlation between fear and neurological, gastrointestinal, and sleep-related symptoms suggests these physical symptoms significantly contribute to emotional distress.

Conclusions: These findings underscore the necessity for gender-sensitive healthcare approaches and highlight new concerns that are not typically covered in standard drug descriptions. Addressing these issues directly in consultations can improve patient outcomes by fostering better adherence to medication regimens and enhancing overall treatment satisfaction. Future research should focus on understanding the unique challenges faced by specific demographics, particularly women in reproductive age who use birth control, to provide more effective and personalized care.

(JMIR Preprints 28/05/2024:62680)
DOI: https://doi.org/10.2196/preprints.62680

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Exploring the Digital Landscape of Antidepressant Use: Demographic Insights, Emotional Analysis, and Topic Modeling

Abstract

Background: Antidepressant medications are widely used to manage depression and anxiety disorders. Given that these medications have been around for many years, their side effects are well documented. Apart from common side effects, there are new emerging data reporting unexpected side effects or stress related challenges that clinicians should be aware of.
Objective: The goal of this study to analyze the demographic profiles, symptomatology, and platform-specific differences in online discussions about antidepressant side effects. Exploring the antidepressant side effects in different gender and age, then the clinical with detailed information to improve the documentation and management of drug side effects.
Methods: We collected dataset from Ask a patient and Reddit platform. A mixed methods approach including quantitative analysis and qualitative analysis. Which we leverage Natural Language Processing method to examine the demographic and symptoms comments, and exploring the language linguistic patterns on word, topic, and theme of posts.
Results: The analysis highlighted significant gender-based differences in antidepressant side effects. Females reported more gastrointestinal, neurological, sleep, general physical, and emotional symptoms, while males reported more sexual health issues. Users frequently discussed withdrawal symptoms and challenges with switching medications, revealing anxiety about the withdrawal process. Unique symptoms such as toe infections, visual changes, and pulsatile tinnitus were also noted. The strong correlation between fear and neurological, gastrointestinal, and sleep-related symptoms suggests these physical symptoms significantly contribute to emotional distress.
Conclusions: These findings underscore the necessity for gender-sensitive healthcare approaches and highlight new concerns that are not typically covered in standard drug descriptions. Addressing these issues directly in consultations can improve patient outcomes by fostering better adherence to medication regimens and enhancing overall treatment satisfaction. Future research should focus on understanding the unique challenges faced by specific demographics, particularly women in reproductive age who use birth control, to provide more effective and personalized care.

Keywords: Antidepressant, AskaPatient, Reddit, Natural Language Processing, BERTopic

Introduction

Antidepressant use in the United States has become increasingly prevalent over the past few decades. According to the National Center for Health Statistics, data from 2015-2018 indicate that 13.2% of adults aged 18 and over reported using antidepressant medication in the past 30 days [1]. This increase reflects ongoing trends in the diagnosis and treatment of depression and other mental health conditions. Factors such as heightened public awareness, reduced stigma surrounding mental health, and advancements in pharmacological treatments have contributed to this upward trend [2]. Understanding patient experiences with antidepressants is crucial, given their significant role in contemporary mental healthcare.

Patient-centered care emphasizes understanding and addressing individual patient needs and preferences. Directly measuring attitudes toward medications from social media reviews can help detect early factors contributing to non-adherence and negative outcomes [3]. This is especially important in psychiatry, where medications are often prescribed by trial-and-error or based on side effect profiles [4]. Research indicates non-adherence is more strongly linked to patient beliefs than to personal attributes or clinical variables. [5]. Online health communities offer patients a platform to share experiences and concerns about treatments, providing valuable insights into their attitudes. Studies show that self-reports of drug effects on social media can improve patient-provider communication and treatment management [5,6]. Analyzing this data allows healthcare providers to understand patient experiences better, inform personalized treatment plans, and address common concerns proactively [4]. Thus, leveraging social media for patient feedback is a vital tool for enhancing patient-centered care and improving adherence outcomes.

Nevertheless, we observe a gap in the comparative analysis of patient-reported experiences with antidepressants across different digital platforms. While significant research has utilized social media data to understand the psychopathological effects of psychiatric medications and monitor adverse drug reactions, there is a lack of comprehensive studies examining the alignment between patient-reported side effects and official drug descriptions. Furthermore, existing research often focuses on isolated platforms such as Reddit without systematically comparing them to more structured forums like AskaPatient [7,8]. This gap highlights the need for a nuanced analysis that considers the demographic and clinical profiles of individuals discussing antidepressant effects, as well as the differences in content between various online forums. Addressing these gaps can enhance our understanding of patient experiences, improve post marketing surveillance, and contribute to the field of precision psychiatry by tailoring treatments based on detailed, real-world data.

Our work seeks to address these gaps and complements existing methodologies for understanding the effects of psychiatric medications. We report a comprehensive study leveraging data from two distinct online platforms: AskaPatient and Reddit, to analyze patient-reported experiences with antidepressants. Specifically, we collected data on the top 10 antidepressants from AskaPatient, encompassing 1,571 reviews from 2015 to 2024, including detailed side effects, comments, and demographic information. Additionally, we filtered Reddit posts from 2018 to 2023 using drug names and demographic patterns, amassing 34,628 posts. Our feature extraction process focused on demographic and linguistic elements, employing Python's Regular Expressions for data parsing and DistilRoBERTa for emotional tone classification. We utilized BERTopic for topic modeling to identify latent discussion themes and KeyBERT alongside sentence transformers for semantic similarity analysis of side effects. By comparing these insights across platforms, our study aims to provide a more nuanced understanding of the real-world impacts of antidepressants, bridging the gap

between patient-reported side effects and official drug descriptions, and enhancing the field of precision psychiatry through detailed, real-world data analysis.

Literature review

Demographic Differences in Antidepressant Use

Research on antidepressant use reveals significant demographic differences, encompassing socioeconomic status, sex, and age. These disparities are critical for understanding the nuanced patterns of antidepressant prescription and efficacy across diverse populations. Firstly, socioeconomic status has been shown to influence antidepressant use. Studies consistently demonstrate that individuals from lower socioeconomic backgrounds exhibit higher rates of antidepressant use, a trend not solely attributable to higher incidences of depression within these groups. For instance, research conducted in Australia revealed that socially disadvantaged individuals reported greater antidepressant use and primary care consultations, suggesting that the universal healthcare system may mitigate healthcare inequalities [9]. Similarly, a longitudinal study in Norway found that low socioeconomic status, indicated by factors such as income and education, was associated with higher antidepressant prescription rates, partly due to increased mental health service utilization among these groups [10].

In addition to socioeconomic factors, sex differences also play a significant role in antidepressant use. Research has revealed significant variations in both the efficacy and pharmacokinetics of these medications between males and females. Studies indicate that females generally exhibit a better response to serotonergic antidepressants compared to males, although postmenopausal women tend to have a diminished response relative to younger females [11]. This variance is attributed to physiological differences such as hormone levels, body fat distribution, and liver metabolism, which affect drug absorption and efficacy [12].

Age differences further complicate the landscape of antidepressant use and response. A mega-analysis of individuals with major depressive disorder (MDD) revealed that patients aged 21 years and younger, as well as those older than 55 years, exhibited slower and less robust responses to pharmacotherapy compared to those aged 21 to 35 years [13]. Additionally, data from the National Health and Nutrition Examination Survey (NHANES) indicated that antidepressant use increases with age, from 3.4% among individuals aged 12-19 to 19.1% among those aged 60 and over [1]. Further, a systematic review and meta-analysis focusing on elderly patients with MDD found that approximately 50.7% of elderly patients responded to antidepressant treatment, a rate comparable to that of younger adults [14]. These findings underscore the necessity for age-specific considerations in the management of depression.

Self-reported Antidepressant Use on social media.

The emergence of social media platforms has provided researchers with a unique opportunity to study self-reported antidepressant use and its associated effects through naturalistic expressions of individuals. Social media has also facilitated the identification of adverse drug reactions at the population level suing self-reports as well as the mentions of side effects of adverse drug reaction [6]. Several studies have leveraged large-scale social media data, particularly from Twitter, to examine the psychopathological effects of psychiatric medications, including antidepressants [15] [16]. These studies have employed machine learning techniques to assess psychopathological states known to be affected by antidepressants, such as mood, cognition, depression, anxiety, and suicidal ideation [15]. By analyzing the language used in social media posts before and after self-reported antidepressant intake, researchers have been able to identify linguistic markers associated with

mental health changes[17].

Social media is also valuable for post marketing drug surveillance [6]. Studies using Twitter data have shown that social media can reveal insights into the side effects and efficacy of antidepressants through linguistic markers and behavioral changes. For example, a study using unsupervised language analysis on Twitter identified keywords related to major side effects like sleep disturbances, weight changes, eating disorders, pain, and sexual dysfunction [17]. Further, analyzing Twitter posts from users under selective serotonin reuptake inhibitor (SSRI) treatment revealed significant behavioral and linguistic changes. These include increased nighttime activity, altered writing styles, and heightened use of first-person singular pronouns, which can indicate depressive symptoms and treatment effects [6]. This study emphasizes the role of social media in monitoring and supporting mental health patients by providing real-time data on their experiences with antidepressants. Additionally, research on Instagram posts has shown that public sentiments towards antidepressant use are predominantly negative. Many users express dissatisfaction due to adverse effects and lack of improvement [7]. A longitudinal study analyzing Twitter data from 2019 to 2020 categorized tweets into medical and non-medical content. It found that discussions about antidepressants mainly focus on health-related areas such as sleep and appetite/weight issues [18]. These findings collectively illustrate the power of social media as a rich data source for understanding self-reported antidepressant use, offering valuable insights into the side effects, efficacy, and public perceptions of these medications.

Moreover, social media data can provide complementary insights into the effects of psychiatric drugs, which is particularly valuable in precision psychiatry. By analyzing pre-treatment signals of mental health states, researchers can potentially predict individual drug success. This approach moves away from the traditional trial-and-error method of medication prescription. It aligns with the principles of digital psychiatry, which uses digital data to improve mental health services and tailor treatments to individual needs [6]. However, some studies have found limitations in using patient comments on social media to identify side effects. These studies often fail to accurately compare online reports of antidepressant effects with the official drug descriptions [17]. Our research aims to fill this gap by examining the alignment between drug descriptions and patient-reported side effects.

Several platforms provide real-time data on experiences with antidepressants. Researchers have utilized professional online health care forums like AskaPatient to identify factors associated with attitudes toward antidepressants [5]. Only one study has focused on social media platforms such as Reddit [8]. Despite the growing body of research utilizing online platforms to understand patient experiences with antidepressants, there remains a notable gap in comparative analyses across different types of digital forums. Specifically, while platforms like AskaPatient offer structured and detailed user reports, broader social media sites such as Reddit encompass a more diverse and conversational range of user interactions. This research seeks to fill this gap by systematically comparing content related to the effects of antidepressants between AskaPatient and Reddit. By doing so, it aims to elucidate differences in user experiences and attitudes across these platforms. The following research questions were proposed:

Research questions:

1) What is the demographic profile of individuals who post antidepression drug symptomatology and effect side online?

2) What are the central antidepression effect side demographic and clinical symptoms as well as depressive symptomatology between depressed patients in different age groups?

3) What differences content in course of antidepression effects between

AskaPatient and Reddit platform?
4) Does online antidepression effect side align to drug description?

Methods

Dataset and Preprocessing

Our study collected data from two online flatform. The first is AskaPatient [19], which allows patients to rate their medicine or view ratings and comments from others. This database includes prescription drugs approved by the FDA, as well as many over-the-counter medicines and shingles, flu and Covid-19 vaccines. We gathered data on the top 10 antidepressants, with 1571 drug reviewers posted on AskaPatient forum from 2015 to 2024. This dataset includes side effects, comments, and demographics information, such as age and gender (refer to Appendix 1, Table 1).
Additionally, we used drug names to filter through Reddit posts from 2018 to 2023, matching them with age and gender patters to extract posts with specified gender and age details. We manually reviewed thousands of posts with specified gender and age details. After numerous experiments to achieve the higher accuracy, we identified specific language patterns (e.g., (17F), [23M], f/19, (F,26), and 23(F)) to filter posts with gender and age information. Using python's Regular expression(RE) operations [20], which define a set of strings matching specific patterns, we successfully collected 34,628 posts using this filtering method. We randomly selected 1000 posts to verify accuracy by our team members, achieving a cross-valuate value of 0.9.

Feature extraction and Transformation

User Profile Features extraction

We extracted demographic data, specifically age and gender, from user profiles to understand demographic influences on antidepressant discussions. This involved scripting in Python using Regular Expression to parse and collect data. Descriptive statistics were used to illustrate the central tendencies and dispersions of demographic variables among forum participants.

Linguistic Features extraction

To study linguistic features [21], we examined the importance of words that users choose to express their feelings in their personal writings.
Emotional Features: The Emotion English DistilRoBERTa-base model [22], a lightweight version of BERT optimized for faster performance, was used to detect and classify emotional tones in text data. The model was trained on six diverse datasets (see Appendix) and predicts Ekman's six basic emotions, plus a neutral class: anger, disgust, fear, joy, neutral, sadness and surprise. The model is trained on a balanced subset from the datasets listed above (2,811 observations per emotion, i.e., nearly 20k observations in total). 80% of this balanced subset is used for training and 20% for evaluation. The evaluation accuracy is 66% (vs. the random-chance baseline of 1/7 = 14%) [23].
Semantic Features: The topic modeling methods such LDA, NMF and BERTopic, [24,25] were used to cluster posts related to similar topics. We used BERTopic [26], an algorithm leveraging transformer models to create more coherent and interpretable topics, to discover latent topics in the dataset. BERTopic is supported by four key models: embedding, uniform manifold approximation and projection, hierarchical density-based spatial clustering of applications with noise, and CountVectorizer. Since default parameters may not offer a one-size-fits-all solution, BERTopic allows for the customization of these key models' parameters to better suit
Specific needs. The two-difference online platform discuss different topics related to the antidepressants, which helps in broadly understand the real adverse effects of drugs.
Domain Specific Features: In this study, we focused on understanding whether the side effects of

antidepressants discussed online align with those described in reputable medical resource. To achieve this, we first extracted the most common side effects (details in Tabel1) from the reputable sources such as WebMD[27] and the national health service (NHS)[28]. These sources are widely recognized for providing credible and accurate health information.

The types of antidepressants commonly prescribed by doctors include selective serotonin reuptake inhibitors (SSRIs), serotonin and norepinephrine reuptake inhibitors (SNRIs), tricyclic antidepressants, atypical antidepressants, and monoamine oxidase inhibitors (MAOIs). We ensured that these categories aligned with the adverse reactions listed for the top drugs from DAILYMED [29]. This alignment forms the basis of our investigation into whether the side effects reported in online forums reflect those mentioned in medical descriptions.

Tabel1 Refined side effects features

Side effects Features	Symptoms
Gastrointestinal symptoms	Nausea, vomiting, Diarrhea, constipation, indigestion, and stomach aches
Neurological Symptoms	Headache, dizziness, Tremors
Sleep Disorders	Insomnia, hypersomnia, sleepy
Sexual Health Issues	Sexual dysfunction, decreased libido, erectile dysfunction, Difficulties achieving orgasm
General Physical Symptoms	Dry mouth, Fatigue, weight gain, Increased sweating
Emotional and Psychological Symptoms	Feeling agitated, shaky, anxious, anxiety, depressed, Loss of appetite

To thoroughly understand the alignment between online discussions of antidepressant side effects and medically documented descriptions, we applied both keyword extraction and semantic similarity analysis techniques.

We applied KeyBERT [30] to generate keywords and keyphrases from users' posts that are most similar to the symptoms of refined side effects features. KeyBERT is a versatile keyword extraction technique that leverages BERT embeddings to identify relevant keywords and phrases from the text.

Next, we utilized the sentence-transformers/all-MiniLM-L6-v2 model [31] to analyze semantic similarities between users' posts and predefined categories of symptoms in Tabel1. This model uses advanced sentence embeddings to capture the contextual meaning of sentences. For each document, we calculated cosine similarity scores against predefined categories such as "Gastrointestinal Symptoms" and "Neurological Symptoms", each associated with specific medical terms. The embeddings for the side effects and their respective category phrases were generated and compared to determine the maximum similarity score. Categories were labeled based on a threshold similarity score of 0.7, with higher scores indicating a significant semantic match between the side effect and the category, facilitating a structured analysis of the side effects data.

By leveraging both keyword extraction and semantic similarity analysis, we were able to achieve a robust assessment of the side effects described in user posts. This analysis helped us observe differences and similarities in the reported symptoms compared to documented side effects.

Correlational Analysis of Feature Interactions

After extracting and synthesizing features, we visualized the relationships among them using heatmaps. This visualization technique is essential for identifying and correlations within the data, providing a clear depiction of how different features interact with one another.

To further investigate the impact of the extracted features on gender classification, we employed SHAP (SHapley Additive exPlanations)[32] values. SHAP values offer a unified measure of feature importance, indicating the contribution of each feature were most influential in predicting gender based on user discussions about antidepressants. SHAP values are particularly useful because they

provide both global and local interpretability of the model. Globally, they show how each feature impacts the overall prediction, while locally, they explain individual predictions by indicating how much each feature contributes to a specific outcome.

Ethical Considerations

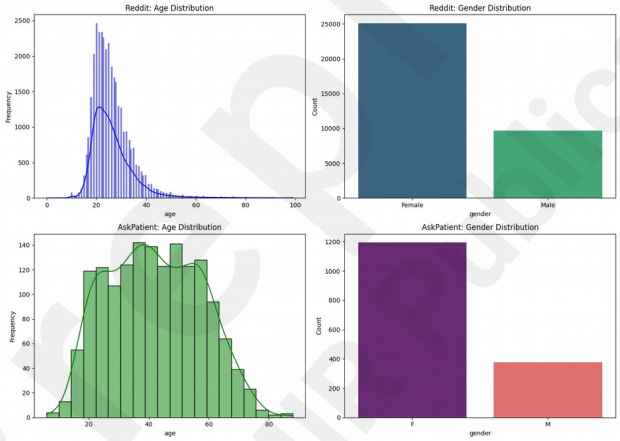
The authors confirm that the research presented in this article met the ethical guidelines, including adherence to the legal requirements, of the United States of America and received approval from the Institutional Review Board (IRB) of Kent State University.

For the dataset obtained from AskaPatient, we signed a data use agreement with the website, ensuring that we complied with all necessary ethical standard and legal obligations. The Reddit data used in this study were sourced from publicly available platforms [33] and, do not contain any identifiable information. Importantly, the data set used in this study is entirely devoid of personal information, such as author names or any other private identifiers.

Results

Feature extraction and Transformation

Figure1 shows the descriptive statistic for age and gender within the two datasets. The Reddit dataset, consisting of 34,628 entries, reveals a mean age of 26.44 years, a standard deviation is 8.62 years, and median age of 25 years.



Firge1 Demographic Distribution of Age and Gender on Reddit and AskaPatient

For the AskaPatient dataset, the total number of observations for the age variable is 1571. The mean age is 42 years, with a standard deviation of 15 years. This measure indicates the amount of variation or dispersion from the mean.

The female distribution is approximately three times that of males within two datasets, as shown in Figure1.

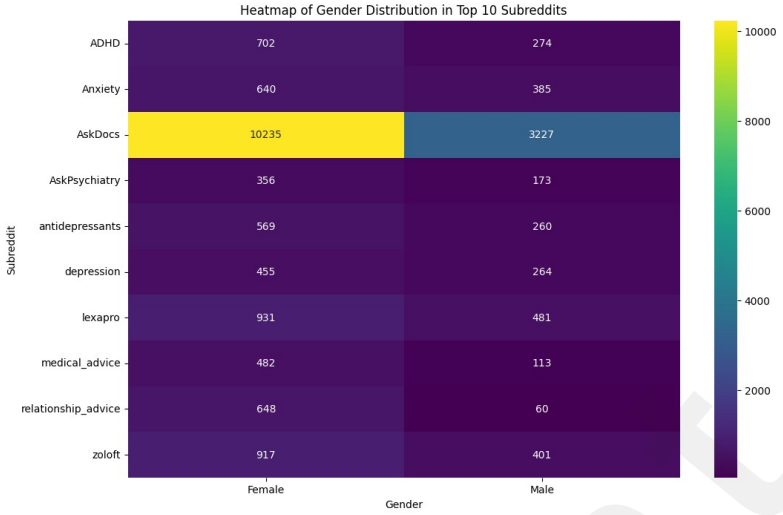


Figure 2 Subreddit distribution Correlation with Gender

Figure2 represents the heatmap of gender distribution in the top 10 subreddits provided insights into the participation of male and female users across Reddit dataset. The Subreddit AskDocs had the highest participation, with a significant majority of female users (10,235) compared to male users (3,227). Subreddits such as ADHD, Anxiety, depression, and lexapro also exhibited higher female participation.

Within the AskaPatient dataset (Detailed in appendix1), the drugs Cymbalta and Lexapro have the highest number of reviews from both female and male users. Additionally, most of the age group falls within the 40-59 range.

Linguistic Features extraction- Emotional Features and Correlation analysis

Figure 3 displays the analysis result of the AskaPatient dataset, revealing district emotional trends across different demographics. The emotion fear has the highest mean value, followed by sadness, neutral, joy, anger, disgust, and surprise. Fear is the dominant emotion for both male and female, with males showing a slightly higher mean value. Females exhibit higher mean values for sadness and neutral emotions compared to males. The emotion joy, anger, disgust, and surprise show relatively lower mean values for both genders.

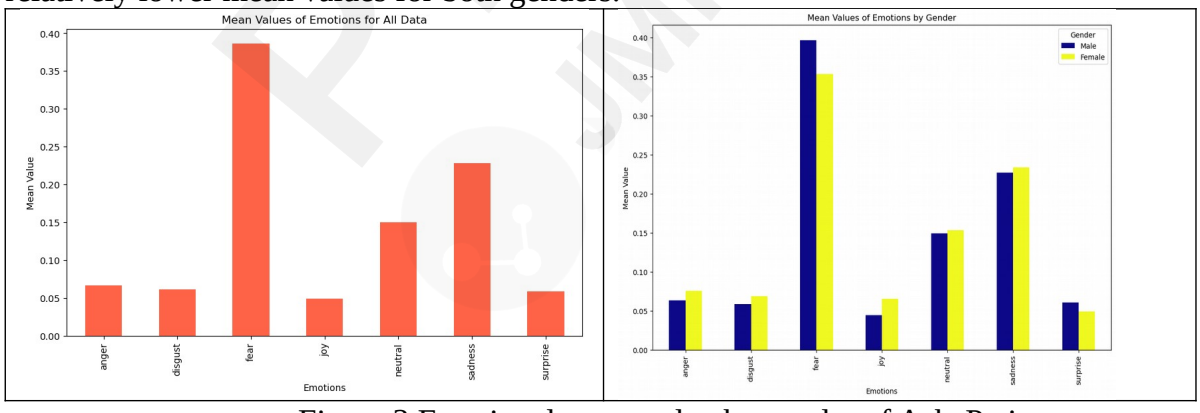


Figure 3 Emotional mean value by gender of AskaPatient

The emotion fear is the most frequently expressed across all age groups, with the highest counts observed in the 35-45 age group, followed by 25-35 age group. The emotion sadness is most frequently in 25-35 age group, followed by 45-55 age group. (Figure details in Appendix1)

Figure 4 represents the emotion features in different demographics within Reddit dataset. The emotion fear has the highest mean value, followed by sadness, neutral, joy, anger, disgust, and surprise. Females have a slightly higher mean value for fear compared to males. In contrast, males have a higher mean value for sadness compared to females.

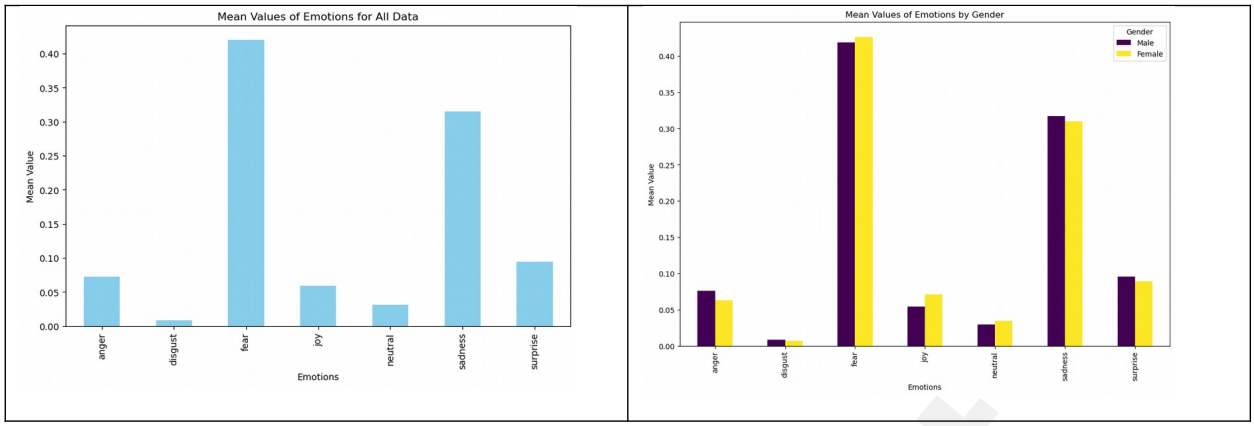


Figure 4 Emotional Mean value by gender of Reddit

For the distribution of different emotion labels across age groups (details in Appendix1), fear and sadness emotions have the highest counts observed in the 15-25 age groups, followed by a significant gap with the 25-35 age groups.

Linguistic Features extraction- Semantic Features and Correlation analysis

Table 2 represents the eight main topics from AskaPatient dataset, each representing different themes related to anxiety medication. The most frequent topic is side effects and experiences of taking anxiety medication, mentioned 331 times (25.3%). The next most frequent topic, with 212 mentions (16.2%), covers the effects of SSRI drugs on sexual health and mental well-being. Lexapro side effects and experiences were mentioned 153 times (11.7%), and Cymbalta's impact on chronic pain and depression over years was discussed 143 times (10.9%). The topic of withdrawal symptoms and challenges of stopping a drug had 132 mentions (10.1%).

Table2 BERTopic result for AskaPatient

Topic	Count	Ratio (%)	CustomName
0	331	25.3	Side Effects and Experiences of Taking Anxiety Medication
1	212	16.2	Effects of SSRI Drugs on Sexual Health and Mental Well-being
2	143	10.9	Cymbalta's Impact on Chronic Pain and Depression Over Years
3	153	11.7	Lexapro side effects and experiences
4	91	7.0	Zoloft for Anxiety Relief - Personal Experiences and Success Stories
5	132	10.1	Withdrawal Symptoms and Challenges of Stopping a Drug
6	91	7.0	Medication Dosage and Side Effects for Anxiety: mg, started, dose, anxiety, side effects
7	69	5.3	Wellbutrin side effects and experiences: anxiety, depression, seizures, weight changes

Zoloft for anxiety relief, including personal experiences and success stories, was mentioned 91 times (7.0%), the same number of mentions as the topic of medication dosage and side effects for anxiety. Finally, Wellbutrin side effects and experiences, including anxiety, depression, seizures, and weight changes, were mentioned 69 times (5.3%). These topics highlight the primary concerns of AskaPatient users regarding medication side effects, the impact of SSRIs on sexual health and mental well-being, and personal experiences with specific drugs like Lexapro, Cymbalta, Zoloft, and Wellbutrin.

The BERTopic analysis on the Reddit dataset yielded several key topics. The top three topics by count are: family relationships and mental health struggles (4,240 mentions, 16.2%), chronic severe headaches and migraines (3,064 mentions, 11.7%), and undiagnosed illnesses, chronic pain, anxiety, chest pain, and high heart rate (2,871 mentions, 11.0%).

Table 3 BERTopic result for Reddit:

Topic	Count	Ratio (%)	CustomName
0	2871	11.0	Undiagnosed illness, chronic pain, anxiety, chest pain, high heart rate
1	4240	16.2	family relationships and mental health struggles
2	3064	11.7	Chronic severe headaches and head pressure with migraines
3	2076	7.9	Issues with sexual intimacy and libido in relationships
4	1845	7.0	Chronic abdominal pain and gastrointestinal issues
5	1842	7.0	Dealing with ADHD Diagnosis and Medication Challenges
6	2519	9.6	Experiences with Lexapro mg for Anxiety
7	1178	4.5	Weight management struggles and challenges with medication, mental health, and lifestyle changes in a busy schedule
8	1290	4.9	Skin rash issues and treatments
9	2138	8.2	Experiences with Zoloft and Anxiety Management
10	1234	4.7	Dealing with Suicidal Thoughts and Narcissistic Relationship Trauma
11	1081	4.1	Persistent abnormal bleeding and challenges with period management while on birth control
12	1080	4.1	Abnormal Blood Test Results with Concerns about Thyroid Function and Testosterone Level
13	1444	5.5	Prozac mg for anxiety and depression - experiences and effects
14	765	2.9	Supportive friend navigating emotional complexities in a recent relationship.
15	1111	4.2	Chronic Sleep Difficulties and Fatigue from Childhood with Various Sleep Medications and Underlying Health Conditions
16	1032	3.9	Experiences with Effexor and Cymbalta Withdrawal and Replacement Options
17	844	3.2	Wellbutrin XL mg for Depression and Anxiety
18	425	1.6	Urinary Symptoms and UTI Management
19	471	1.8	Personal experiences with drug use, addiction, and mental health
20	383	1.5	Toe infection treatment and self-care concerns
21	349	1.3	Eye Condition Variances and Visual Changes
22	381	1.5	Alcohol-Related Anxiety and Withdrawal Symptoms
23	355	1.4	Seizure Symptoms and Medication Concerns
24	268	1.0	Pulsatile Tinnitus and Anxiety in Female with History of Loud Noise Exposure
25	229	0.9	challenges with intimacy in a year-long relationship
26	227	0.9	Mental health struggles and tragic incidents involving Elisa Lam and Sophie Elliott

Additionally, sex-related topics are prominently discussed, including issues with sexual intimacy and libido in relationships (2,076 mentions, 7.9%), persistent abnormal bleeding and challenges with period management while on birth control (1,081 mentions, 4.1%), dealing with suicidal thoughts and narcissistic relationship trauma (1,234 mentions, 4.7%), supportive friends navigating emotional complexities in recent relationships (765 mentions, 2.9%), challenges with intimacy in a year-long relationship (229 mentions, 0.9%), and mental health struggles and tragic incidents involving Elisa Lam and Sophie Elliott (227 mentions, 0.9%). These topics highlight the significant concerns of Reddit users regarding family dynamics, chronic pain, mental health, sexual health, and intimate relationships.

Linguistic Features extraction- Domain Specific Features and Correlation analysis

The result of domain specific features shows a gender-based distribution, detailed in figures (see Appendix). Females report a higher count of most symptoms compared to males, including

gastrointestinal symptoms, neurological symptoms, sleep disorders, general physical symptoms, and emotional and psychological symptoms. Conversely, males report more sexual health issues than females.

The figure5 presents a SHAP (SHapley Additive exPlanations) analysis within AskaPatient dataset, showcasing the impact of different side effect categories on a specific outcome, likely related to gender. Gastrointestinal symptoms and neurological symptoms have a strong positive impact in determining the female (1). Sleep disorder and sexual health issues show slight positive impact, reflecting their contribution to male (0).

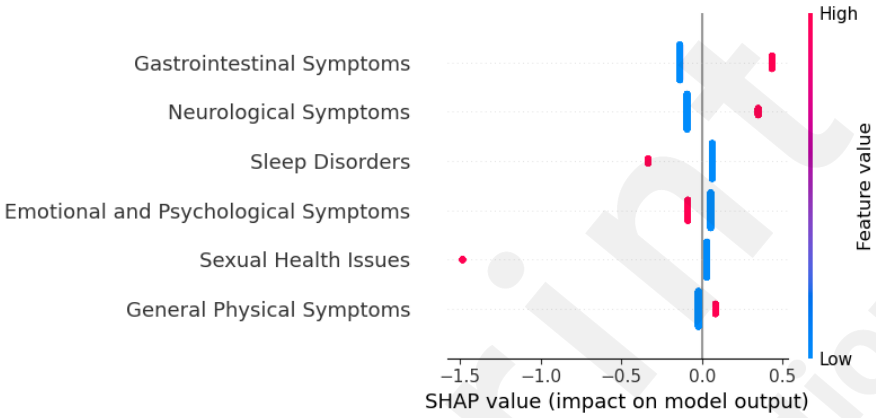


Figure 5 SHAP analysis within AskaPatient

Figure6 shows the correlation between various side effects and emotion in the AskaPatient dataset. The emotion fear is strongly positively related to neurological symptoms, gastrointestinal symptoms, and sleep disorder. The emotion disgust shows a slightly positive correlation with gastrointestinal symptoms.

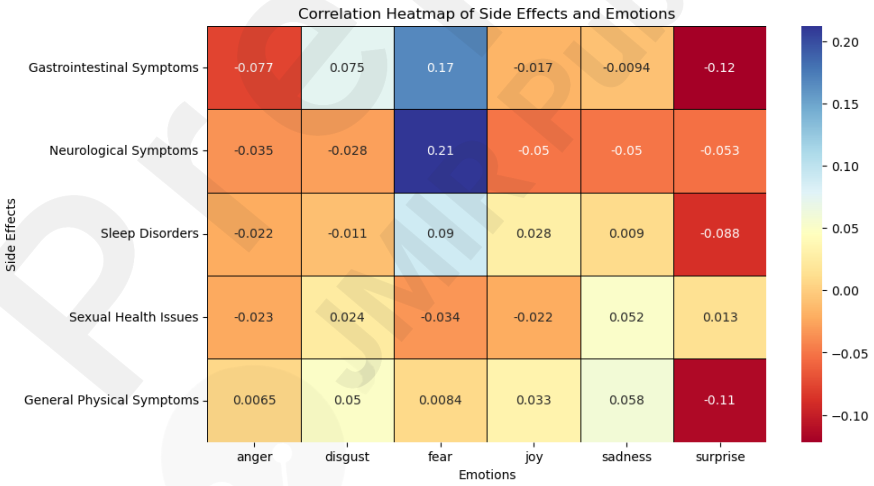


Figure 6 Correlation between side effects and emotions in the AskaPatient

The SHAP (Figure7) plot demonstrates the impact of different side effects on the prediction of the model's prediction of gender within Reddit, with male (0) and female (1). Emotional and psychological symptoms, sexual health issues, and sleep disorder shows a slightly positive impact in predicting gender, indicating a slightly association with males. Gastrointestinal symptoms, general physical symptoms and neurological symptoms show a slightly negative impact, indicating a slight association with females.

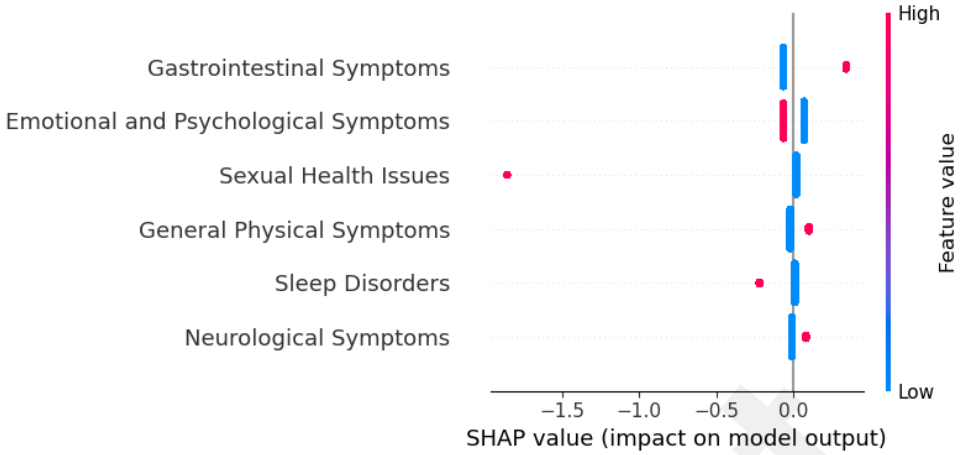


Figure7 SHAP analysis within Reddit

The heatmap (Figure8) illustrates the correlation between different side effects and various emotions within the Reddit dataset. Gastrointestinal symptoms are positively correlated with fear (0.085) and disgust (0.035). Neurological symptoms are positively correlated with fear (0.093). Sleep disorders are positively correlated with fear (0.038). General physical symptoms show a slight positive correlation with fear (0.022) and sadness (0.021).

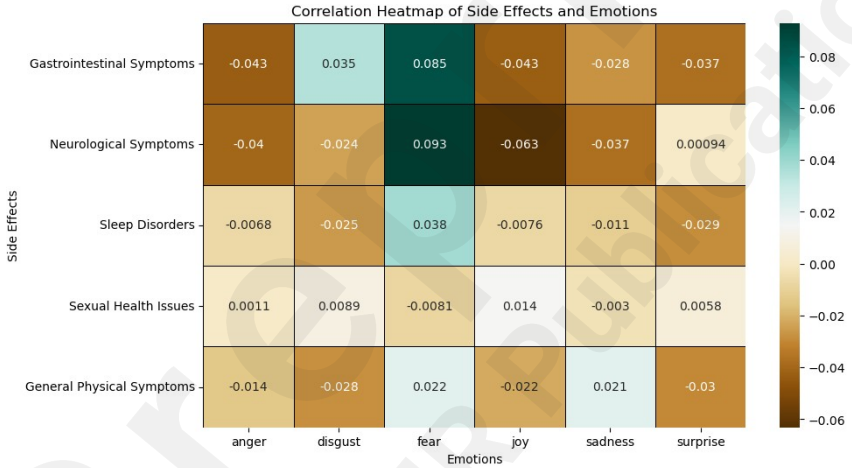


Figure8 Correlation heatmap of side effects and emotions

Discussion

Principal Results

The descriptive statistics of the Reddit and AskaPatient datasets provide significant insights into the demographic characteristics of the participants discussing antidepressants online. The Reddit dataset indicates a younger demographic, while older individuals are more likely to review antidepressants on AskaPatient. Gender distribution analysis reveals that females are significantly more active on both platforms, with their participation being approximately three times that of males. This could be due to the higher prevalence of conditions such as anxiety and depression among women, which often necessitate medication[34]. Additionally, on AskaPatient, drugs like Cymbalta and Lexapro receive numerous reviews from both female and male users, but the overall participation still skews heavily female. Most users fall within the 40-59 age range, suggesting that middle-aged women are particularly engaged in sharing their medication experiences. The higher female participation on both platforms underscores the need for gender-sensitive approaches in health communication and support. This gender disparity may reflect broader trends in health communication, where women are generally more likely to seek health information and support online [35]. Women's greater engagement with health-related content online indicates their proactive role in managing health,

which can be leveraged to design more effective digital health interventions.

The emotion analysis reveals several significant findings about the emotional landscape in discussions of antidepressant use. The emotion fear has the highest mean value overall and is particularly prevalent among both genders and across all age groups. This suggests that fear is a dominant emotion among individuals discussing antidepressants, potentially reflecting concerns about side effects and efficacy associated with mental health issues. One research reveals that even before people take their first antidepressant, they are influenced by the controversial publicity surrounding the medication. For instance, everyone seems to have friends and relatives with strong opinions about antidepressants, including fears of addiction [36]. Sadness follows closely, indicating that individuals often express lower emotions when discussing antidepressants online. Healthcare providers should be aware of the predominant emotions of fear and sadness and consider these factors when supporting patients. Tailoring communication and interventions to address these emotional responses could improve patient outcomes and adherence to treatment.

The BERTopic analysis of the AskaPatient and Reddit datasets highlights significant themes and concerns regarding antidepressant use, showing both common and unique symptoms reported by users. The analysis identified domain side effects six features, particularly focusing on gastrointestinal issues, neurological symptoms, sleep disorders, and sexual health problems. In addition to these commonly reported symptoms, our analysis uncovered several unique concerns not typically highlighted in traditional drug descriptions.
Additionally, users frequently discussed withdrawal symptoms and challenges of replacing one antidepressant with another, revealing anxiety about the withdrawal process. This concern is critical as it can lead to medication non-adherence or discontinuation without proper medical guidance. Furthermore, issues related to ADHD diagnosis and medication challenges were found to exacerbate depression and anxiety symptoms, indicating the need for careful management of these conditions in conjunction with antidepressants. The analysis also highlighted heightened anxiety related to alcohol and substance use during the withdrawal process, emphasizing the importance of addressing substance use issues as part of comprehensive care.
Other unique symptoms such as toe infections, visual changes (eye conditions), and pulsatile tinnitus were mentioned by users, which are not typically emphasized in standard drug descriptions. Additionally, users reported concerns about sexual intimacy, the impact of birth control on period management, and abnormal bleeding while on birth control, which are not commonly discussed in general side effect listings. Women in reproductive age who use birth control pills need to be aware of how these antidepressants can affect the action of their contraceptive pills in order to avoid side effects and prevent the reduced effectiveness of both[37].

The analysis of domain-specific features highlights significant gender-based differences in the reporting of symptoms related to antidepressant use. Females report a higher incidence of gastrointestinal, neurological, sleep, general physical, and emotional and psychological symptoms compared to males. This disparity suggests that women may experience or are more likely to report a broader range of side effects from antidepressants. Conversely, males report more issues related to sexual health, indicating a gender-specific concern that

might affect treatment adherence and overall satisfaction with antidepressant therapy. The SHAP analysis underscores these findings by demonstrating that gastrointestinal and neurological symptoms are strong indicators for female patients. In contrast, sleep disorders and sexual health issues are more predictive for male patients. This differentiation in symptom reporting and impact highlights the necessity for gender-sensitive healthcare approaches. For instance, recognizing that women may have a broader spectrum of side effects can guide healthcare providers to offer more comprehensive monitoring and support. Emotional responses also vary significantly with gender. The correlation between fear and neurological, gastrointestinal, and sleep-related symptoms is particularly strong. This suggests that these physical symptoms may contribute significantly to the emotional distress experienced by patients. In contrast, disgust is more specifically associated with gastrointestinal issues, indicating a nuanced emotional response to certain side effects.

These insights are crucial for tailoring patient care. Healthcare providers should be mindful of the broader range of side effects reported by female patients and the specific concerns related to sexual health in male patients. Addressing these issues directly in consultations can improve patient outcomes by fostering better adherence to medication regimens and enhancing overall treatment satisfaction. Moreover, understanding the emotional correlates of physical symptoms can help in providing more holistic care that addresses both the physical and psychological needs of patients.

Limitations

The findings of this study should be interpreted with caution due to several limitations. Firstly, the data analyzed from the AskaPatient, and Reddit datasets may not be generalizable to the broader population, as users of these platforms might have different characteristics and motivations compared to the public. Secondly, there is a gender imbalance in the data, with a significantly higher number of female participants. In general, it is important to acknowledge potential bias introduced by the overall demographics of Reddit users, who are predominantly White, male, and college-educated individuals who are 18-29 years of age and living in the United States[38]. These biases could affect the features of the findings, particularly concerning gender-specific side effects and emotional responses to antidepressants. Additionally, the self-reported nature of the data means that the accuracy and completeness of the reported symptoms cannot be verified. Future research should strive to include more balanced gender representation and explore the impact of antidepressant use in specific contexts, such as birth control interactions and postpartum depression in women.

Conclusions

The BERTopic analysis of the AskaPatient and Reddit datasets has provided valuable insights into the common and unique concerns regarding antidepressant use reported by users. While the analysis identified several well-documented side effects, such as gastrointestinal, neurological, and sexual health issues, it also uncovered unique concerns not typically highlighted in traditional drug descriptions. These include issues related to medication withdrawal, ADHD diagnosis, alcohol and substance use during withdrawal, and specific physical conditions like toe infections, visual changes, and pulsatile tinnitus. The study emphasizes the importance of gender-sensitive approaches in healthcare, highlighting the broader range of side effects reported by female patients and the specific concerns of male patients related to sexual health. These findings underscore the need for comprehensive monitoring and support tailored to the diverse experiences of individuals using antidepressants.

Conflicts of Interest

The authors confirm that these financial relationships did not influence the design, execution, interpretation, or reporting of this research. All authors have disclosed these interests to the journal and have adhered to the journal's policies on conflicts of interest.

References

1. Brody D, Gu Q. Antidepressant Use Among Adults: United States, 2015-2018. Published 2020. Accessed May 23, 2024. https://www.cdc.gov/nchs/products/databriefs/db377.htm
2. Olfson M, Marcus SC. National trends in outpatient psychotherapy. *Am J Psychiat*. 2010;167(12):1456-1463. doi:10.1176/appi.ajp.2010.10040570
3. Fix GM, VanDeusen Lukas C, Bolton RE, et al. Patient-centred care is a way of doing things: how healthcare employees conceptualize patient-centred care. *Health Expectations*. 2018;21(1):300-307. doi:10.1111/hex.12615
4. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiat*. 2006;163(1):28-40. doi:10.1176/appi.ajp.163.1.28
5. Zolnoori M, Fung KW, Fontelo P, et al. Identifying the underlying factors associated with patients' attitudes toward antidepressants: Qualitative and quantitative analysis of patient drug reviews. *JMIR Ment Health*. 2018;5(4):e10726. doi:10.2196/10726
6. Lardon J, Abdellaoui R, Bellet F, et al. Adverse drug reaction identification and extraction in social media: a scoping review. *J Med Internet Res*. 2015;17(7):e4304. doi:10.2196/jmir.4304
7. Gupta R, Ariefdjohan M. Mental illness on instagram: a mixed method study to characterize public content, sentiments, and trends of antidepressant use. *J Ment Health (Abingdon Engl)*. 2021;30(4):518-525. doi:10.1080/09638237.2020.1755021
8. Scheuer K, Salihu EY. Discussions of antidepressant side effects and withdrawal on reddit. Published online March 23, 2023. doi:10.31234/osf.io/7psfz
9. Butterworth P, Olesen SC, Leach LS. Socioeconomic differences in antidepressant use in the PATH through life study: evidence of health inequalities, prescribing bias, or an effective social safety net? *J Affect Disorders*. 2013;149(1-3):75-83. doi:10.1016/j.jad.2013.01.006
10. Soest T von, Bramness JG, Pedersen W, Wichstrøm L. The relationship between socio-economic status and antidepressant prescription: a longitudinal survey and register study of young adults. *Epidemiol Psychiatr Sci*. 2012;21(1):87-95. doi:10.1017/S2045796011000722
11. Sramek JJ, Murphy MF, Cutler NR. Sex differences in the psychopharmacological

treatment of depression. *Dialogues Clin Neurosci*. 2016;18(4):447-457.
12. Seifert J, Führmann F, Reinhard MA, et al. Sex differences in pharmacological treatment of major depressive disorder: results from the AMSP pharmacovigilance program from 2001 to 2017. *J Neural Transm*. 2021;128(6):827-843. doi:10.1007/s00702-021-02349-5
13. Strawn JR, Mills JA, Suresh V, et al. The impact of age on antidepressant response: a mega-analysis of individuals with major depressive disorder. *J Psychiatr Res*. 2023;159:266-273. doi:10.1016/j.jpsychires.2023.01.043
14. Gutsmiedl K, Krause M, Bighelli I, Schneider-Thoma J, Leucht S. How well do elderly patients with major depressive disorder respond to antidepressants: a systematic review and single-group meta-analysis. *BMC Psychiatry*. 2020;20(1):102. doi:10.1186/s12888-020-02514-2
15. Saha K, Sugar B, Torous J, Abrahao B, Kıcıman E, De Choudhury M. A social media study on the effects of psychiatric medication use. *Proc Int AAAI Conf Weblogs Soc Media, Int AAAI Conf Weblogs Soc Media*. 2019;13:440-451.
16. Lien L, Bonsaksen T, Stea TH, Kleppang AL, Steigen AM, Leonhardt M. Time trends in self-reported depressive symptoms, prescription of antidepressants, sedatives and hypnotics and the emergence of social media among norwegian adolescents. *PLoS One*. 2023;18(12):e0295384. doi:10.1371/journal.pone.0295384
17. Saha K, Torous J, Kiciman E, Choudhury MD. Understanding side effects of antidepressants: large-scale longitudinal study on social media data. *JMIR Ment Health*. 2021;8(3):e26589. doi:10.2196/26589
18. de Anta L, Alvarez-Mon MA, Ortega MA, et al. Areas of interest and social consideration of antidepressants on english tweets: a natural language processing classification study. *J Pers Med*. 2022;12(2):155. doi:10.3390/jpm12020155
19. Ask a Patient - Drug Ratings and Patient Reviews: a Forum on Medications. Accessed May 19, 2024. https://www.askapatient.com/
20. re — Regular expression operations. Python documentation. Accessed May 19, 2024. https://docs.python.org/3/library/re.html
21. Park M, Cha C, Cha M. Depressive Moods of Users Portrayed in Twitter.
22. Hartmann J. Emotion english distilroberta-base (2022).
23. j-hartmann/emotion-english-distilroberta-base · Hugging Face. Accessed May 20, 2024. https://huggingface.co/j-hartmann/emotion-english-distilroberta-base
24. Egger R, Yu J. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify twitter posts. *Front Sociol*. 2022;7. doi:10.3389/fsoc.2022.886498
25. Argyle LP, Busby EC, Fulda N, Gubler J, Rytting C, Wingate D. Out of one, many: Using language models to simulate human samples. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. ; 2022:819-862. doi:10.18653/v1/2022.acl-long.60
26. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. Published online March 11, 2022. doi:10.48550/arXiv.2203.05794
27. Rausch SL. Managing the side effects of depression treatment. WebMD. Accessed May 8, 2024. https://www.webmd.com/depression/features/coping-with-side-effects-of-depression-treatment
28. Side effects - antidepressants. nhs.uk. Published February 5, 2021. Accessed May 8, 2024. https://www.nhs.uk/mental-health/talking-therapies-medicine-treatments/medicines-and-psychiatry/antidepressants/side-effects/
29. DailyMed. Accessed May 8, 2024. https://dailymed.nlm.nih.gov/dailymed/index.cfm
30. KeyBERT - KeyBERT. Published March 15, 2024. Accessed March 13, 2024. https://maartengr.github.io/KeyBERT/api/keybert.html

31. sentence-transformers/all-MiniLM-L6-v2 · hugging face. Published January 18, 2024. Accessed May 8, 2024. https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2
32. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. 2020;2(1):56-67. doi:10.1038/s42256-019-0138-9
33. Reddit comments/submissions 2005-06 to 2022-12 : reddit : free download, borrow, and streaming. Internet Archive. Accessed February 22, 2024. https://archive.org/details/pushshift_reddit_200506_to_202212
34. Albert PR. Why is depression more prevalent in women? *J Psychiatry Neurosci*. 2015;40(4):219-221. doi:10.1503/jpn.150205
35. Duggan SF and M. Health Online 2013. Pew Research Center. Published January 15, 2013. Accessed May 26, 2024. https://www.pewresearch.org/internet/2013/01/15/health-online-2013/
36. Research shows that patients fear antidepressants a “dirty little habit.” Monash University. Published November 4, 2015. Accessed May 26, 2024. https://www.monash.edu/news/articles/collaborative-research-shows-that-patients-fear-antidepressants-a-dirty-little-habit
37. vikramkadmin. Taking Birth Control Pill And Antidepressant At The Same Time. Her Smart Choice. Published March 12, 2020. Accessed May 26, 2024. https://hersmartchoice.com/blog/is-it-safe-to-take-birth-control-pills-and-antidepressants-at-the-same-time/
38. Mitchell MB Galen Stocking, Jesse Holcomb and Amy. 1. Reddit news users more likely to be male, young and digital in their news preferences. Pew Research Center. Published February 25, 2016. Accessed May 27, 2024. https://www.pewresearch.org/journalism/2016/02/25/reddit-news-users-more-likely-to-be-male-young-and-digital-in-their-news-preferences/