

Digital Conversations about Severe Depression Symptoms Across Different Ethnic and Racial Groups: A Big-data, Machine Learning Analysis

Ruby Castilla-Puentes^{1,2,*}, Silvia Ferrer^{2,3}, Ariceli Alfaro^{2,4}, Laura Jimenez-Parrado^{2,5}, Whitney S. Cordoba-Grueso^{2,6}, Sebastian Castilla^{2,7}

¹University of Cincinnati, 2600 Clifton Ave, Cincinnati, OH 45221, USA

²WARMI Mental Health, 7709 Cornell Rd, Cincinnati, OH 45242, USA

³Universidad Del Norte, Puerto Colombia, Colombia, Área metropolitana de, Kilómetro 5, Vía Puerto Colombia, Barranquilla, Atlántico, Colombia

⁴University of California, Irvine, CA 92697-5705, USA

⁵Albert Einstein College of Medicine - PRIME. Bronx, New York, 1225 Morris Park Avenue, NY 10461, USA

⁶Indiana University, 107 S. Indiana Avenue, Bloomington, IN 47405, USA

⁷University of Alabama, 600 University Blvd. Suite C, Tuscaloosa, Alabama 35487, USA

*Correspondence should be addressed to Ruby Castilla-Puentes, castilrb@mail.uc.edu

Received date: October 23, 2024, **Accepted date:** January 03, 2025

Citation: Castilla-Puentes R, Ferrer S, Alfaro A, Jimenez-Parrado L, Cordoba-Grueso WS, Castilla S. Digital Conversations about Severe Depression Symptoms Across Different Ethnic and Racial Groups: A Big-data, Machine Learning Analysis. Arch Psychiatry. 2025;3(1):1-11.

Copyright: © 2025 Castilla-Puentes R, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Background: Although studies have explored the use of technology and social media to help minorities suffering from depression, prior research has not thoroughly analyzed the racial and ethnic variation in the digital conversations related to symptoms of severe depression across racial/ethnic groups in the United States (U.S.).

Method: Machine-learning methods were used to extract open-source online conversations in the US from February 1, 2019, to November 1, 2020. The information included self-identified racial/ethnic groups: Hispanics, Non-Hispanic whites (NHW), African Americans and Asian Americans. Symptoms of Severe Depression were defined by the term "depression" and included at least two of the pre-determined severity adjectives described by the users in the conversation. Analyses were conducted for four domains: 1) Topics Generated, 2) Sentiments, 3) Mindset, and 4) Path to Treatment.

Results: A total of 1.3 million unique conversations referring to symptoms of severe depression posted during the selected period were analyzed. Conversations were most frequent among NHW 54%, Hispanics 21%, African Americans 20%, and 6% Asian Americans. Conversations were different across racial and ethnic groups: NHW talked more about diagnosis, making their conversations along the path to treatment more balanced out between the stages. They were more proactive than any other racial/ethnic groups. Depression was perceived as a more social phenomenon among African Americans. Asian Americans had the highest percentage of positive sentiment oriented toward the world and the future. Hispanics were less proactive, more negative, symptomatic, and less involved in treatment when compared with the conversations of other individuals of other racial/ethnic groups.

Conclusions: In conclusion, we have shown that conversations referring to symptoms of depression differs by race/ethnicity, and that these results highlight opportunities for culturally competent approaches to address areas amenable to change that could impact the ability of people to seek and receive mental health support. Future studies identifying ethnic/racial variations in severe depression symptoms may help to improve equity in mental health care.

Keywords: Artificial Intelligence, Severe, Depression, Symptoms, Ethnic, Racial

Introduction

Depression is a common but serious mental disease that ranges widely in severity. It affects 5% of adults worldwide, and approximately 280 million people in the world suffer from the condition [1]. The National Center for Health Statistics reported in 2019 that 2.8% of adults had symptoms of severe depression, 4.2% had moderate depression, and 11.5% had mild depression, in the two weeks recorded [2]. One way to early detect symptoms of depression is to analyze what people are talking about - their conversations. Symptoms of severe depression include insomnia or excessive sleeping, irritability, loss of interest in activities that used to be enjoyable, hopelessness, persistent thoughts of something bad happening, thoughts of death or suicide, and suicide attempts. In very severe cases, psychotic symptoms, such as hallucinations or delusions, and the inability to take care of oneself in things such as eating, bathing, or fulfilling family or work responsibilities are present. Researching for symptoms of severe depression should be prioritized because it results in high suicide rates, influences mental health-related disabilities, negatively affects family dynamics due to loss of income, and increases healthcare costs due to frequent emergency room visits [3].

Race and ethnicity are complex multidimensional constructs comprising heterogeneous societal and cultural factors, and thus, in some contexts, they may appear as a proxy for social determinants of health [4]. Evidence indicates that race/ethnicity disparities exist in both the burden and the treatment of depression [5]. There is a known association between stressful life events (SLEs) and depression. SLEs can be seen in racial minorities from lower socioeconomic backgrounds, whose lives are compounded by abject or perceived racism, a dearth of education, communal violence, single-family households, or substance abuse [6].

Our group has investigated Hispanics' digital conversations and found that Hispanics engage in more negative conversations and have attitudes of resignation and hopelessness towards depression more often than non-Hispanics [7]. Although studies have explored the use of technology and social media to help minorities suffering from depression, prior research has not thoroughly analyzed the racial and ethnic variation in the digital conversations related to severe symptoms of depression.

Symptoms of depression are present differently across racial and ethnic boundaries. As literature suggests, disparities are present in all facets of the disease – in risk factors, presentation of disease, type and severity of symptoms, and modalities of care offered [8]. To our knowledge, Artificial Intelligence (AI) has not been used to investigate symptoms of severe depression. The analysis of digital conversations provides a unique opportunity to understand the patient's real-world

concerns and perspectives. Our study uses open-source digital conversations to explore the topics generated by AI (sentiments, mindset, and path to treatment) to evaluate the characteristics of conversations regarding severe depression symptoms among racial/ethnic groups. This project addressed the following research questions:

1. Are there ethnic variations in prevalence of symptoms of severe depression?
2. How do symptoms of severe depression vary across different ethnic groups?

Our hypothesis is that digital conversations related to symptoms of severe depression will differ across racial and ethnic groups.

Methods

The focus for this cross-sectional study was the conversations of adults from different ethnic origins. They were categorized into four different groups (based on their self-identification during digital conversations or on their public online profile): 1) Hispanics (including Latino, Latinx, Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish cultures or origin regardless of race), 2) non-Hispanic whites (NHw), 3) African Americans and 4) Asian Americans. Those who reported symptoms of severe depression or engaged in conversations about severe depression in the United States were included in the analysis. CulturIntel™ used its advanced AI-based tools to mine and structure the unstructured, qualitative online data on the topic of interest (i.e., symptoms of severe depression). The search included a total of 4.2 million conversations about depression from February 1, 2019, to November 1, 2020, originating from US Internet Protocol (IP) addresses. Almost 70% of these conversations were excluded from the final analysis due to depression not being severe, leading to a final sample of 1.3 million conversations. No data were obtained from private accounts or from discussions behind firewalls. De-identified digital conversations in English were gathered from various online, open-source topical sites (e.g. Depression and Bipolar Support Alliance), message boards (e.g. Beyond Blue), social networks (e.g. Facebook), and blogs, all of which are in the public domain. Categories of digital conversations were described using counts and percentages. Chi square tests were used for bivariate analyses among conversation topics. Fisher's exact test was used for bivariate analysis among sentiments.

Social media data can detect an individual's depression based on their posted negative opinions [9]. Big data can be processed in two approaches: big data batch processing and big data stream processing. Big data batch processing is suitable for creating a depression detection model, while big data stream processing is necessary for real-time depression detection [10].

Park *et al.* proposed a framework for analyzing sentiments from Twitter and searching for features that identified depression. The experimental results gathered from 69 participants found that posting negative messages, messages expressing depression, and the negative Twitter emoticon, or emotion icon, was significantly associated with the occurrence of depression symptoms [11]. Symptoms of severe depression were defined by the term “depression” and included at least two of the pre-determined severity adjectives described by the users in the conversation (e.g., severe, debilitating, relentless, inescapable, terrible, awful, dreadful, serious, horrible, dire, worst, misery, constant, non-stop, never-ending, horrendous, frightening, hellacious, appalling, extreme, hopelessness, suicidal, despair, thought interference, helplessness or uncontrollable). This text identification follows published recommended approaches, given the current lack of a standard model [12].

Ethics consideration

All the information gathered from the different online, open sources (topical sites [eg, Depression and Bipolar Support Alliance], message boards [eg, Beyond Blue], social networks [eg, Facebook], and blogs) is in the public domain and is deidentified. The study was exempt from Institutional Review Board approval as it used publicly available, deidentified information. However the following key ethical principles were included: transparency, privacy protection, accountability, data quality, minimization of harm, and ensuring the appropriate use of collected data; meaning clearly explaining how data is being used, safeguarding sensitive information, taking responsibility for the analysis, maintaining data accuracy, minimizing potential negative impacts, and only using data for its intended purpose.

CulturIntel™ big data and AI suite of tools ‘scrapes and listens’ to open-source conversations online. Data mining and collection occurs across various sites (topical sites, blogs, social network, and message boards) where relevant discussions are taking place. It encompasses a complete range of social discussion channels, including sites directed toward selected segments and directed by (but not limited to) their predefined topic of interest (racial/ethnic groups and severe depression). Advanced search techniques of web crawlers/scrapers (Google technology that archives (copies and stores) a pre-programmed collection of website/topic/discussion data, as it generated on the internet) were applied. CulturIntel™ then extracted the topical data, tagged these data with their origin and user, which were then de-identified. Subsequently a large, unstructured ‘big’ dataset is created. After the completion of this comprehensive dataset, NLP, and text analytics were employed to examine previously described and undescribed patterns in the data. NLP analyses were supervised, whereby the authors worked with CulturIntel™ team in analyzing and tagging a sub-set of the ‘big’ dataset. In this process, digital conversations were reviewed by authors and based on understanding of English language; each conversation was

tagged for the positive or negative tone, attitude, etc. towards depression. This random sub-set was used for initial training, testing, and reviewing of a larger dataset by NLP tools, followed by multiple iterations and repetitions of the process leading to supervised machine learning. In this thematic analysis, authors tagged and sorted the data, determined key motivations of topics being discussed, and assigned underlying drivers, attitudes, and topics.

Data analysis

To avoid duplicity of posts/conversations, multiple postings by an individual were included in the analysis only if their post was a unique comment. A single user with multiple posts/comments within a conversation was only counted once. A single comment repeatedly appearing through sharing/linking was counted and analyzed only once as well. However, users posting multiple unique comments across discussions/posts/sites were counted for each comment.

Rather than using keywords, conversations were mined by the topic of depression. Discussions were identified and included if they were related to depression in general (defined by the use of the term “depression” and its adjacencies, such as “feeling depressed”), seeking help for depression (defined by the use of terms such as “help,” “looking for,” “support,” and “assistance”), and severe depression (defined by including “depression” and at least 2 adjectives, e.g., “debilitating, relentless, inescapable, terrible). The digital conversations were primarily analyzed by race/ethnicity in four domains: 1) Topics Generated, 2) Sentiments, 3) Mindset and 4) Path to Treatment [7]. The topics and their stages were selected using natural language processing (NLP) and machine learning algorithms to automatically determine the emotional tone behind online conversations. There are different algorithms used in the analysis models, depending on how many data need analysis and how accurate the model would be [7,13,14]. Conversations about depression reflect different levels of approach to the condition, from focusing on describing the impact to making sense of it and finding ways to manage it. The conversations were further classified based on their tone: positive, neutral, or negative. The conversations with positive and negative tones were then analyzed into heuristic drivers (leading factors) behind the tone. It is understood that the conversations may differ among individuals based on their status in the natural course and self-management of depression. Consequently, the topics and attitudes were analyzed in a framework that mapped the digital conversations into possible stages: suspect (concerned about the possibility of depression), diagnosis, treating (undergoing active treatment changes), coping (just enduring depression, and lacking the constructive perspective to manage it). Of note, not all subjects may go through all these stages.

The categorization of digital conversations is described in detail in **Figure 1**. These methods were used in conjunction with social media data mining to examine patterns in data and perform a thematic analysis [15].

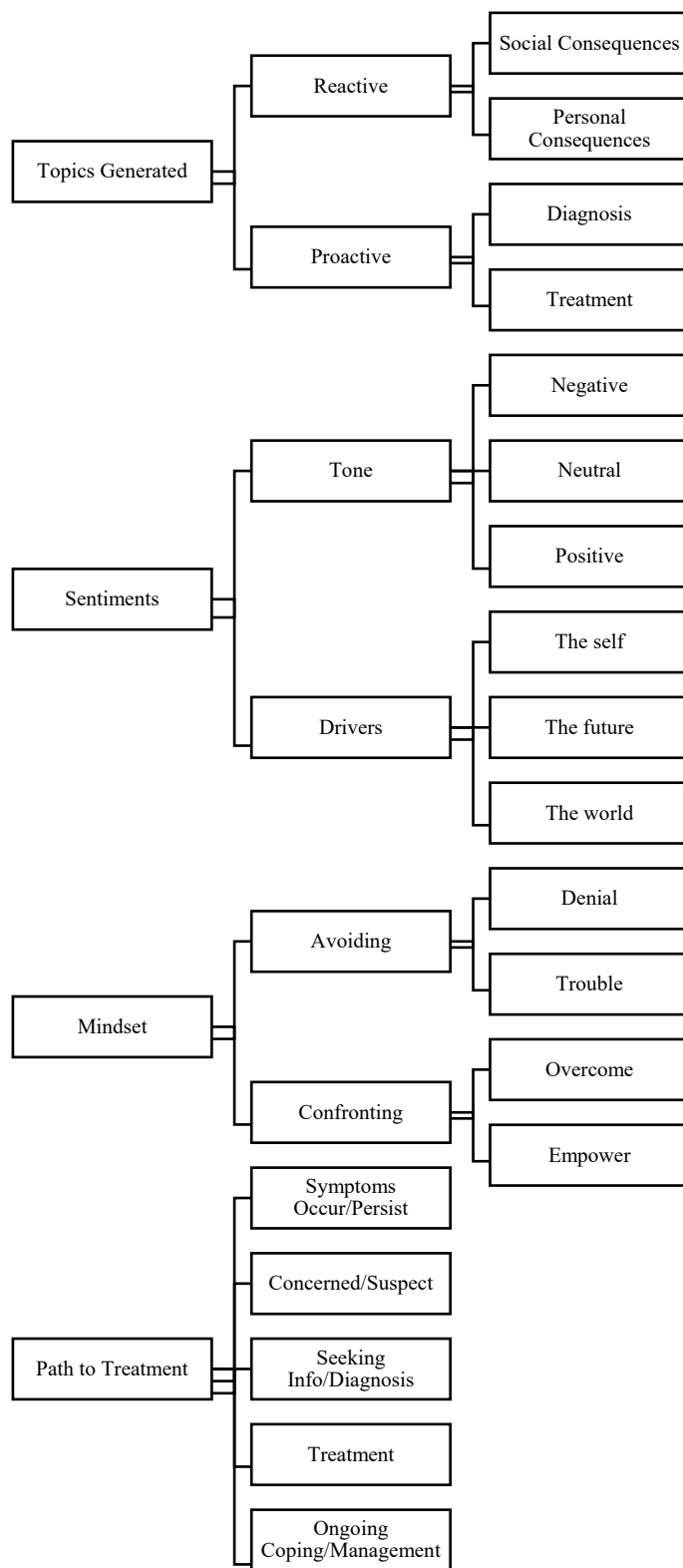


Figure 1. Categorization of digital conversations related to symptoms of severe depression.

Results

A total of 1.3 million unique conversations about symptoms of severe depression were posted during the selected period of the analysis (from February 1, 2019, to November 1, 2020). Race/ethnic analysis of conversations revealed that approximately 54% were posted by NHw, compared to 21% for Hispanics, 20% for African Americans, and 6% for Asian Americans. The distribution is presented in **Figure 2**.

According to source analysis, 38.0% of conversations originated from topical websites (such as Depression and Bipolar Support Alliance—<https://www.dbsalliance.org>), followed by message boards with 28%, social networks with 13%, blogs with 9%, comments with 8%, and content sharing with 4%, respectively. **Figure 3** provides a complete breakdown of the source of these conversations.

Topics generated by racial/ethnic group

Content analysis of conversations revealed similar proportions of reactive and proactive topics by NHw and Asian Americans (50% vs 50%) but, NHw were less concerned about their impact on others and more interested in understanding their diagnosis. Conversations about reactivity were higher among African Americans (59%) and Hispanics (64%), with more focus on personal consequences than social effects among the latter.

Table 1 summarizes the topics generated and provides the percentages of reactive (social and personal consequences) vs. proactive (diagnosis and treatment) themed conversations by the different racial/ethnic groups.

Sentiments

Sentiments users had towards severe depression were classified in tones (negative, neutral, positive) and drivers of negative sentiment (the self, the future, and the world). All the groups showed a predominantly negative tone, with self-perception being the dominant driver in all of them, followed by neutral tone, and finally positive tone. NHw had a negative tone in 68% of conversations. The scarce positive sentiment was driven by getting effective emotional support and improving their quality of life. Their goal was to regain the lost sense of agency and to feel more in control of the future, while also focusing on finding support. The impact of depression on African Americans with severe depression was focused on negative beliefs about themselves (72%), especially on not being able to function and having a poor quality of life, and the positive sentiments were addressed to the world by getting the needed support and treatment in 52% of conversations with positive sentiment. There was a negative tone among 70% of Asian Americans' conversations, driven by negative beliefs about themselves and feeling their condition negatively impacts others. They also had the highest percentage of positive sentiment oriented toward the world

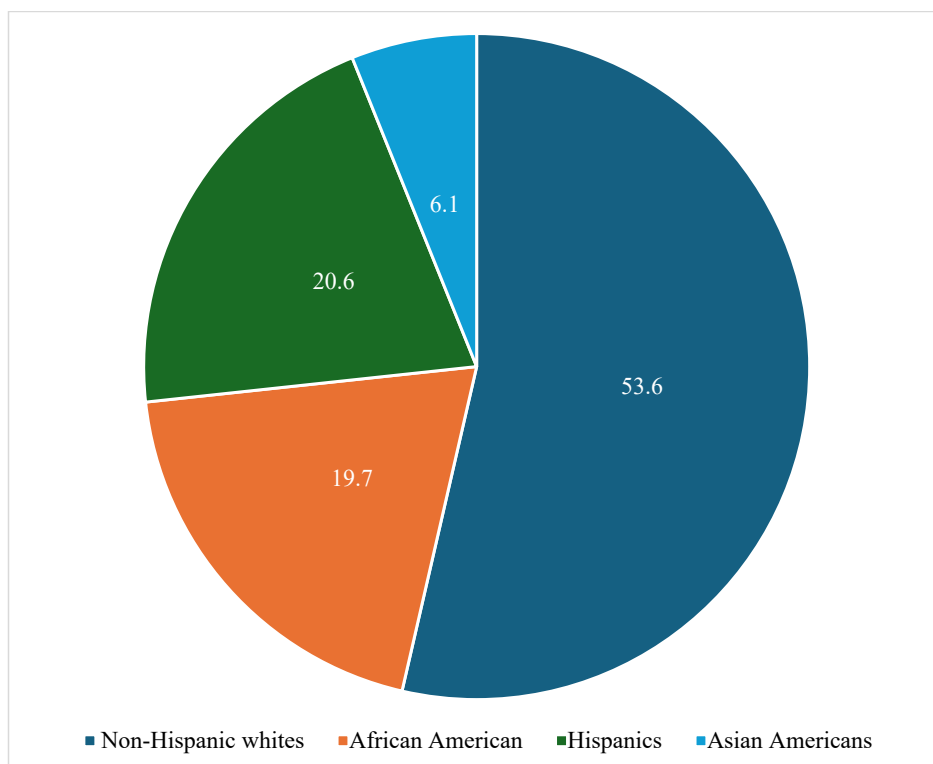


Figure 2. Racial/Ethnic distribution of conversations in which symptoms of depression were determined (N=1.3 million).

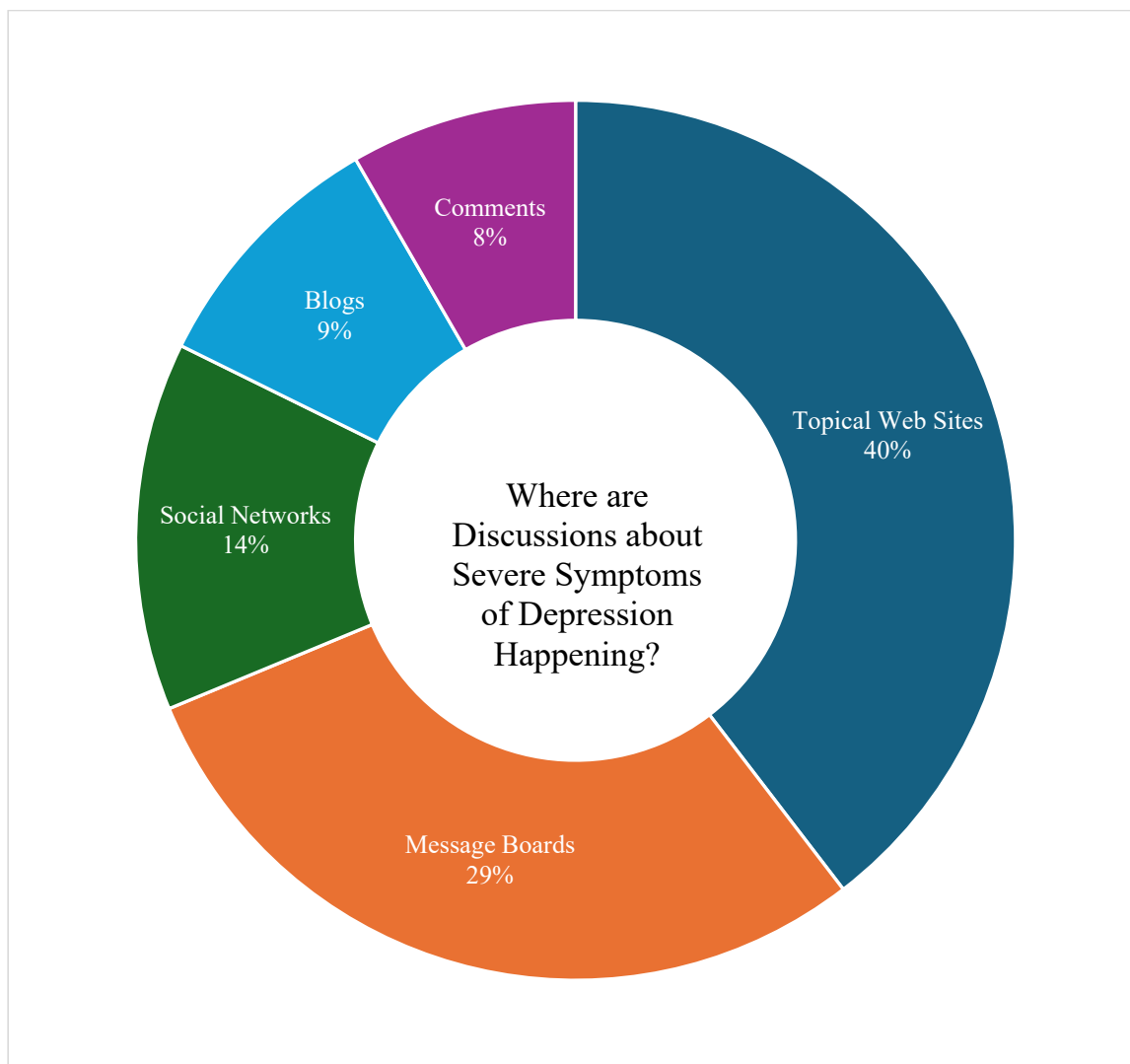


Figure 3. Source of conversations (%) in which symptoms of severe depression were determined.

Table 1. Topics generated in the conversations about severe depression grouped by race/ethnicity – (percentages of all conversations).

Race/Ethnicity		NH whites	African Americans	Asian Americans	Hispanics	X ² , p-value
Reactive	Social Consequences	124,981 (17.91%)	45,006 (17.53%)	11,323 (14.21%)	46,400 (29.00%)	6256, p<0.001
	Personal Consequences	210,829 (30.22%)	98,592 (38.39%)	27,572 (34.59%)	56,948 (35.59%)	
	Total	335,810 (48.13%)	143,598 (55.92%)	38,895 (48.80%)	103,348 (64.59%)	
Proactive	Diagnosis	221,140 (31.70%)	82,500 (32.13%)	25,504 (32.00%)	39,911 (24.94%)	6254, p<0.001
	Treatment	140,750 (20.17%)	30,702 (11.96%)	15,301 (19.20%)	16,741 (10.46%)	
	Total	361,890 (51.87%)	113,202 (44.08%)	40,805 (51.20%)	56,652 (35.41%)	

Note: Topics generated (sentiments, mindset, and path to treatment) were classified further in reactive and proactive. Significance p<.001.

and the future, which means that they sought assistance and access to different types of treatment. Hispanics have the highest negative tone compared to the other ethnic groups. They were feeling stigmatized by others 2.1 times more than NHw, related to Hispanics suffering from severe depression

having a similar neutral sentiment towards their condition than other Hispanics and NHw understanding their situation and looking for information. **Table 2** and **Figure 4** present the sentiments associated and provides the percentages grouped by race/ethnicity.

Table 2. Sentiments associated with the conversations on symptoms of severe depression, grouped by race/ethnicity – (percentages of conversations).

Race/Ethnicity		NH whites N=697,700	African Americans N=256,800	Asian Americans N=79,700	Hispanics N=160,000	Fisher's exact test, p value	
Sentiments Associated	Negative	474,400 (67.99%)	187,000 (72.82%)	55,300 (69.39%)	120,600 (75.38%)	p<0.001	
	Neutral	195,400 (28.01%)	66,400 (25.86%)	21,300 (26.73%)	39,400 (24.63%)		
	Positive	27,900 (4.00%)	3,300 (1.29%)	2,900 (3.64%)	0 (0%)		
Drivers Group 1	Drivers of the Negative Sentiment	The self	288,972 (60.91%)	94,655 (50.62%)	32,837 (59.38%)	62,485 (51.81%)	p<0.001
		The future	97,148 (20.48%)	34,595 (18.50%)	7,455 (13.48%)	15,900 (13.18%)	
		The world	88,280 (18.61%)	57,750 (30.88%)	15,008 (27.14%)	42,215 (35.00%)	
	Drivers of the Neutral Sentiment	The self	71,791 (36.74%)	25,411 (38.27%)	7,455 (35.00%)	15,119 (38.57%)	p<0.001
		The future	59,641 (30.52%)	21,658 (32.62%)	6,735 (31.62%)	13,697 (34.94%)	
		The world	63,968 (32.74%)	19,331 (29.11%)	7,110 (33.38%)	10,384 (26.49%)	
	Drivers of the Positive Sentiment	The self	10,212 (36.60%)	1,485 (45.00%)	667 (23.00%)	N/A	N/A
		The future	4,867 (17.44%)	99 (3.00%)	232 (8.00%)	N/A	
		The world	12,821 (45.95%)	1,716 (52.00%)	2,001 (69.00%)	N/A	

Note: Sentiment associated (i.e., negative, neutral, or positive) with the conversations about severe depression was mapped to drivers of negative sentiments toward the self, the future, and the world (i.e., losing the quality of life, being unable to function, feeling like a burden, feeling hopeless, dealing with uncertainty, feeling stigmatized, or feeling a lack of support). Drivers of a positive, neutral or negative sentiment toward severe depression related to the self, the future, and the world (i.e., improving the quality of life, making progress, having a sense of agency, finding support, or getting access to treatment) were mapped to the construct of perceived benefit. All groups p<.001.

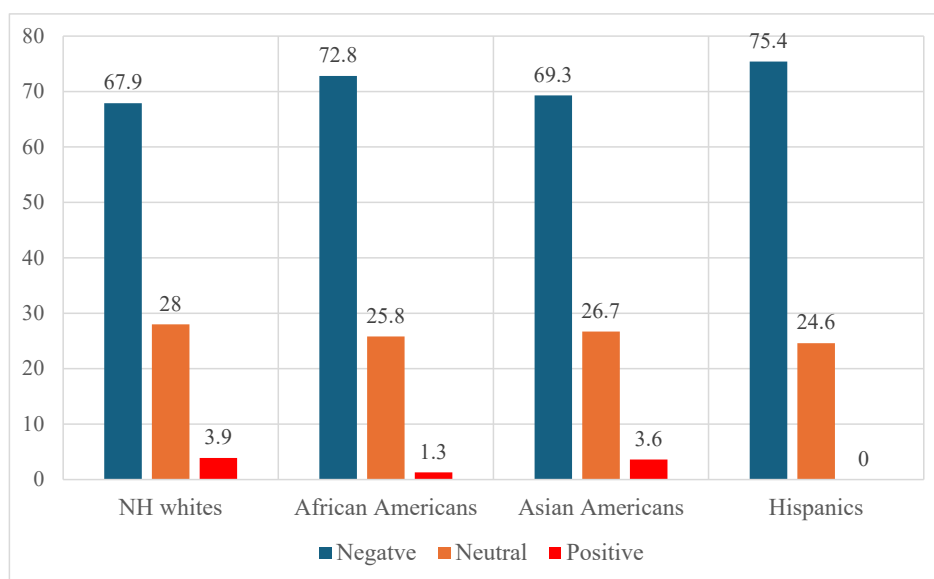


Figure 4. Sentiments associated with the conversations on symptoms of severe depression, grouped by race/ethnicity – (Percentages of conversations).

Mindsets

Conversations were analyzed as avoiding, which means denying or having trouble with one's depressive mindset, and confronting, which means overcoming and empowering one's depressive mindset. NHw in conversations related to severe depression symptoms tended to have a troubled mindset (44%) instead of an empowerment mindset (31%). Most African Americans had an avoidance attitude (70%) and displayed a troubled mindset in 52% of conversations. They felt the burden but did not act, unless it was out of necessity to get better. They had an inability to feel hopeful about the future. Conversations in Asian Americans were balanced between a confronting mindset (51%) and an avoiding mindset (49%). Confrontation was encouraged by an empowered mindset (40%) that led them to take action to overcome the situation. In contrast, the denial mindset was usually driving their avoiding attitude (26%). Hispanics had an avoidant approach to the condition at the highest rate of the four groups studied (84%). They carried the burden without acting and tended to deny it at a high rate (35%). In summary, Hispanics/Latinos are far less likely to talk about severe depression symptoms than the other groups. **Table 3** summarizes the mindset and provided percentages grouped by race/ethnicity.

Path to treatment

NHw talked more about diagnosis, making their conversations along the path to treatment more balanced out between the stages. They focused more on the early stages than they did on treatment and ongoing management. As they move forward in the path to treatment, they develop a more confronting approach, especially during the treatment stage, and this approach remains high during the ongoing management stage. Their proactive approach peaked at the treatment stage, and they were more proactive than any other racial/ethnic group during the concern and seeking information stages.

African Americans talked less about treatment than NHw and Asian Americans and slightly less about diagnosis than NHw and Asian Americans (**Table 2**). However, they talked about treatment at the same rate as Hispanics and talked about diagnosis more than Hispanics. They started the path to treatment with a highly proactive approach but also

displayed a high level of denial. This denial decreased as they moved across the path, becoming very low by the last two stages, where they were more empowered. They had a significantly troubled mindset throughout the whole path, which intensified in the concerned and diagnosis stages, where they also had a more reactive approach. During the ongoing management stage, they were confronting but also very reactive.

Asian Americans went online to discuss the path to treatment right at the time they felt symptoms and continued doing so up until the treatment stage. During the ongoing management, they discussed online at the same rate as NHw and African Americans but more than Hispanics. They started the path to treatment attempting to make sense of their situation but also displaying troubled and denial mindsets. This denial tended to decrease as they moved across the path to treatment, while the troubled mindset remained steady with a slight peak at the stage concerned. As they moved forward in the path, they developed some capacity for confrontation, especially in the late stages, where they put a lot of effort into feeling in control of their situation.

Hispanics engaged in conversations about severe depressive symptoms had a path to treatment similar to that of Hispanics with milder depression. The number of conversations, however, were more balanced across stages, having 1.6 times as many conversations in the stages of concern, seeking information and ongoing management. Hispanics also started the path to treatment trying to make sense of what happened to them, but all other groups were still more confronting than they were at the early stages. Their confronting approach didn't increase until the treatment stage. There was a considerable gap in the confronting approach between Hispanics (only 3%), probably because Hispanics were more in denial and unable to act. Sixty-four percent of the conversations between Hispanics had a proactive approach to the first stage. This dropped to 34% and 36% by the second and third stages. In treatment, however, 74% maintained a proactive approach. This fell back to 38% for those in the final stage of managing their condition. Overall, conversations between Hispanics maintained a proactive approach at about the same rate as Asian Americans and African Americans in the first, second, third, and fifth stages. They maintained a proactive approach about 7% less often than NHw at every stage except the first. Conversations

Table 3. Conversations about severe depression: mindset- grouped by race/ethnicity – (provided in %).

Race/Ethnicity		NH whites	African Americans	Asian Americans	Hispanics	
Avoiding vs Confronting		59% vs.41%	70% vs. 30%	49% vs.51%	84% vs.16%	
Mindset	Avoiding	Denial	15%	18%	26%	35%
		Trouble	44%	52%	23%	49%
	Confronting	Overcome	10%	3%	11%	0%
		Empower	31%	27%	40%	16%

were also examined to determine the mindsets along the path to treatment. The five stages of the path to treatment are as follows: the occurrence or persistence of symptoms, concern or suspicion that the individual may have symptoms of severe depression, seeking information or a diagnosis, treatment, and ongoing coping and depression management.

Discussion

Race/ethnicity is an important variable in chronic disease evaluation, but it remains a poorly understood concept. This is the first study that uses Big Data to analyze thousands of online digital conversations about severe depression across racial and ethnic groups. These digital environments differ from the usual clinical or community setting of depression research, so this study provides a different perspective than clinical research. Digital platforms for people living with severe depression offer anonymity, peer support, and the opportunity to interact in a judgment-free environment that is seemingly free of real-life consequences; in this way, people could express their thoughts and mindsets openly. This AI tool grants researchers the ability to analyze large amounts of data beyond human capability, leading to a more normal distribution of various results for English-speaking persons from the US. This is the major contribution of this study to literature. Hispanics have the highest negative tone compared to the other ethnic groups. This aligns with our previously published studies highlighting the internal stigma surrounding depressive symptoms within the Hispanic community [14-16]. While our study is enlightening, a more in-depth analysis considering gender, race, SES variations in depression could be crucial. External barriers to seeking treatment for depression included stigma and lack of resources. Stigma was particularly relevant for Hispanics and African Americans, aligning with previous studies that demonstrated how depression symptoms are perceived among members of these racial/ethnic groups [17]. Although Black-White differences between depressive symptoms and hopelessness (with Black persons showing a weaker association compared to their White counterparts) have been reported [18], our study didn't include the possible roles that attitudes about self, others, and future play on severe depression across various ethnic groups. Additional racial/ethnic-specific findings warrant further discussion. In accordance with previous research [19], evidence from the present study supports the African Americans' health paradox (defined as better subjective health of African American population, despite their worse objective health and other adversities). This phenomenon reflects the resilience of African American populations, particularly older adults who have a high number of chronic medical condition. Various scholars have attributed this observation to the growth and flourishing in the presence of adversity [19]. Accordingly, our study supports existing literature suggesting that Black individuals experiencing depression may retain higher levels

of positive emotions, such as hope, compared to their White counterparts [6]. Further studies investigating whether social media discussions by Black individuals on depression also involve words associated with positive emotions, such as joy and hope are warranted. A model to understand neurological and cultural interaction proposed by Kitayama, *et al.* [20] may address why terms related to severe depression appear more social among Black individuals. This conceptual framework hypothesizes that the brain serves as a crucial site that accumulates effects of cultural experience, insofar as neural connectivity is likely modified through sustained engagement in cultural practices. The model is supported by evidence regarding (a) collective-level factors involved in both production and adoption of cultural values and practices and (b) neural changes that result from engagement in cultural practices. According to Qin *et al.*, [21], compared with the NHw, social terms were more used during conversations about symptoms of severe depression by African American people. This can be explained by the "resource mobilization framework," which suggests that individuals who are more negatively affected by discrimination (more severe depressive symptoms) are more likely to reach out to friends and family to cope with discrimination. Moreover, the association between discrimination and depressive symptoms over time was substantially stronger among respondents who reported high levels of support from family. Further validation of the proxies used to study various groups handling severe depression symptoms is needed. The AI tool requires language expansion to better serve a diverse population. Such analytical power can also be applied to measure differences where public health interventions occur by using IP addresses. It is also worth exploring the benefits of these digital discussions, despite their indication of ongoing barriers and stigmas. Like previous work from Assari *et al.* [22] indicating that race alters how depression is linked to changes in evaluation of self over time, our study demonstrates differences in the topics generated, sentiments associated, and path to treatment between Hispanics and non-Hispanics and supports prior work done with this AI. In this study we cannot deny that the "Diminished Gain" phenomenon can influence our results. "Diminished Gain" is a phenomenon wherein the health effects of certain socioeconomic resources and psychological assets are systematically smaller for Blacks compared to NHw. These patterns are robust, with similar findings across different resources, assets, outcomes, settings, cohorts, and age groups [23]. Finally, due to the complexity around the construct of "race," we want to mention that determining the exact terminology, and genetic bases of racial/ethnic differences falls outside the scope of this paper. Excluding the word "race" will be a difficult transition, requiring substantial resources to re-educate the population, and commitment from the stakeholders involved to model the change by omitting the word "race" from their vocabulary [24].

Limitations of the study

The novelty of this technology leads to a few limitations. The conversations could only be gathered from participants using the English language, leaving a substantial portion of Hispanic or Latin-identifying populations within the US out of the analysis. Since only public conversations were used, any protected conversations behind firewalls or in private profiles were not included. Due to the nature of digital conversations, we used a proxy for symptoms of severe depression rather than a validated instrument such as the nine-item Patient Health Questionnaire (PHQ-9) [2]. The nature of social media data and NLP analysis introduces bias to our findings, which could impact the reliability of our outcomes. Social media content is unlikely to represent the broader population due to demographic biases in technology uptake, barriers to access, and regional social media platform preferences [25]. A proxy also exists between every user and their digital representation of themselves, a quandary that affects every Internet-related data mining project. The AI could not relay gender-related data, residency, or generational status either. Although experts on language expression were used to train this program, we cannot account for biases written into AI programming or that all data points were from anonymous voices. The conversations of each Racial/Ethnic group were also not of same sample size, so the results may be biased because of higher numbers of conversations in NHw than other individuals.

Since the study primarily employs the terms “depression” plus the severe adjectives for search queries, alternative inclusive expressions such as “nervous breakdown”, “ataque de nervios,” and other diverse terms around mental health within the communities of color were not included.

Finally, considering the complex interactions between severity of depression with variables like gender, race, and socioeconomic status (SES), a more in-depth analysis considering the interaction between gender and race and the interaction between gender and SES variations in severe depression is warranted. Notwithstanding these limitations, the present study holds significant implications and strengths. It should be emphasized that this may be one of the few studies examining the digital conversations related with severe depression symptoms across four major racial/ethnic groups in the U.S., especially using a very large, data set.

Conclusions

Utilizing the digital conversations made by individuals during routine daily lives that are available in the public domain, our study reveals that conversations regarding symptoms of severe depression are different across racial and ethnic groups. As this research suggests, variations exist in all facets of the conversations: topics generated, sentiments associated, mindset, and path to treatment. Notable differences found

in this study were that Hispanics were less proactive, were more negative, were more symptomatic, and showed less involvement in treatment when compared with the conversations of other individuals of other racial/ethnic groups. Likewise, this study supports that depression is perceived as a more social phenomenon in digital conversations among African Americans when they are compared with other ethnic/racial groups.

Our study supports earlier investigations by Markus and Kitayama about the influence of cultural factors in differences in self-construal between individuals [26]. Unfortunately, racial and ethnic inequalities exist in relation to depression diagnosis and treatment within the United States. The recent Surgeon General’s report – Mental Health: Culture, Race, and Ethnicity [27]– discussed the issues of racial and ethnic disparities in mental health treatment, describing that racial/ethnic minority populations have less access to mental health treatment, greater need for services, and receive a lower quality of care when treated [28].

Future research may explore how differential hopelessness may contribute to ethnic variation in causes, courses, and consequences of severe depression across ethnic groups, which have implications for the elimination of health disparities. Our findings highlight opportunities for culturally competent and targeted approaches to address areas amenable to change that could impact the ability of people to seek and receive mental health support. Given the severely debilitating nature of the disease, clinicians need to be cognizant of these disparities to better serve communities. When faced with the notion that ethnic minorities are less likely to seek mental health care than their NHw counterparts, it is also important to consider various social constructs ingrained in the culture that may limit approaches and treatment modalities when addressing depression in diverse populations. Future studies identifying ethnic/racial variations in severe depression symptoms may help to improve equity in mental health care. This knowledge should particularly be used for formulating strategies to engage diverse communities in the awareness of severe depressive disorders.

Acknowledgements

The authors would like to express sincere gratitude to Dr. Tatiana Falcone and Dinorah Villanueva, for providing valuable insights and critical feedback during the study design and manuscript preparation, as well as Liliana Gil Valleta, for facilitating access to the Culturintel database used in this investigation. We also thank Therese Deiseroth for assistance with the manuscript editorial review. This research was supported by a grant from the National Center for Advancing Translational Sciences of the National Institutes of Health, under Award Number UL1TR001425. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

1. National Institute of Mental Health (NIMH). Major Depression; [updated July 2023; cited 2023 Aug 23]. Available from: <https://www.nimh.nih.gov/health/statistics/major-depression>
2. NCHS Data Brief No. 379 September 2020 <https://www.cdc.gov/nchs/data/databriefs/db379-H.pdf>. Accessed on Jan 20 2023.
3. Villarroel MA, Terlizzi EP. Symptoms of depression among adults: United States, 2019. NCHS Data Brief, no 379. Hyattsville, MD: National Center for Health Statistics. 2020. <https://www.cdc.gov/nchs/products/databriefs/db379.htm#ref2>
4. LaVeist TA. Beyond dummy variables and sample selection: what health services researchers ought to know about race as a variable. *Health Serv Res*. 1994 Apr;29(1):1-16.
5. Simon GE, Coleman KJ, Waitzfelder BE, Beck A, Rossom RC, Stewart C, et al. Adjusting Antidepressant Quality Measures for Race and Ethnicity. *JAMA Psychiatry*. 2015 Oct;72(10):1055-6.
6. Bailey RK, Mokongho J, Kumar A. Racial and ethnic differences in depression: current perspectives. *Neuropsychiatr Dis Treat*. 2019 Feb 22;15:603-9.
7. Castilla-Puentes R, Dagar A, Villanueva D, Jimenez-Parrado L, Valleta LG, Falcone T. Digital conversations about depression among Hispanics and non-Hispanics in the US: a big-data, machine learning analysis identifies specific characteristics of depression narratives in Hispanics. *Ann Gen Psychiatry*. 2021 Nov 29;20(1):50.
8. González HM, Tarraf W, Whitfield KE, Vega WA. The epidemiology of major depression and ethnicity in the United States. *J Psychiatr Res*. 2010 Nov;44(15):1043-51.
9. De Choudhury M, Counts S, Horvitz E. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th annual ACM web science conference 2013* May 2 (pp. 47-56).
10. Kolajo T, Daramola O, Adebisi A. Big data stream analysis: a systematic literature review. *J Big Data*. 2019;6(47):1-30.
11. Park M, Cha C, Cha M. Depressive moods of users portrayed in twitter. 2012. p. 1-8.
12. Aiello AE, Renson A, Zivich PN. Social Media- and Internet-Based Disease Surveillance for Public Health. *Annu Rev Public Health*. 2020 Apr 2;41:101-18.
13. Falcone T, Dagar A, Castilla-Puentes RC, Anand A, Brethenoux C, Valleta LG, et al. Digital conversations about suicide among teenagers and adults with epilepsy: A big-data, machine learning analysis. *Epilepsia*. 2020 May;61(5):951-8.
14. Castilla-Puentes R, Pesa J, Brethenoux C, Furey P, Gil Valletta L, Falcone T. Applying the Health Belief Model to Characterize Racial/Ethnic Differences in Digital Conversations Related to Depression Pre- and Mid-COVID-19: Descriptive Analysis. *JMIR Form Res*. 2022 Jun 20;6(6):e33637.
15. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative research in psychology*. 2006 Jan 1;3(2):77-101.
16. Misra S, Jackson VW, Chong J, Choe K, Tay C, Wong J, et al. Systematic Review of Cultural Aspects of Stigma and Mental Illness among Racial and Ethnic Minority Groups in the United States: Implications for Interventions. *Am J Community Psychol*. 2021 Dec;68(3-4):486-512.
17. Alegría M, Chatterji P, Wells K, Cao Z, Chen CN, Takeuchi D, et al. Disparity in depression treatment among racial and ethnic minority populations in the United States. *Psychiatr Serv*. 2008 Nov;59(11):1264-72.
18. Assari S, Lankarani MM. Depressive Symptoms Are Associated with More Hopelessness among White than Black Older Adults. *Front Public Health*. 2016 May 4;4:82.
19. Assari S, Najand B, Bazargan M. Ethnic variation in the association between objective and subjective health in older adults. *International Journal of Travel Medicine and Global Health*. 2022 Sep 1;10(3):127-33.
20. Kitayama S, Uskul AK. Culture, mind, and the brain: current evidence and future directions. *Annu Rev Psychol*. 2011;62:419-49.
21. Qin W, Nguyen AW, Mouzon DM, Hamler TC, Wang F. Social Support, Everyday Discrimination, and Depressive Symptoms Among Older African Americans: A Longitudinal Study. *Innov Aging*. 2020 Aug 10;4(5):igaa032.
22. Assari S, Lankarani MM. Reciprocal Associations between Depressive Symptoms and Mastery among Older Adults; Black-White Differences. *Front Aging Neurosci*. 2017 Jan 5;8:279.
23. Assari S. Health disparities due to diminished return among black Americans: Public policy solutions. *Social Issues and Policy Review*. 2018 Jan;12(1):112-45.
24. Schaare D, Abenavoli L, Boccuto L. Race: How the Post-Genomic Era Has Unmasked a Misconception Promoted by Healthcare. *Medicina (Kaunas)*. 2023 Apr 28;59(5):861.
25. Kim G, Sellbom M, Ford KL. Race/ethnicity and measurement equivalence of the Everyday Discrimination Scale. *Psychol Assess*. 2014 Sep;26(3):892-900.
26. Markus HR, Kitayama S. Cultures and Selves: A Cycle of Mutual Constitution. *Perspect Psychol Sci*. 2010 Jul;5(4):420-30.
27. Office of the Surgeon General (US); Center for Mental Health Services (US); National Institute of Mental Health (US). *Mental Health: Culture, Race, and Ethnicity: A Supplement to Mental Health: A Report of the Surgeon General*. Rockville (MD): Substance Abuse and Mental Health Services Administration (US); 2001 Aug. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK44243/>
28. Shim RS, Ye J, Baltrus P, Fry-Johnson Y, Daniels E, Rust G. Racial/ethnic disparities, social support, and depression: examining a social determinant of mental health. *Ethn Dis*. 2012 Winter;22(1):15-20.