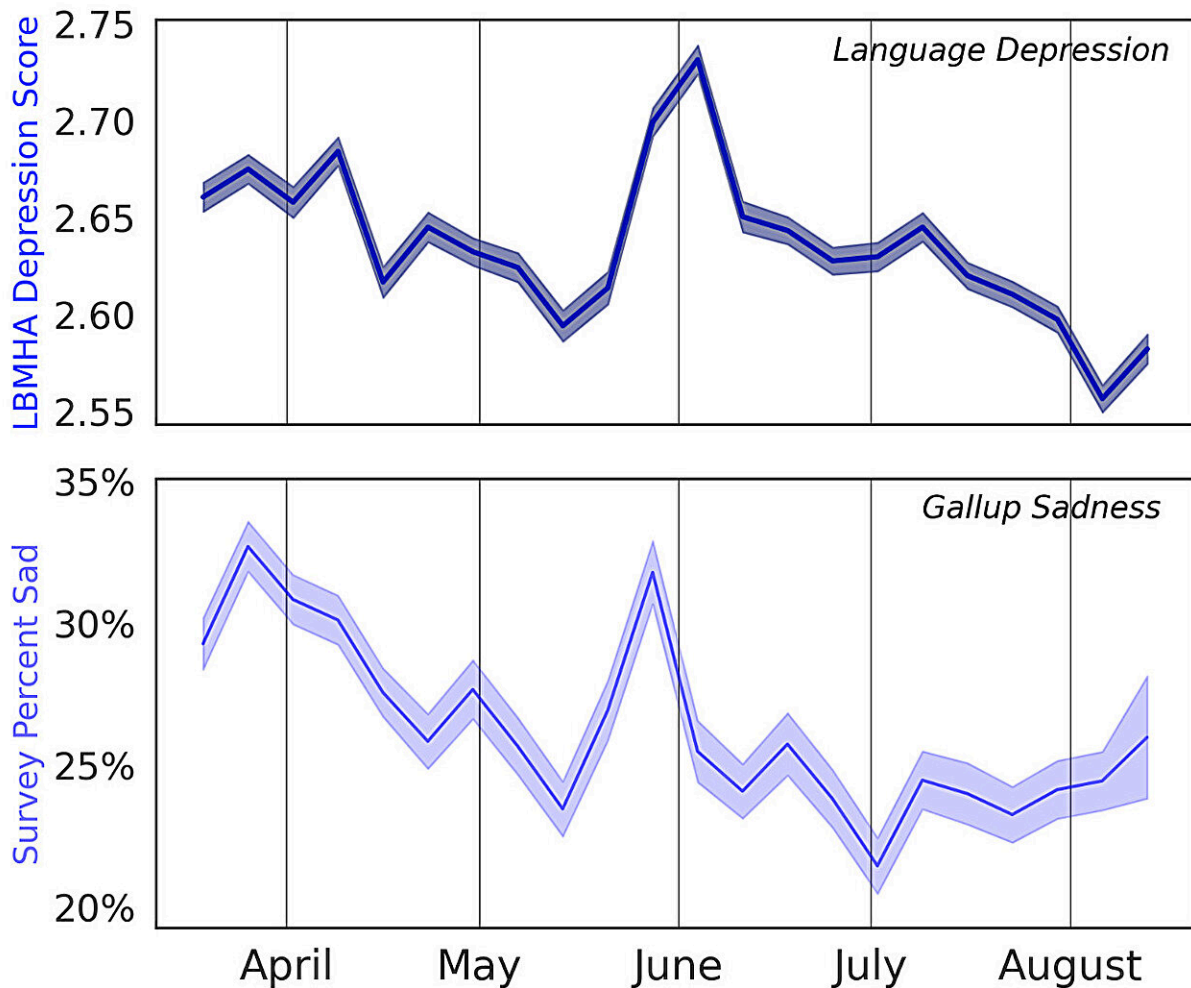# Using AI and social media to track depression in communities could offer more reliable assessments than surveys

May 7 2024

Weekly depression in the U.S. for 2020 as measured using LBMHA, and compared to Gallup polling for sadness during the same period of time. The

central lines represent the average. The higher the line, the more depression or sadness, respectively. Credit: Sid Mangalik

A study that used artificial intelligence (AI) and social media posts to assess the rates of depression and anxiety in nearly half of American counties found that the AI-generated measurements produced more reliable assessments than population surveys.

Led by Stony Brook University researchers in collaboration with computer scientists and psychologists at Stanford University and the University of Pennsylvania, the study assessed rates at a weekly level and enabled the team also to track changes for smaller regions. The findings are published in *npj Digital Medicine*.

Depression and anxiety are the two leading mental health conditions in society. According to the National Center for Health Statistics (NCHS) and Census Bureau, 10.8 percent of American adults suffered from anxiety or depression in 2019. Mental health professionals say that poor mental health plays a central role in recent increases in suicide rates and opioid-related deaths.

Typically, to measure the health of populations, expensive phone surveys ask people if they experienced "sadness" or "worry." Yet, according to the study, such reports rarely have enough data to track population changes in local communities.

The study incorporated one of the largest datasets of depression and anxiety measurements that has been used publicly. The term for these mental health measurements is Language-based mental health assessments (LBMHAs), a new AI system for measuring community-level mental health using social media language. With this system, the

researchers analyzed nearly one billion tweets from over two million users living in 1,418 U.S. counties from all 50 states.

According to Senior Author H. Andrew Schwartz, Ph.D., Associate Professor of Computer Science at Stony Brook University, LBMHAs had more reliability than the biggest public polls completed for public well-being. LBMHAs also indicate a lot of external validity in that they predict other community measurements commonly associated with mental health, such as mortality rates, better than surveys. The AI-produced scores were also more predictive of other social, economic, and political variables.

Schwartz, along with Lead Author Siddharth Mangalik, a Ph.D. student in Computer Science at Stony Brook, and Johannes C. Eichstaedt, Assistant Professor in Psychology and Human-Centered AI at Stanford, created the LBMHAs.

The LBMHAs system is the culmination of nearly a decade of work creating robust mental health assessments. This includes geo-locating Twitter/X users, determining the language use patterns of users from Tweets/X posts, combining these estimates into regions, and adapting AI models that analyze language to estimate mental health to work well on Twitter/X.

This method matched rates of depression with sadness and rates of anxiety with worry, collected through representative phone surveys in 2020 with surprising accuracy.

"The main result of this study was a comparison of how well our AI model's predictions lined up with survey-based methods and how [computational methods](link) enable new resolutions of mental health studies that were previously not possible," explains Mangalik.

They found that the proposed system outperformed survey methods by 10 percentage points in correlating with external factors like education, housing, income, and socialization.

Mangalik and his co-authors recognize the difficulty of capturing signals of psychopathology through language behavior and how users present themselves on social media.

"Social media measures allow us to track depression and anxiety—in principle—in real-time. Social media platforms are constantly changing in their leadership, policies, and how researchers can access data for the common social good," says Eichstaedt.

A new tool for mental health experts, public health assessments?

Co-author Sean Clouston, Ph.D., a Professor in the Program in Public Health and in the Department of Family, Population, and Preventive Medicine in the Renaissance School of Medicine (RSOM) at Stony Brook University, points out that listening to people's speech is a better way to understand their emotional state including feelings of worry or sadness than straight survey responses.

"This study provides a new tool to describe and understand public mental health in a way that was unimaginable just five years ago," says Clouston. "We hope it can soon be used by clinicians, mental health providers, and others to help improve public mental health in the future."

The authors recommend that public health officials consider language-based assessments alongside their survey data to understand the health of communities in closer to real-time. They explain that the inclusion of observed rather than expressed psychological states sidesteps the

tendency of participants to under-report less desirable or stigmatized traits, such as the presence of mental conditions.

Schwartz says they will continue to assess this AI-generated system to verify its usefulness for years to come. As language and platforms change over time, social media assessments will likewise need to evolve and adapt to the changing landscape.

More information: Siddharth Mangalik et al, Robust language-based mental health assessments in time and space through social media, *npj Digital Medicine* (2024). DOI: 10.1038/s41746-024-01100-0

Provided by Stony Brook University