Uncovering Patient Narratives of Opioid Use and Recovery Using Large Language Models for Topic and Emotion Analysis of Social Media

MSR208

Ashwin Rai,[1] Victoria Ikoro,[2] Devika Bhandary,[3] Andre Ng[3]

[1] Thermo Fisher Scientific, Waltham, MA, USA; [2] Thermo Fisher Scientific, Ontario, Canada; [3] Thermo Fisher Scientific, London, UK

Background

- Opioid use disorder (OUD) imposes substantial clinical and societal burden, and risks evolve with prescription and illicit opioids.[1,2]
- Understanding patient perspectives is essential for value-based, patient-centered care.
- Social media forums contain large-scale, unfiltered patient narratives about withdrawal, pain management, recovery, and emotional challenges, but the data are unstructured and difficult to analyze at scale.[3,4]
- Large language models (LLMs) enable scalable extraction of topics and emotions from these narratives, providing real-world evidence to inform interventions and policy.[5]

Objectives

- **To apply LLM–based topic modeling to patient-authored opioid-related discussions from online forums, identifying major themes in lived experience.**
- **To perform LLM-based emotion analysis across narratives, characterizing emotional burdens such as fear, sadness, hope, and relief.**
- **To generate scalable, interpretable insights from unstructured patient narratives that can inform tailored interventions, patient support resources, and health policy**

Methods

- We analyzed more than 5,000 opioid-related posts and comments from the /r/opiates subreddit to identify emotional and thematic patterns in patient narratives.

Data collection

- Source: Reddit's /r/opiates forum
- Sample: >5,000 English-language posts and comments
- Inclusion: Texts related to opioid use, withdrawal, recovery, relapse, or support
- Data were de-identified and pre-processed to reduce noise.

Text embedding and topic clustering

- Each post and comment was embedded using a pre-trained LLM to capture semantic content.
- Unsupervised clustering (e.g., density-based) was applied to these embeddings to group semantically similar texts.
- The resulting clusters were manually labelled based on top keywords and representative samples, yielding 20 interpretable topic clusters.

Emotion classification

- We used a transformer-based classifier trained on the **GoEmotions** dataset to assign emotions to each post and comment.
- The model's native 28-label output was mapped to 10 emotion categories: *sadness, fear, anger, disgust, anxiety, relief, hope, gratitude, joy, neutral*.
- Neutral labels were excluded from further analysis.

Emotion aggregation and topic emotion profiling

- Within each topic cluster, we calculated the **share of posts/comments** associated with each non-neutral emotion.
- The **dominant emotion** per topic was defined as the most frequent high-confidence label.
- Emotional profiles were visualized to highlight variation across opioid-related themes.

Figure 1. LLM-based pipeline for topic discovery and emotion analysis

Abbreviation: LLM = large language model

Results

- The analysis identified 17 interpretable topic clusters, each reflecting a distinct dimension of patient experience:
  - Physiological experiences – *withdrawal & dosing, prescription tolerance, overdose*; highlighting the physical challenges and risks of opioid use.
  - Emotional hardship – *grief & family loss, chronic pain, shame & self-loathing*; capturing the psychological toll on both individuals and families.
  - Community and recovery narratives – *recovery milestones, gratitude & peer support, use cycles*; emphasizing hope, connection, and the role of online communities in sustaining recovery.

Results (cont.)

Emotions by topic cluster

- As shown in **Figure 2**, most opioid-related discussions were dominated by negative emotions, particularly anxiety (35% of all posts) and sadness (20%). Positive tones were less frequent overall, but joy (15%) and gratitude (6%) highlighted the supportive role of online communities.
- **Figure 4** illustrates how these emotions varied by topic cluster. Anxiety dominated high-risk themes such as *Overdose & Narcan Use, Withdrawal & Dosing*, and *Pill Authenticity & Safety*. Sadness defined grief-related clusters (*Grief & Family Loss, Pain, Grief & Recovery*), reflecting the depth of bereavement and suffering. In contrast, recovery and peer-support clusters (*Recovery Milestones & Hope, Gratitude & Community Praise*) surfaced more positive emotions, with gratitude, joy, and hope balancing the overall negative climate.
- These findings underscore the emotional heterogeneity of patient narratives and identify potential touchpoints for designing patient-centered support interventions in OUD.

Figure 2. Overall emotional distribution

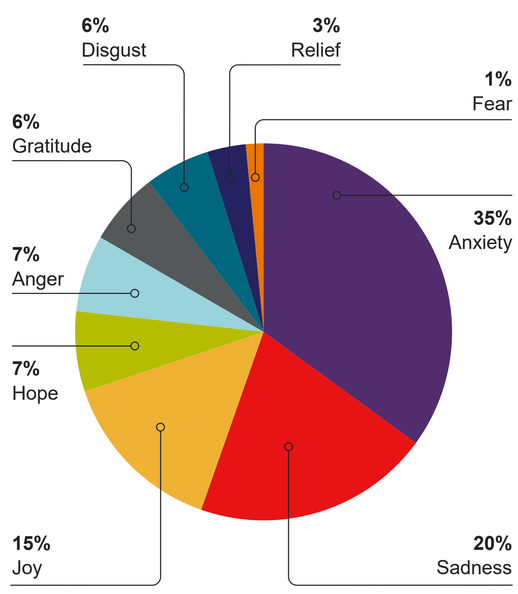
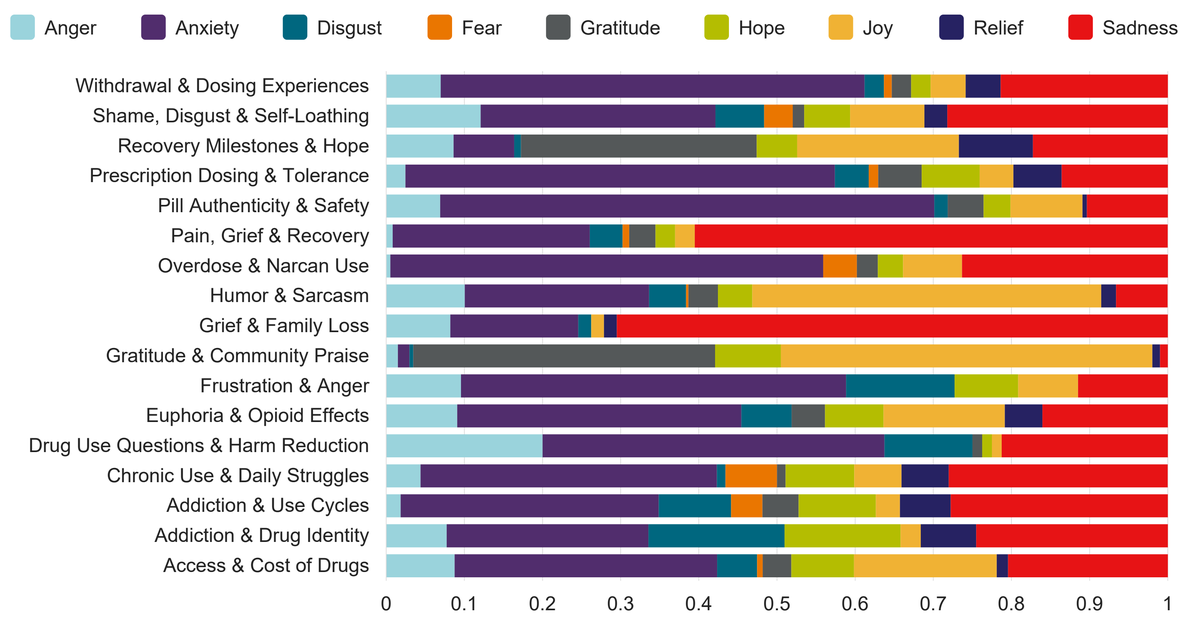


Figure 3. Illustrative sample quotes

Figure 4. Emotion distribution across 17 topic clusters



Conclusions

- **LLM-based clustering and emotion analysis successfully uncovered 17 interpretable topics in opioid-related Reddit discussions, spanning withdrawal, overdose, stigma, recovery, and support.**
- **Anxiety was the most frequent dominant emotion, especially in high-risk themes such as overdose, pill safety, and withdrawal.**
- **Sadness was prominent in grief-related topics, while gratitude and joy were most common in recovery and peer-support discussions.**
- **Emotion–topic mapping from unstructured social data offers a scalable way to surface patient-reported experiences and inform targeted interventions in opioid care.**

References

1. Volkow ND, Blanco C. *Mol Psychiatry*. 2021;26(1):218-33. **2.** Nichols VP, et al. *BMJ open*. 2020;10(2):e032988. **3.** Sarker A, et al. *J Med Internet Res*. 2017;19(10):e361. **4.** Carpenter KA, et al. *PLOS Digital Health*. 2025;4(4):e0000842. **5.** Thirunavukarasu AJ, et al. *Nat Med*. 2023;29(8):1930-40.

Disclosures

AR, VI, DB and **AN** are employees of PPD™ Evidera™ Real-World Data & Scientific Solutions, Thermo Fisher Scientific. Funding provided by Thermo Fisher Scientific.

Acknowledgments

Editorial and graphic design support was provided by Caroline Cole and Shani Berger of Thermo Fisher Scientific.