Using social listening to understand the sleep lexicon in online PWS communities

Zachary Cline[1], Amee Revana[2], Aaron Chidekel[3], Kimberly Renner[3], Kerry Smallacombe[3], Erin Cooper Carter[4], Christopher DeFelice[1], Maria Picone[1]
[1]TREND Community, Philadelphia, PA, USA [2]Texas Children's Hospital, Houston, TX, USA [3]Nemours Children's Health, Wilmington, DE, USA [4]Westside Behavioral Health, Westlake, OH, USA

#99

Background

Prader-Willi syndrome (PWS) is a rare genetic disorder that can lead to multiple health challenges that can be physical, mental, and behavioral in nature. Sleep disorders are among those possible challenges.

For people with chronic health conditions—as well as their caregivers—social media has become a tool to create a sense of community, ask questions, and share experiences.

For those caring for people with rare conditions like PWS, these social media groups are even more vital for connecting with the community. The conversations taking place in these groups present a rich source of insight into patient and caregiver experiences, which can be tapped by social listening and natural language processing.

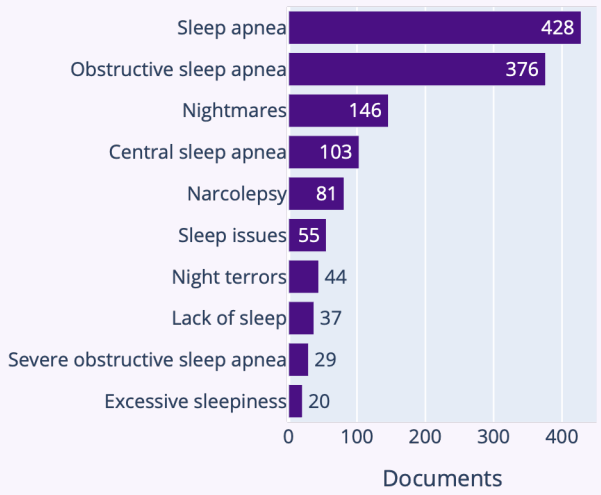
We applied these methods to conversations relating to PWS to better understand the patient/caregiver language around sleep issues and to highlight key concerns for intervention in the community.

Methods

We analyzed 226,333 documents (posts/comments) from Facebook groups devoted to PWS, mostly for caregivers of young children, and 3280 documents from reddit conversations mentioning PWS. A proprietary engine analyzed documents for mentions of clinical concepts and 14 daily living attributes, including sleeping. Documents were lemmatized with spacy, and Term Frequency-Inverse Document Frequency (TF-IDF) was used to find words/phrases associated with subsets of documents. Vector semantics was used to determine phrases semantically similar to "sleepiness" and "narcolepsy". The vectors for words/phrase were created using dimensionality reduction on a context-distribution smoothed Positive Pointwise Mutual Information (PPMI) matrix of n-grams and context words up to 7 words away. The counts of n-gram and context cooccurrences were based on the full corpus of PWS social media documents and nothing else.

Results

Figure 1: Most frequently mentioned sleep-related diseases/symptoms



The most frequently mentioned sleep-related diseases/symptoms were, in order, *Sleep apnea* (428 documents), *Obstructive sleep apnea* (376), *Nightmares* (146), *Central sleep apnea* (103), and *Narcolepsy* (81). In paragraphs about sleep that did not mention a specific disease/symptom, TF-IDF revealed the following associated lemmatized words/phrases: "sleep study", "nap", "wake", "night", "day", "sleepy", "bed", "fall asleep", "start gh", "growth hormone", "sleep [a] lot", "sleep sack", "sleep through the night", "pulse ox".

Table 1. Words/phrases most semantically similar to "narcolepsy"

Word/Phrase	Sim. Score	Word/Phrase	Sim. Score	Word/Phrase	Sim. Score
cataplexy	0.981	neurological	0.866	correlation	0.857
symptom	0.886	prader willi syndrome	0.864	willi syndrome	0.854
history	0.870	mental illness	0.860	psychological	0.851
epilepsy	0.869	described	0.859	reported	0.849
linked	0.868	genetic disorder	0.858	asd	0.848
associated	0.866	behavioral issues	0.858	condition	0.847

The most semantically similar phrases to "narcolepsy", as measured by cosine similarity, were "cataplexy", "symptom", "history", "epilepsy", "linked", "associated", and "neurological".

Table 2. Words/phrases most semantically similar to "sleepiness"

Word/Phrase	Sim. Score	Word/Phrase	Sim. Score	Word/Phrase	Sim. Score
daytime sleepiness	0.936	significant	0.833	improvement	0.818
fatigue	0.907	muscle tone	0.833	since starting	0.815
daytime	0.846	significantly	0.832	overall	0.813
alertness	0.841	improves	0.829	infancy	0.810
wakefulness	0.838	noticeable	0.827	regulation	0.810
excessive	0.834	reduction	0.827	energy levels	0.810

For "sleepiness", the most similar phrases were "daytime sleepiness", "fatigue", "daytime", "alertness", "wakefulness", "excessive", "significant", and "muscle tone".

Conclusions

Social listening revealed "sleep apnea" to be the most discussed sleep issue in these online communities for young children with PWS, with sleep studies for starting growth hormone therapy seeming to drive much of that discussion. This methodology also allowed us to identify words associated with sleep issues in the community's lexicon, which would be much more difficult to achieve through other research methods.

"Try getting in touch with PWSAUSA. We were told by our pulmonologist that the results of the sleep study were just a baseline and wouldn't prevent her from recommending growth hormone. We had obstructive and central."

"Her narcolepsy is having a far worse impact on her life than everything else. It ties in with the eating in that she can be so tired that she forgets she ate and will keep asking for more."

Title: Using social listening to understand the sleep lexicon in online PWS communities

Authors and institutions: Zachary Cline[1], Amee Revana[2], Aaron Chidekel[3], Kimberly Renner[3], Kerry Smallacombe[3], Erin Cooper Carter[4], Christopher DeFelice[1], Maria Picone[1]

[1]TREND Community, Philadelphia, PA, USA
[2]Texas Children's Hospital, Houston, TX, USA
[3]Nemours Children's Health, Wilmington, DE, USA
[4]Westside Behavioral Health, Brockton, MA, USA

Total Word Count: 350/350 (excluding title, authors and institutions)

Introduction:

Sleep disorders are among the possible health challenges people with Prader-Willi Syndrome (PWS) face. Social listening of online PWS discussion reveals patient/caregiver language around sleep issues and can highlight key concerns for intervention in this community.

Methods:

We analyzed 226,333 documents (posts/comments) from Facebook groups devoted to PWS, mostly for caregivers of young children, and 3280 documents from reddit conversations mentioning PWS. A proprietary engine analyzed documents for mentions of clinical concepts and 14 daily living attributes, including sleeping. Documents were lemmatized with spacy, and Term Frequency-Inverse Document Frequency (TF-IDF) was used to find words/phrases associated with subsets of documents. Vector semantics was used to determine phrases semantically similar to "sleepiness" and "narcolepsy".

Results:

The top sleep conditions were *Sleep apnea* (428 documents), *Obstructive sleep apnea* (376), *Nightmares* (146), *Central sleep apnea* (103), and *Narcolepsy* (81). TF-IDF revealed the following lemmatized words/phrases were associated with paragraphs about sleep that didn't mention a specific disease/symptom: "sleep study", "nap", "wake", "night", "day", "sleepy", "bed", "fall asleep", "start gh", "growth hormone", "sleep [a] lot", "sleep sack", "sleep through the night", "pulse ox". The most similar phrases to "narcolepsy" were "cataplexy", "symptom", "history", "epilepsy", "linked", "associated", and "neurological". For "sleepiness", the most similar phrases were "daytime sleepiness", "fatigue", "daytime", "alertness", "wakefulness", "excessive", "significant", and "muscle tone".

Conclusion:

This analysis of over 229,000 community posts highlights sleep apnea as the most frequently discussed sleep issue in these online caregiver communities of young

children with PWS, with sleep studies for starting growth hormone therapy seeming to drive much of that discussion. Narcolepsy also emerged as a key topic, with related terms like "cataplexy" and "daytime sleepiness" appearing in the discussions, suggesting an awareness of its potential connection to PWS. The use of social listening revealed a rich, nuanced vocabulary around sleep challenges in PWS that may be overlooked by traditional research methods, providing valuable insights into patient and caregiver experiences. By tapping into real-world conversations, social listening offers a unique opportunity to complement clinical data and better understand the lived realities of those affected by PWS, potentially guiding more targeted interventions and improving patient care.

Disclosures:

ZC is an employee of and owns stock options in TREND Community. Clients of TREND Community are pharmaceutical and biotechnology companies including, but not limited to, Horizon Therapeutics, Chiesi Global Rare Disease, Novartis, Harmony Biosciences, and Avadel Pharmaceuticals.

AR is the principal investigator for the phase 2 clinical trial for pitolisant to treat excessive daytime sleepiness in Prader–Willi syndrome (NCT04257929), sponsored by Harmony Biosciences.

AC is a consultant to BioMarin Pharmaceutical.

KR No disclosures are listed for this author.

KS No disclosures are listed for this author.

ECC No disclosures are listed for this author.

CD is an owner of and owns stock options in TREND Community. Clients of TREND Community are pharmaceutical and biotechnology companies including, but not limited to, Horizon Therapeutics, Chiesi Global Rare Disease, Novartis, Harmony Biosciences, and Avadel Pharmaceuticals.

MP is an owner of and owns stock options in TREND Community. Clients of TREND Community are pharmaceutical and biotechnology companies including, but not limited to, Horizon Therapeutics, Chiesi Global Rare Disease, Novartis, Harmony Biosciences, and Avadel Pharmaceuticals.