

Emotional Expression, Treatment Discussions, and Peer Support in Online Fibromyalgia Communities: Large-Scale Natural Language Processing Analysis

Onyinye Favour Ezike

Birmingham City University <https://orcid.org/0009-0002-9139-7113>

Fatuma Cisman

Birmingham City University <https://orcid.org/0009-0002-3587-6080>

Stephen P. Young

University of Birmingham <https://orcid.org/0000-0002-6355-3361>

Andrew Sean Wilson

andrew.wilson@bcu.ac.uk

Birmingham City University <https://orcid.org/0000-0001-7064-6681>

Research Article

Keywords: Natural language processing, topic modelling, large language models, emotion classification, social media, Fibromyalgia

Posted Date: October 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11306553/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Emotional Expression, Treatment Discussions, and Peer Support in Online Fibromyalgia Communities: Large-Scale Natural Language Processing Analysis

Onyinye Favour Ezike MSc¹
Orcid ID:0009-0002-9139-7113

Fatuma Cisman BSc¹
Orcid ID:0009-0002-3587-6080

Stephen P. Young PhD²
Orcid ID:0000-0002-6355-3361

Andrew Sean Wilson PhD^{1*}
Orcid ID:0000-0001-7064-6681

¹Department of Computer Science, Birmingham City University,
STEAMHouse,
Belmont Row, Birmingham B4 7RQ, United Kingdom

²School of Infection, Inflammation and Immunology, College of
Medicine and Health, University of Birmingham, United
Kingdom

*Author for correspondence: Andrew.Wilson@bcu.ac.uk

Keywords: Natural language processing; topic modelling; large language models; emotion
classification; social media; Fibromyalgia

ABSTRACT

Purpose: Fibromyalgia is a long-term medical condition associated with widespread pain, fatigue, diagnostic uncertainty, and persistent unmet patient needs, including feeling misunderstood, inadequately supported, and uncertain about treatment. Individuals often turn to online communities for information, validation, and peer support, generating digital health data that can provide insights into lived experience beyond traditional clinical settings.

Methods: A total of 66,611 publicly available posts published between 2013 and 2025 from an online fibromyalgia community were analysed. Three topic-modelling approaches (LDA, NMF and BERTopic) were comparatively evaluated and integrated with a custom 12-category, domain-specific emotion and communication taxonomy developed specifically for fibromyalgia and chronic-illness discourse. LLM-based classification was used to characterise thematic, emotional and peer-support patterns at scale. Medication discourse was additionally analysed for emotional associations and geographic variation.

Results: Patient discussions linked symptom burden with emotional distress and treatment experiences. Frustration, anxiety, uncertainty, and support-seeking frequently accompanied discussions of pain, healthcare experiences, and medication use. Highly active members contributed disproportionately to peer support. Ninety-six distinct medications and supplements were identified, with low-dose naltrexone the most frequently discussed treatment (3,977 mentions). Medication discussions showed distinct emotional patterns, with conventional treatments frequently associated with distress and some complementary or experimental therapies associated with hope, resilience and peer support.

Conclusions: Large-scale computational analysis of patient-generated digital health data identified patterns of fibromyalgia experience consistent with those previously observed through manual qualitative research, while enabling their systematic examination across substantially larger volumes of patient discourse. This scalable approach provides insight into patient priorities, treatment perceptions, and unmet support needs that may be less visible during conventional clinical encounters, with relevance to patient-centred digital health services and clinician–patient communication.

Introduction

Fibromyalgia is a long-term medical condition characterised by widespread musculoskeletal pain and fatigue that substantially impact physical, psychological, and social wellbeing [1], affecting approximately 2–4% of the global population [2]. Despite its prevalence, fibromyalgia remains poorly understood and is not always recognised as a legitimate clinical diagnosis [3]. Symptoms may be dismissed as psychological rather than physical, contributing to emotional distress [4]. There is consequently increasing recognition of the need for improved fibromyalgia education and awareness among healthcare professionals [5].

Pharmacological management remains important but often contentious. Standard medications, including pregabalin, duloxetine, and amitriptyline, provide partial relief for some patients but may have side effects and limited long-term efficacy [6,7]. Individuals also use online communities to share experiences of prescribed, off-label and complementary treatments [8,9]. Such discussions reflect the interplay between self-management, emotion, and peer support.

Uncertainty about both their condition and healthcare can lead patients to online communities for emotional validation, information sharing, and peer connection [10,11]. These discussions provide a rich source of information that could help inform healthcare and improve patient–clinician communication. However, the volume of patient-generated data can make manual analysis impractical. Natural language processing (NLP) enables systematic analysis of these data to identify recurring themes and emotional patterns [12, 13, 17]. Previous large-scale computational analysis of patient-generated discourse in myositis has demonstrated the potential of these methods to identify lived experiences and peer-support patterns at scale [13]. General-purpose sentiment frameworks may be insufficiently sensitive to context-specific emotional and communicative expressions in chronic illness discourse, providing a rationale for evaluating a domain-specific approach. Advances in large language models (LLMs) provide opportunities to analyse context-dependent expressions within patient narratives, including descriptions of pain and emotion. Their use for large-scale analysis of emotional, treatment and peer-support patterns within online fibromyalgia communities nevertheless remains underexplored.

This study investigates patient-generated fibromyalgia discourse using comparative topic modelling, a custom domain-specific emotion and communication taxonomy, and LLM-based classification. It examines emotional expression, treatment experiences, healthcare interactions, and peer-support dynamics within a large online community, providing insights relevant to patient-centred healthcare and digital health technologies.

Methods

Data Source and Ethical Considerations

Data were collected from fibromyalgiaforums.org, an online health community where individuals discuss symptoms, diagnosis and treatment. Publicly available English-language posts published between 2013 and 2025 were retrieved using automated web scraping (Puppeteer, Node.js). Usernames and other personally identifiable information were excluded. Ethical approval was obtained from Birmingham City University (UK; Reference: Bamansoor/#12594/sub1/Am/2025/May/CEBE FAEC). The study followed established ethical guidance for research using publicly accessible online health data [14–16].

The analytical workflow integrated thematic analysis, emotion classification, treatment discourse and peer-support patterns to characterise the lived experience of fibromyalgia from patient-generated digital health data.

Data Extraction and Pre-processing

Posts were exported as comma-separated value (CSV) files and processed using Python (v3.11) with pandas, spaCy, and scispaCy. Pre-processing followed established procedures for health-related text analysis [12,17], including lowercasing, removal of punctuation, HTML tags and URLs, tokenisation, stop-word removal, lemmatisation, and removal of duplicate or empty posts.

Biomedical entities were extracted using the scispaCy model (en_ner_bc5cdr_md) and mapped to Unified Medical Language System (UMLS) concepts to incorporate domain-specific terminology (e.g. fibro-fog, neuropathic pain, low-dose naltrexone) into subsequent analyses.

Exploratory Data Analysis

Exploratory analysis characterised forum activity, user participation, and discussion structure. Posts were classified as original posts (initiating threads), or replies that were contributions to existing discussions. Post length distributions were examined to identify long-form narrative contributions using the interquartile range (IQR) method, with posts exceeding $1.5 \times \text{IQR}$ above the third quartile (>318 words) classified as long-form.

Participation inequality was quantified using the Gini coefficient, and Simpson's Diversity Index was used to estimate the probability that two randomly selected posts were authored by different users. Super-users were defined as contributors who had published more than 100 posts. Activity patterns of the ten most active contributors, a subset of these super-users, were examined to assess their role within the community.

Geographic information was derived from self-declared user profile data where available.

Topic Modelling

Thematic patterns were identified using three topic-modelling approaches: Latent Dirichlet Allocation (LDA), Non-negative Matrix Factorisation (NMF), and BERTopic. Comparing frequency-based (LDA/NMF) and semantic embedding-based (BERTopic) approaches enabled assessment of thematic consistency across modelling frameworks.

For LDA and NMF, models were generated across topic numbers ($k = 5-50$), with topic coherence (c_v) used to guide model selection. Models were also applied to a UMLS-enriched corpus generated through biomedical entity extraction.

Two BERTopic configurations were evaluated using all-mpnet-base-v2 embeddings (bigram-only and combined unigram-bigram representations), with the combined representation selected based on improved thematic coherence.

Development of Emotion and Function Taxonomy

To capture emotional and communicative dimensions beyond simple sentiment polarity, a domain-specific taxonomy was developed. A sample of 100 posts was manually reviewed and coded using NVivo (Lumivero) to identify recurring emotional expressions and communication patterns, with 99 posts available for subsequent thematic–emotional integration following preprocessing (Supplementary Figure S1).

BERTopic-derived themes were grouped into higher-level meta-themes (e.g. symptom experience and treatment management), informing the custom 12-category taxonomy comprising emotional states (e.g. frustration, anxiety/worry and hope/optimism) and communicative functions (e.g. support-seeking, support-offering and information-sharing) (Supplementary Table S1).

Multi-label annotation captured co-occurring emotional and functional expressions. Within the 100-post manually reviewed subset, a co-occurrence matrix linked these categories with thematic meta-themes.

Evaluation of LLM-based Emotion Classification

GPT-4 (OpenAI API, accessed May 2025) was used with few-shot prompting to assign emotion and communication labels at scale. Three classification strategies were compared: (1) the custom taxonomy; (2) the general-purpose GoEmotions taxonomy; and (3) unconstrained emotion extraction.

This enabled comparison of domain-specific and general-purpose approaches and assessment of the custom taxonomy's contextual relevance for health-related communication. Agreement was assessed using label overlap and dominant-emotion consistency across approaches, with exemplar posts and predefined labels used for few-shot prompting.

A random subset of outputs was manually reviewed for alignment between model-generated labels and post content. Evaluation focused on qualitative agreement and consistency rather than predictive performance, as the aim was to characterise emotional expression at scale rather than develop a supervised classifier.

Medication Mention Extraction

Medication and supplement mentions were identified using lexicon-based extraction with regular expressions and strict word-boundary matching. A custom dictionary was developed from fibromyalgia treatment guidelines, commonly discussed off-label and complementary therapies, and corpus-derived terms.

Brand and generic names were treated as distinct entities. Frequencies were calculated across the corpus and stratified by country. Within the annotated subset, medication mentions were linked to custom taxonomy emotion labels to generate a row-normalised medication–emotion co-occurrence matrix (Supplementary Figure S2).

Potentially ambiguous terms (e.g. “iron”) were manually reviewed to confirm that mentions referred to supplements rather than unrelated uses.

Figure 1 summarises the analytical pipeline.

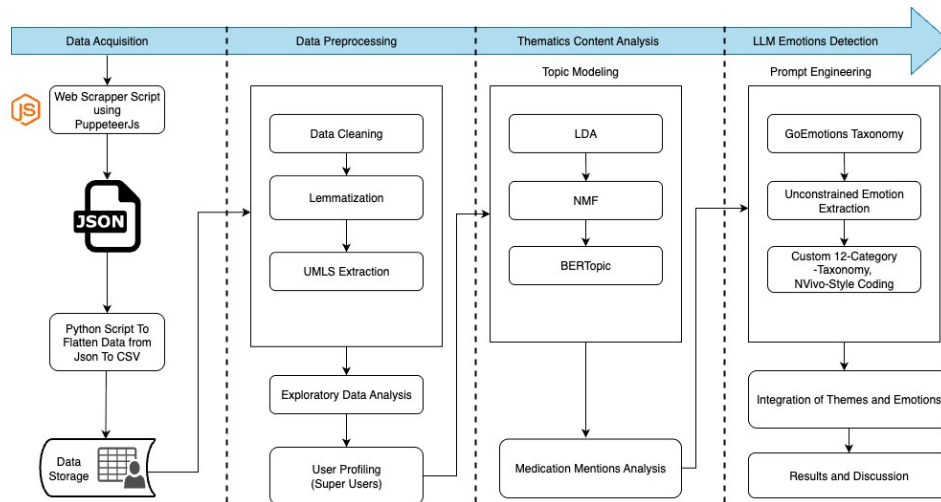


Figure 1. Workflow diagram.

Alt text: Flowchart illustrating the study workflow from data collection and preprocessing through topic modelling, emotion taxonomy development, GPT-4–based emotion classification, medication extraction, and interpretation of the findings.

Results

Community Structure and User Activity

The dataset comprised 66,611 posts, including 8,269 original discussion threads and 58,342 replies. Users were predominantly from English-speaking countries, with the United States (35,000 posts) and the United Kingdom (10,000) contributing the largest proportions. A total of 4,260 posts (6.4%) were classified as long-form (>318 words), with thematic patterns broadly consistent with the wider corpus (Supplementary Figure S3).

Participation was highly concentrated (Gini coefficient = 0.818), with 75 “super-users” (1.3%) generating 36,406 posts (54.7%) (Table 1).

Table 1. User participation table

User Activity Level	Number of Users (n, %)	Total Posts (n, %)	Posts per User Range
Super Users	75 (1.3%)	36,406 (54.7%)	101–4,008
High Activity	292 (5.1%)	11,939 (17.9%)	21–100
Moderate Activity	956 (16.8%)	9,513 (14.3%)	6–20
Low Activity	2,224 (39.0%)	6,600 (9.9%)	2–5
Single Post	2,153 (37.8%)	2,153 (3.2%)	1

Super-users primarily functioned as responders, sustaining existing conversations rather than initiating discussions. They also contributed a higher proportion of support-offering posts than the broader community (52% vs 43%) and fewer help-seeking posts (14% vs 29%), indicating a predominantly peer-support role.

Thematic Modelling Analysis

Topic modelling identified recurring themes including symptom burden, pain management, treatment experiences, healthcare navigation, peer support, and disability-related concerns.

LDA and NMF identified broad themes but produced considerable topic overlap (Supplementary Tables S2–S3). Incorporating scispaCy biomedical entity extraction and UMLS concept mapping improved clinical specificity, particularly for medication use, symptoms, and comorbid conditions (Supplementary Tables S4–S5).

The combined approach produced more coherent and clinically interpretable themes (Figure 2; Supplementary Table S6), providing the foundation for subsequent analyses of emotional expression and medication discourse.

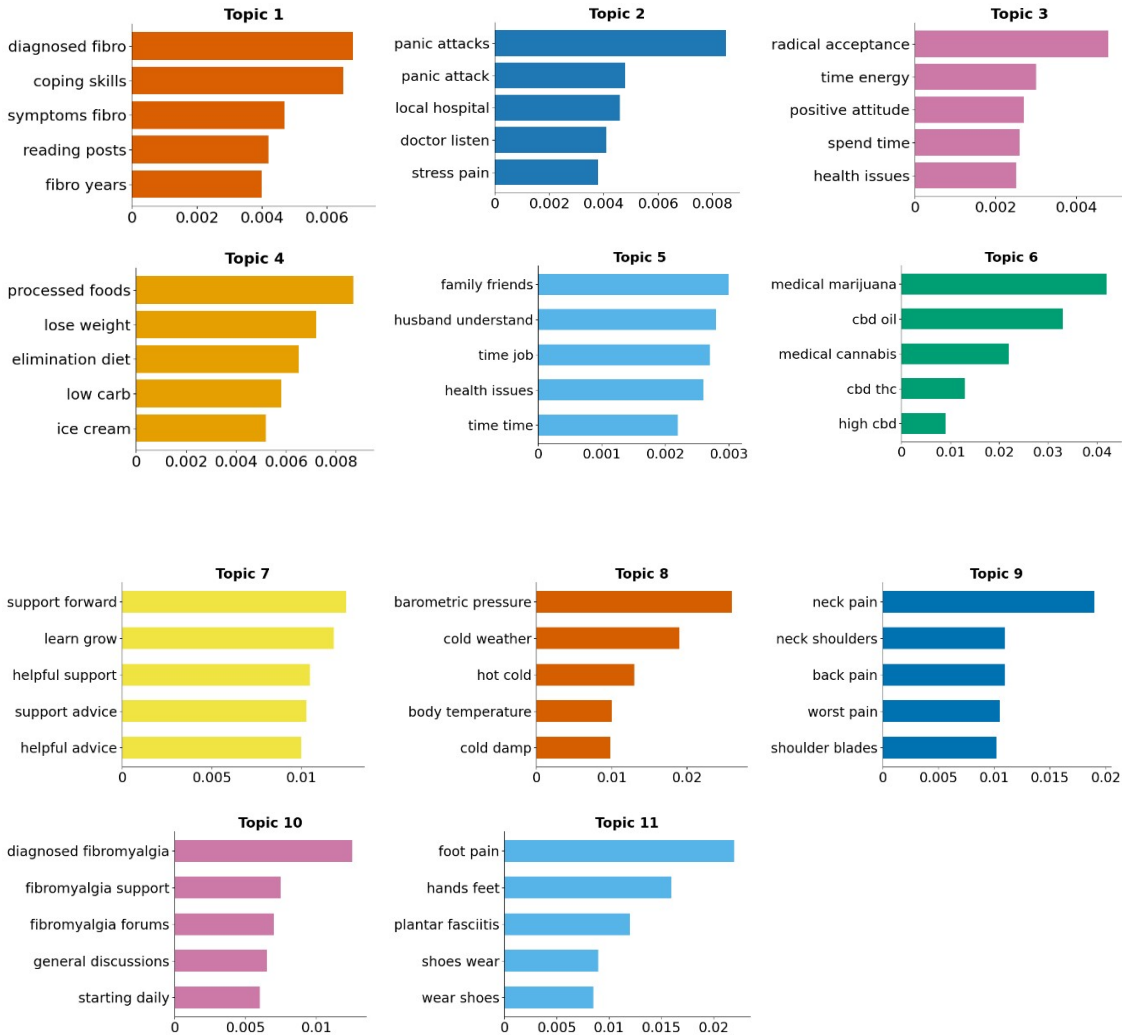


Figure 2. BERTopic top word scores for the top eleven topics.

Alt text: Bar charts showing the highest-weighted words for each BERTopic-derived topic, illustrating the principal themes discussed by individuals with fibromyalgia, including symptoms, treatments, diagnosis and peer support.

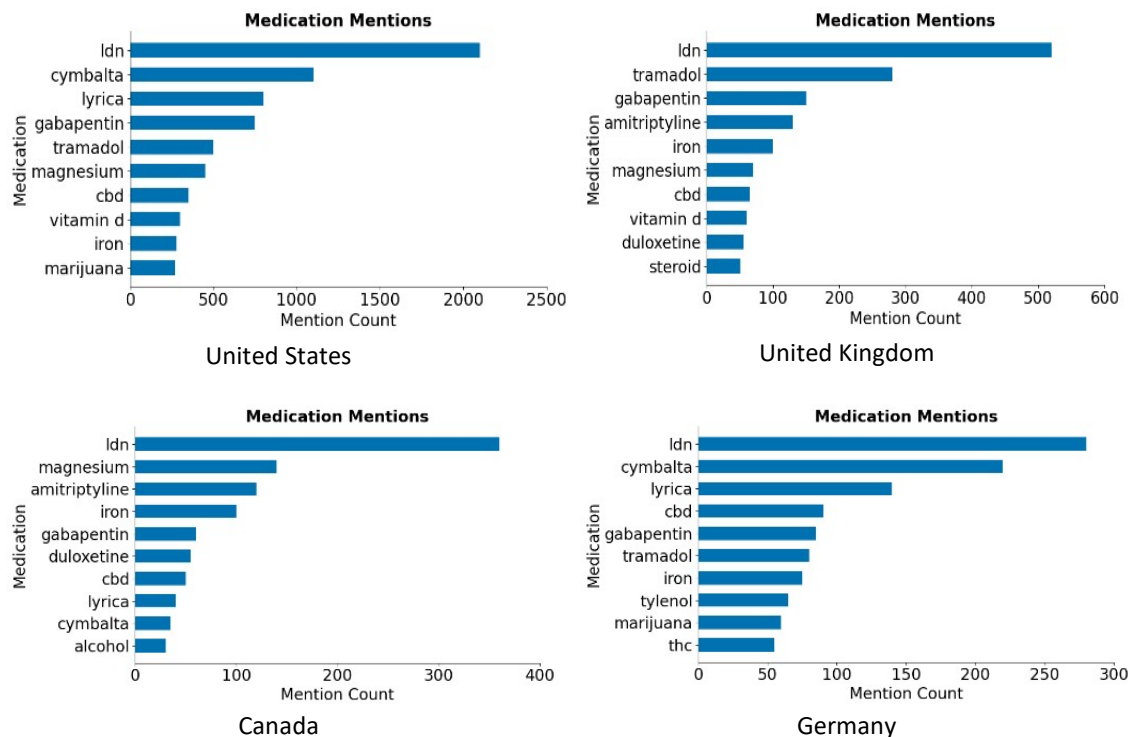
Medication Mentions and Geographic Variation

Forum discussions referenced 96 distinct medications and supplements. Low-dose naltrexone (LDN) was the most frequently mentioned treatment (3,977 mentions), followed by Cymbalta (1,531), Lyrica (1,237), Gabapentin (1,150), and Tramadol (992) (Table 4).

Table 4. Frequency of treatment and substance mention

Treatments and Substances	Number of Mentions
LDN (low-dose naltrexone)	3,977
Cymbalta	1,531
Lyrica	1,237
Gabapentin	1,150
Tramadol	992
Magnesium	911
Iron	693
CBD	673
Vitamin D	513
Amitriptyline	455
Marijuana	454
THC	446
Alcohol	412
Steroid	348
Cannabis	326
Caffeine	307
Ibuprofen	293
Tylenol	281
Savella	268
Duloxetine	251

Geographic variation was observed in treatment discussions (Figure 3). LDN was the most frequently discussed treatment in both the United States and United Kingdom. Germany showed greater prominence of magnesium and amitriptyline, while Canadian and Australian discussions showed broadly similar patterns to the United States, although with lower volumes.



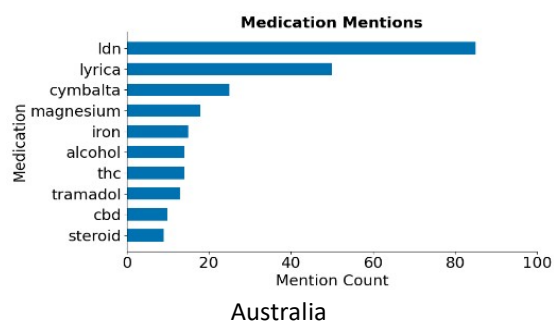


Figure 3. Most frequently mentioned treatments and substances across the five countries with the highest representation.

Alt text: Five grouped bar charts comparing the most frequently discussed medications among users from the United States, United Kingdom, Canada, Germany and Australia. Low-dose naltrexone is the most frequently discussed treatment across most countries, with regional differences in other medications.

Within the manually reviewed subset, emotion–medication associations varied across treatments. Paracetamol was associated with anxiety/worry, despair/hopelessness, and pain-related distress, whereas cannabis and caffeine were more often associated with hope/optimism, acceptance/resilience, and support-offering. Duloxetine was associated with both anxiety/worry and hope/optimism, while tramadol showed a more even distribution across emotional categories. Alcohol was associated with both anxiety/worry and acceptance/resilience.

These patterns show how medication discussions were intertwined with emotional experiences, including patients’ accounts of treatment, adverse effects, and coping with fibromyalgia.

Emotion Classification Outcomes

Comparison of the three emotion classification approaches showed differences in their characterisation of fibromyalgia-related discourse (Figure 4). GoEmotions captured broad emotional states but did not represent several dimensions identified within the domain-specific taxonomy, including isolation, despair, support-seeking and support-offering. The custom taxonomy captured these chronic-illness-specific emotional and communicative dimensions, while unconstrained classification generated a wider but less consistent range of descriptors.

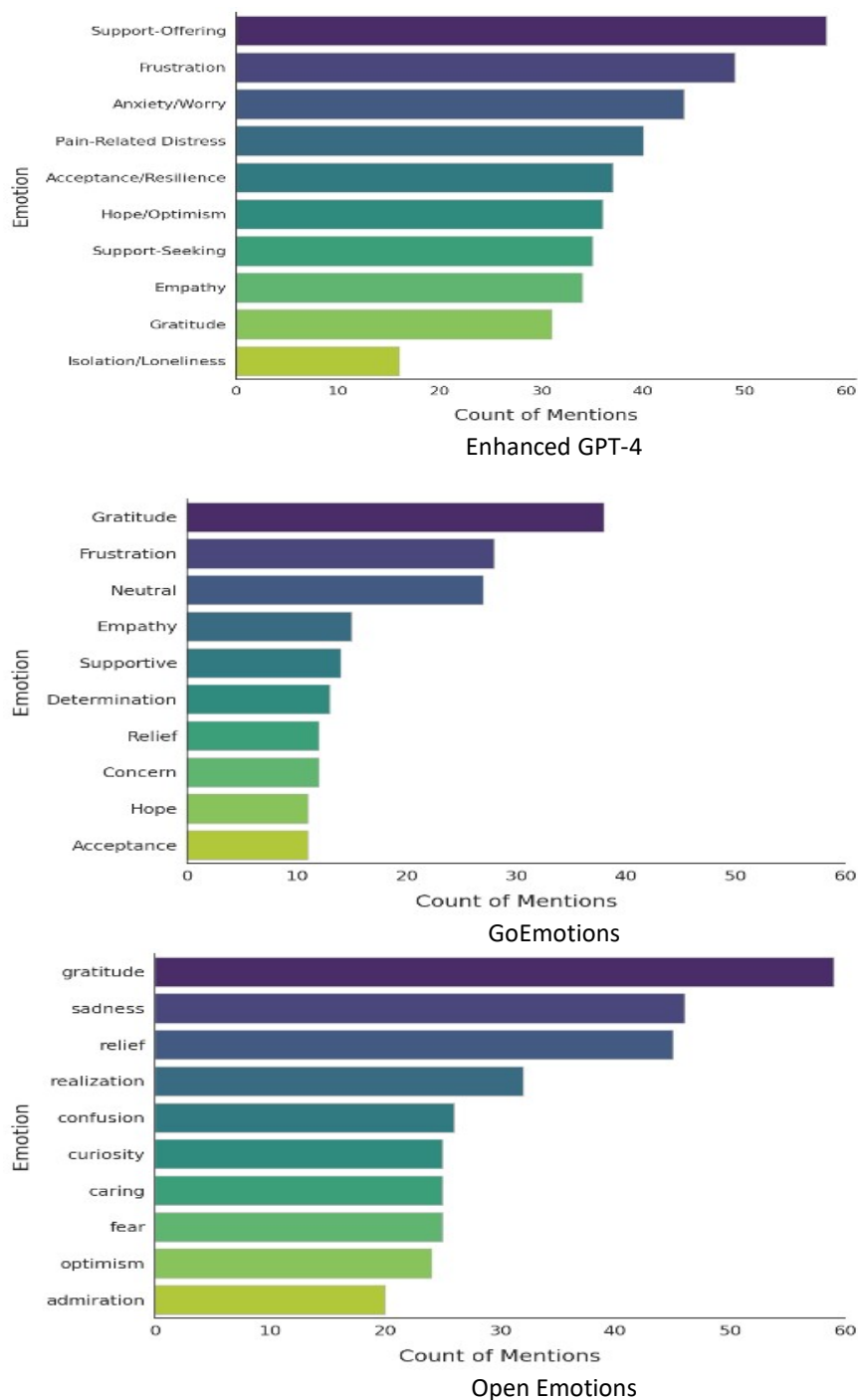


Figure 4. Comparison of top emotions from the three LLM approaches.

Alt text: Three visualisations comparing classifications produced using the custom taxonomy, the GoEmotions taxonomy, and the unconstrained GPT-4 approach.

These findings suggest that domain-specific emotion frameworks provide greater contextual relevance for analysing patient-generated health data than general-purpose emotion taxonomies.

Thematic–Emotional Integration

Distinct relationships emerged between discussion themes and emotional expression (Supplementary Figure S4). Pain, sleep disturbance, and symptom burden were most frequently associated with frustration, despair, and anxiety. In contrast, peer-support and coping discussions were characterised by empathy, support-offering, acceptance, and hope. Healthcare-related discussions showed mixed emotional profiles, including uncertainty and frustration associated with diagnosis, treatment, and ongoing management.

These findings show that emotional expression was closely linked to discussion context, with patients using the forum not only to describe symptoms but also to seek validation, exchange practical advice, and provide mutual support.

Discussion

Interpretation of Thematic and Emotional Findings

The findings show that different dimensions of fibromyalgia experience are associated with distinct emotional responses. Pain and sleep disturbance were associated with distress, while peer-support interactions showed more supportive emotional expressions. Healthcare-related discussions reflected greater uncertainty and emotional variability.

The custom 12-category taxonomy captured dimensions particularly relevant to chronic illness discourse, including support-seeking, isolation, and pain-related distress, which were less well represented by general emotion taxonomies developed from non-clinical discourse, such as GoEmotions.

The findings are consistent with previous qualitative studies based on manual analysis of fibromyalgia experiences, including diagnostic uncertainty and perceived dismissal [4,18,19], while demonstrating that similar patterns can be identified at scale in patient-generated discourse. Online communities appear to function as spaces for validation, uncertainty management, and exchanging experiential knowledge, supporting previous research on their role in information sharing and peer connection [10,11], while our findings further show that highly active users contribute disproportionately to this peer support.

Medication Discourse and Treatment Context

Conventional pharmacological therapies such as duloxetine and tramadol were frequently associated with frustration and pain-related distress, whereas discussions of complementary or experimental therapies, including cannabis and LDN, were more often associated with hope, resilience, and peer support.

The prominence of LDN, despite its off-label status, suggests a divergence between established treatment pathways and patient-driven exploration of complementary therapies.

Geographic variation in medication discussions may similarly reflect differences in prescribing practices, treatment availability, and patient preferences, although these factors cannot be determined from forum data alone.

Implications for Health Informatics

This study demonstrates how large-scale analysis of patient-generated digital health data can complement clinical evidence by revealing aspects of fibromyalgia experience that may be less visible during routine healthcare encounters. This is particularly important in fibromyalgia, where patients report diagnostic uncertainty, perceived dismissal, and difficulty having their experiences understood within healthcare [18,19]. The prominence of frustration, uncertainty, isolation, treatment concerns, and support-seeking identified through this analysis highlights patient needs that extend beyond symptom management alone.

From a clinical perspective, these findings reinforce the importance of recognising patients' lived experiences and improving communication around diagnosis, treatment expectations, and ongoing management. Systematically analysing how patients describe their condition in their own words could help identify recurring informational, emotional, and support needs that may otherwise remain poorly captured by conventional clinical data.

The findings also highlight the importance of peer networks. Highly active community members contributed disproportionately to support-offering while engaging less frequently in help-seeking, suggesting that experienced patients play an important role in sustaining online support communities. Understanding these interactions could inform the design and co-design of patient-centred digital health resources that better support information exchange, validation, and peer connection.

Building on previous work analysing patient-generated discourse in myositis [13], this study extends the approach through domain-specific emotion classification, thematic analysis, and treatment discourse. Such methods offer a scalable means of bringing patient-generated experience into health informatics and could be adapted to other long-term conditions where aspects of patient experience and unmet need are inadequately represented by traditional clinical data.

Limitations and Future Work

Several limitations should be considered when interpreting these findings. The manually reviewed subset used to develop the domain-specific taxonomy was coded by a single researcher, preventing assessment of inter-rater agreement. Furthermore, thematic–emotional integration was based on this subset and should therefore be considered exploratory. The study was designed to characterise patient-generated discourse at scale rather than develop or benchmark a supervised classifier, and large-scale manual annotation was not undertaken. Future work should validate the taxonomy and classification outputs using larger independently annotated datasets and multiple researchers, enabling assessment of inter-rater agreement and classification performance while retaining the scalability of the automated approach.

The cross-sectional analysis does not capture how patient experiences, emotional expression, or treatment discussions evolve over time. Longitudinal analysis could identify changes in patient priorities, treatment perceptions, and support needs.

Finally, the dataset comprised English-language posts from a single online fibromyalgia community with self-selected participation and may not represent all individuals living with

fibromyalgia. Replication across online communities, platforms and cultural settings would strengthen generalisability.

Future work should therefore extend validation, longitudinal analysis, and cross-platform evaluation while preserving the principal advantage of the approach: the ability to analyse patient-generated health discourse at a scale impractical through manual analysis alone.

Conclusion

This large-scale analysis of patient-generated fibromyalgia discourse identified close relationships between symptom burden, emotional expression, treatment experiences, and peer support. Pain, healthcare, and treatment discussions were frequently associated with frustration, anxiety, and uncertainty, while peer interactions provided substantial emotional and practical support.

The findings reflect patterns previously identified through manual qualitative research, while demonstrating that computational methods can identify these experiences systematically across substantially larger volumes of patient-generated data. This scalable approach provides a means of incorporating patient perspectives and unmet needs into health informatics, with potential to improve clinician–patient communication, support patient-centred care, and inform digital health technologies responsive to the lived experience of fibromyalgia.

Abbreviations

CBD: Cannabidiol
GPT: Generative Pre-trained Transformer
IQR: Interquartile range
LDN: Low-dose naltrexone
LLM: Large language model
LDA: Latent Dirichlet Allocation
NLP: Natural language processing
NMF: Non-negative matrix factorisation
POS: Part-of-speech
UMLS: Unified Medical Language System

Author Contributions

Conceptualization: Andrew Sean Wilson.
Methodology: Onyinye Ezike, Andrew Sean Wilson.
Software: Onyinye Ezike.
Investigation: Onyinye Ezike.
Data Curation: Onyinye Ezike.
Formal Analysis: Onyinye Ezike, Andrew Sean Wilson.
Validation: Andrew Sean Wilson, Stephen P. Young.
Writing – Original Draft: Onyinye Ezike.
Writing – Review & Editing: Andrew Sean Wilson, Stephen P. Young, Fatuma Cisman.
Supervision: Andrew Sean Wilson.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

The data underlying this study are derived from a publicly accessible online forum but are not publicly shared due to platform terms of service and institutional ethical restrictions. Aggregated data supporting the findings are included within the article and its supplementary materials.

Consent to Participate

Individual informed consent was not obtained because the study involved secondary analysis of publicly accessible online forum posts, without direct participant recruitment or contact.

AI Disclosure

GPT-4 (OpenAI) has been used as an analytical tool to perform emotion and communication classification of patient-generated text using structured prompting. AI-generated outputs were reviewed, interpreted, and integrated into the analyses by the authors. GPT-4 was also used to assist with language editing and improve the clarity of the manuscript. All scientific content, interpretation of the findings, and the final manuscript were reviewed and approved by the authors, who accept full responsibility for the published work.

References

- 1 Bhargava J, Goldin J. Fibromyalgia. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan–. Updated 2025 Jan 31. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK540974/> [accessed 2026-10-08].
- 2 Marques AP, Santo ASDE, Berssaneti AA, Matsutani LA, Yuan SLK. Prevalence of fibromyalgia: literature review update. *Rev Bras Reumatol Engl Ed*. 2017;57(4):356–363. doi:10.1016/j.rbre.2017.01.005
- 3 Häuser W, Fitzcharles MA. Facts and myths pertaining to fibromyalgia. *Dialogues Clin Neurosci*. 2018;20(1):53–62. doi:10.31887/DCNS.2018.20.1/whauser
- 4 Crump L, LaChapelle D. "My fibro family!" A qualitative analysis of Facebook fibromyalgia support groups' discussion content. *Can J Pain*. 2022;6(1):95–111. doi:10.1080/24740527.2022.2078183
- 5 Agarwal A, Emary PC, Gallo L, et al. Physicians' knowledge, attitudes, and practices regarding fibromyalgia: a systematic review and meta-analysis of cross-sectional studies. *Medicine (Baltimore)*. 2024;103(31):e39109. doi:10.1097/MD.00000000000039109.
- 6 Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Häuser W, Fluß E, Choy E, Kosek E, Amris K, Branco J, Dincer F, Leino-Arjas P, Longley K, McCarthy GM, Makri S, Perrot S, Sarzi-Puttini P, Taylor A, Jones GT. EULAR revised recommendations for the management of fibromyalgia. *Ann Rheum Dis*. 2017;76(2):318–328. doi:10.1136/annrheumdis-2016-209724
- 7 Moore A, Bidonde J, Fisher E, Häuser W, Bell RF, Perrot S, Makri S, Straube S. Effectiveness of pharmacological therapies for fibromyalgia syndrome in adults: an overview of Cochrane Reviews. *Rheumatology (Oxford)*. 2025;64(5):2385–2394. doi:10.1093/rheumatology/keae707
- 8 Cipolletta S, Tomaino SCM, Lo Magno E, Faccio E. Illness experiences and attitudes towards medication in online communities for people with fibromyalgia. *Int J Environ Res Public Health*. 2020;17(22):8683. doi:10.3390/ijerph17228683
- 9 Gewandter JS, Edwards RR, Hill KP, Wasan AD, Hooker JE, Lape EC, Besharat S, Cowan P, Le Foll B, Ditre JW, Freeman R. Cannabinoid therapy: attitudes and experiences of people with chronic pain. *Clin J Pain*. 2023;39(6):249–258. doi:10.1097/AJP.0000000000001109
- 10 Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. *BMC Health Serv Res*. 2016;16:442. doi:10.1186/s12913-016-1691-0
- 11 Coulson NS. How do online patient support communities affect the experience of inflammatory bowel disease? An online survey. *JRSM Short Rep*. 2013;4(8):2042533313478004. doi:10.1177/2042533313478004
- 12 Schoene AM, Basinas I, van Tongeren M, Ananiadou S. A narrative literature review of natural language processing applied to the occupational exposome. *Int J Environ Res Public Health*. 2022;19(14):8544. doi:10.3390/ijerph19148544

- 13 Wilson AS, Bin Sajid H, Gupta L. Reddit and rare diseases: what myositis communities tell us about support and struggle. *Oxford Open Digital Health*. 2026;4:oqag007. doi:10.1093/oodh/oqag007
- 14 Vayena E, Dzenowagis J, Brownstein JS, Sheikh A. Policy implications of big data in the health sector. *Bull World Health Organ*. 2018;96(1):66–68. doi:10.2471/BLT.17.197896
- 15 Ford E, Shepherd S, Jones K, Hassan L. Toward an ethical framework for the text mining of social media for health research: a systematic review. *Front Digit Health*. 2021;2:592237. doi:10.3389/fdgth.2020.592237
- 16 Harris J, Germain J, McCoy E, Schofield R. Ethical guidance for conducting health research with online communities: a scoping review of existing guidance. *PLoS One*. 2024;19(5):e0302924. doi:10.1371/journal.pone.0302924
- 17 Khanbhai M, Anyadi P, Symons J, Flott K, Darzi A, Mayer E. Applying natural language processing and machine learning techniques to patient experience feedback: a systematic review. *BMJ Health Care Inform*. 2021;28(1):e100262. doi:10.1136/bmjhci-2020-100262
- 18 Asbring P, Närvänen AL. Women's experiences of stigma in relation to chronic fatigue syndrome and fibromyalgia. *Qual Health Res*. 2002;12(2):148–160. doi:10.1177/104973230201200202
- 19 Werner A, Malterud K. It is hard work behaving as a credible patient: encounters between women with chronic pain and their doctors. *Soc Sci Med*. 2003;57(8):1409–1419. doi:10.1016/s0277-9536(02)00520-8

Supplementary Material: A Large-Scale Natural Language Processing Analysis of Emotional Expression, Treatment Discussions, and Peer Support in Online Fibromyalgia Communities

This supplementary material provides additional detail supporting the analyses presented in the main manuscript, including the emotion taxonomy, topic modelling outputs, and extended visualisations of patient discourse patterns.

Supplementary Tables

The taxonomy was developed through manual review of forum posts and captures both emotional states and communicative functions observed in patient discourse.

Supplementary Table S1. Emotion and Function Taxonomy

Category	Definition	Paraphrased Example
Pain-Related Distress	Expressions of physical suffering, discomfort, and the direct sensory experience of pain.	"The burning sensation in my legs is unbearable today, it feels like my nerves are on fire."
Frustration	Expressions of annoyance, anger, or exasperation with illness, treatments, or the healthcare system.	"I'm so tired of going to doctors who just dismiss my symptoms and tell me it's all in my head."
Anxiety/Worry	Expressions of fear, nervousness, or uncertainty about the future, symptoms, or treatment outcomes.	"I have a new symptom, and I can't stop worrying that it might be something more serious."
Despair/Hopelessness	Expressions of deep sadness, hopelessness, or the feeling that things will not get better.	"I've tried every treatment and nothing works. I feel like I'm never going to have a normal life again."
Isolation/Loneliness	Expressions of feeling alone, misunderstood, or disconnected from others due to the illness.	"No one in my life truly understands what this is like. It's such a lonely experience."
Acceptance/Resilience	Expressions of coming to terms with the condition, focusing on coping strategies, and demonstrating strength.	"I've learned to listen to my body and pace myself. It's not the life I planned, but I'm making it work."
Hope/Optimism	Expressions of positivity, belief in improvement, or encouragement for the future.	"I started a new therapy and I'm finally starting to feel a little better. There is hope."
Gratitude	Expressions of thankfulness towards the community, specific members, or for small victories.	"Thank you all so much for your support. I don't know what I'd do without this group."
Empathy	Expressions of understanding and sharing the feelings of another member.	"I'm so sorry you're going through that. I've had those same pains and I know how awful it is."
Support-Offering	The functional act of providing advice, encouragement, or emotional support to others.	"Sending you hugs. Have you tried using a heating pad? That sometimes helps me."
Support-Seeking	The functional act of asking for help, advice, information, or emotional validation.	"Has anyone else experienced this side effect? I'm not sure what to do and could use some advice."
Information-Sharing	The functional act of providing factual information, research, or personal experience in a neutral, informative tone.	"My doctor prescribed LDN. It's an off-label medication that works on the immune system."

Supplementary Table S2. Examples of LDA topics

Topic ID	0	4	6	14	16
Topic Label	Understanding Fibromyalgia Symptoms	Strategies for Managing Fibromyalgia	Managing Fibromyalgia Symptoms	Understanding Fibromyalgia Symptoms and Management	Fibromyalgia Pain Management
Topic Items	fibro pain	manage fibromyalgia	pain management	fibro symptom	hot water
	trigger point	advice manage	cider vinegar	night sleep	drink water
	back pain	live fibromyalgia	apple cider	anxiety depression	pain killer
	muscle pain	trial error	heating pad	pain med	water bottle
	mental health	follow basic	pain level	weight gain	anti-depressant

Supplementary Table S3. Examples of NMF topics

Topic ID	5	7	10	11	16
Topic Label	Exploring Diet and Health Connections	Sensory Symptoms in Fibromyalgia	Cognitive Impairment and Pain in Fibromyalgia	Managing Symptoms and Wellness Strategies	Interpreting Autoimmune Lab Results
Topic Items	low fat	rib cage	brain fog	blood pressure	lyme disease
	fat diet	numbness tingling	fibro brain	beta carotene	metabolic panel
	herxheimer reaction	skin sensation	joint pain	epsom salt	negative lyme
	fatigue syndrome	burn skin	nerve pain	methylation detox	titer pattern
	chronic fatigue	heartbeat	tender point	fizzy drink	rheumatoid factor

Supplementary Table S4. Examples of LDA topics (UMLS)

Topic ID	0	2	4	5	17
Topic Label	Managing Symptoms in Fibromyalgia	Pain and Comorbidities in Fibromyalgia	Associations Between Pain, Vitamins, and Supplements	Interplay of Pain, Fatigue, and Emotional Health	Pain Management and Comorbidities
Topic Items	pain lyrics	interstitial cystitis	hair loss	pain depression	arthritis pain
	sleep apnea	spinal stenosis	steroid pain	myofascial pain	celiac disease
	migraine headaches	pelvic pain	magnesium malate	worry fibromyalgia	lipoic acid
	arnica cream	degenerative disc	degenerative disc	pain guilt	neuropathy pain

Supplementary Table S5. Examples of NMF topics (UMLS)

Topic ID	2	4	6	14	17
Topic Label	Psychological Impact of Fibromyalgia	Diverse Pain Experience in Fibromyalgia	Medication Management for Fibromyalgia Pain	Nerve Pain Management and Associated Symptoms	Managing Pain and Side Effects in Fibromyalgia
Topic Items	anxiety depressed	knee pain	pain paracetamol	muscle spasms	pain lyrica
	anxiety worry	neck pain	cymbalta tramadol	nerve damage	cymbalta pain
	ptsd pain	chest pain	tapentadol tramadol	peripheral neuropathy	weight gain
	anxiety guilt	emotional pain	pain amitriptyline	pain amitriptyline	memory loss
	fibromyalgia panic	pelvic pain	pain oxycodone	pain osteoarthritis	dry mouth

Supplementary Table S6. Examples of BERTopic Topics

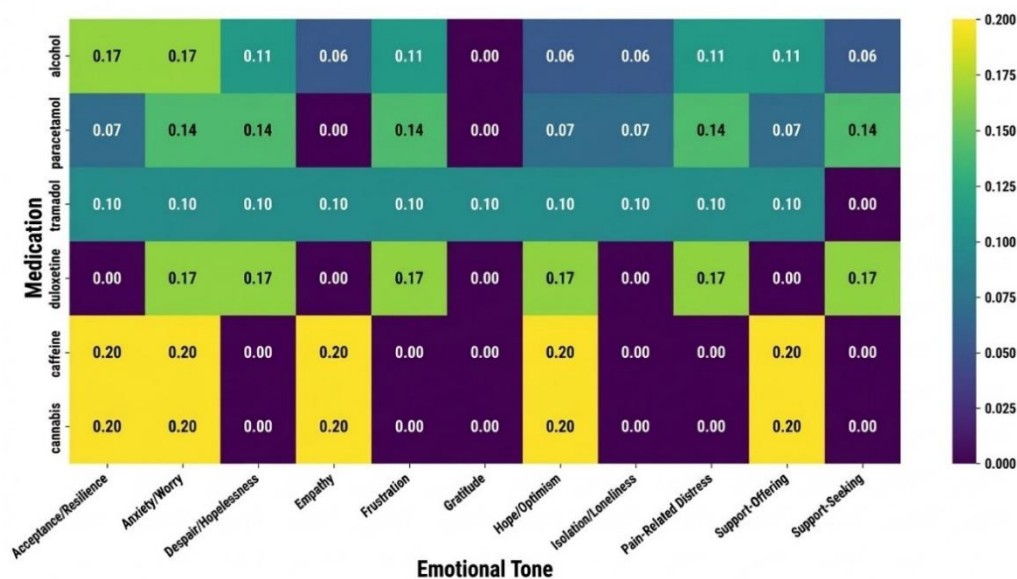
Topic ID	2	3	4	5	8
Topic Label	Dietary Strategies for Weight Management	Exploring Cannabis for Pain Relief	Sensitivity to Temperature Changes	Supportive Family Dynamics in Fibromyalgia	Navigating Social Security Disability
Topic Items	processed foods	medical marijuana	night sweats	husband understand	social security
	healthy diet	CBD oil	body temperature	family friends	back pay
	olive oil	CBD thc	cold weather	emotional support	security disability
	eat meat	CBD products	hot water	husband supportive	applied ssd
	fruits vegetables	cannabis legal	cold feet	big hugs	denied time

Supplementary Figures

Pain-Related Distress Anxiety/Worry Support-Seeking Gratitude Hope/Optimism Acceptance/Resilience Frustration Despair/Hopelessness Isolation/Loneliness Empathy Support-Offering			
Post ID	Post	Emotions Detected	
post-302809	hi everyone. this may or may not be of use, but i've been thinking a lot about some commonalities between fibro symptoms and symptoms for pppd, aka perpetual postural perceptual dizziness. i have both, and find, not real surprisingly, that they exacerbate each other. symptoms for pppd are skewed depth perception, constant recalibration of balance whenever i turn my head or stoop down, chronic fatigue, ear-ringing, the tingling sensation in muscles all over body, double vision. that's a good start. does anyone else endure these symptoms? i also have multiple chemical sensitivities, which also exhibits the same kind of	Pain-Related Distress, Frustration, Isolation/Loneliness, Empathy, Support-Seeking, Gratitude	
post-304819	symptoms, excepting that i have to leave wherever i'm being exposed. for instance, indoor pools, or even outdoors, that really load up on the chlorine. awk. our poor bodies. they try so hard to do the right thing. thank you for listening. i do hope no one else has these additional blessings! thank you so much. i have been suffering with chronic pain for over a year. i have all symptoms of this except for one. i did have a doctor who knew nothing about my suffering, so i did hundreds of hours of research. one of the things i noticed is that since i have ppsd and was dealing with this great, something traumatic happened to me, it started worse than ever. i thought i would never walk again or feel good. it made me determined to do my own research, and i did it. god bless you. i have found another doctor and will continue to do everything to get better.	Pain-Related Distress, Frustration, Despair/Hopelessness, Anxiety/Worry, Hope/Optimism, Gratitude, Support-Seeking, Acceptance/Resilience	
post-303933	hi, i am 46 years old and i was diagnosed with fibro a year ago. my pain is from the age of 20 but no one knew what i had. i also had good times and bad times. the only thing that helped me was sports and optimism. since the pandemic, the pain has worsened. i have back, shoulder, arm, elbow pain, my fingers are always stuck, especially in the morning. treatment: 300 mg gabaran 1 * 3 / day and duloxetine in the evening 2 cps. i also take a quarter of trico for sleep. i can sleep for 8 hours, but i started to wake up at night with pain. for several days i have had an intercostal stab that kills me. i took paracetamol, 60 mg arcoxia and midocalm, but the pain persists. i do not know what to do. if i try to move, the pain is greater the next day. any advice is welcome. god help me.	Pain-Related Distress, Frustration, Despair/Hopelessness, Anxiety/Worry, Hope/Optimism, Support-Seeking	

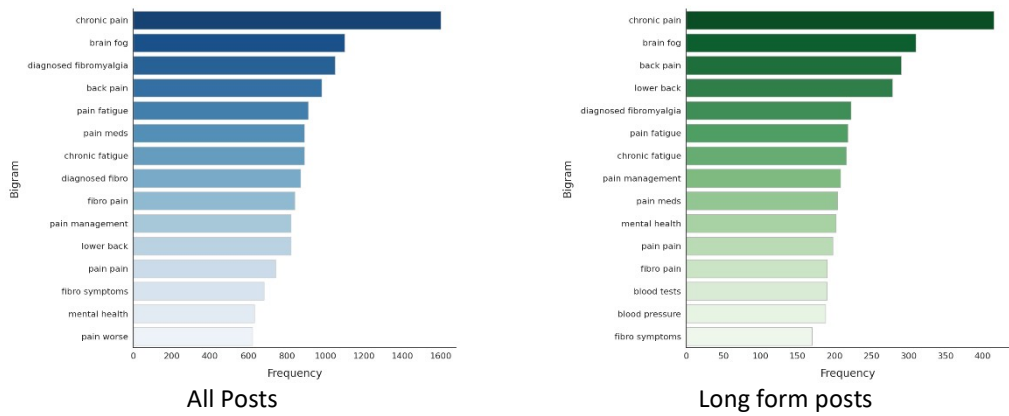
Supplementary Figure S1. Example NVivo coding of three posts

Alt text: Screenshot of three anonymised forum posts manually coded in NVivo. Coloured highlights illustrate the assignment of multiple emotion and communication categories, demonstrating how the domain-specific taxonomy was applied during manual annotation.



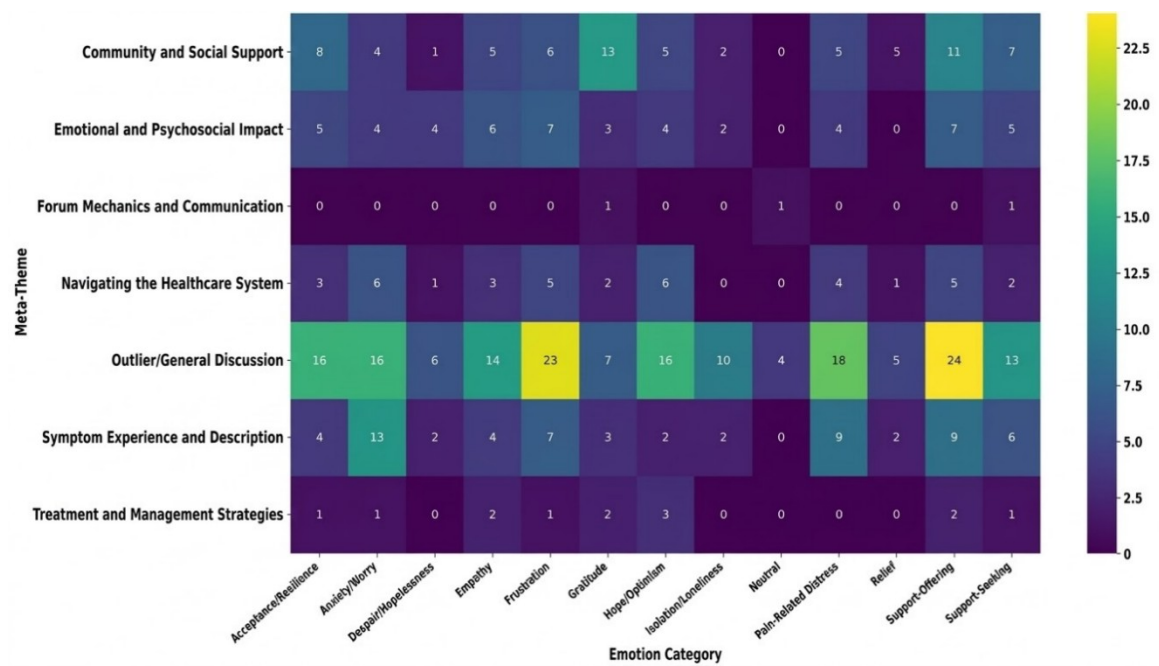
Supplementary Figure S2. Row-normalised heatmap showing the emotional distribution associated with medication mentions in the manually reviewed subset of posts (n = 100).

Alt text: Heatmap showing the distribution of emotion categories associated with medication mentions in the manually reviewed subset of posts. Colour intensity represents the relative frequency of emotion–medication associations, illustrating how different treatments are discussed in distinct emotional contexts.



Supplementary Figure S3. Comparison of the most frequent bigrams between the full corpus and long-form posts (>1.5×IQR). Thematic overlap indicates that extended posts largely reflect the same discussion topics as the broader forum corpus.

Alt text: Two horizontal bar charts comparing the most frequent bigrams in the full corpus and in long-form posts. Similar terminology across both bar charts indicates that longer posts expand on the same core discussion topics rather than introducing substantially different themes.



Supplementary Figure S4. Heatmap showing the frequency of emotion categories across thematic meta-themes derived from BERTopic modelling. The analysis was conducted on a manually reviewed subset of 100 posts ($n = 99$ following preprocessing). Values represent counts of post–emotion co-occurrences.

This figure illustrates the co-occurrence of emotional categories with thematic meta-themes, highlighting the association between pain-related discussions and negative emotional expressions, and between peer-support discussions and supportive emotional responses.

Alt text: Heatmap illustrating the co-occurrence of emotion categories and thematic meta-themes derived from BERTopic analysis. Pain- and symptom-related themes are predominantly associated with frustration and anxiety, whereas peer-support themes are predominantly associated with empathy, hope, and supportive communication.