

Original Paper

Contemporary English Pain Descriptors as Detected on Social Media Using Artificial Intelligence and Emotion Analytics Algorithms: Cross-sectional Study

Ming Yi Tan, BDS, MDS; Charlene Enhui Goh, BDS, MPH, DrPH; Hee Hon Tan, BDS, MDS

Faculty of Dentistry, National University of Singapore, Singapore, Singapore

Corresponding Author:

Ming Yi Tan, BDS, MDS

Faculty of Dentistry, National University of Singapore

9 Lower Kent Ridge Road

Singapore, 119085

Singapore

Phone: 65 67725340

Email: dentmy@nus.edu.sg

Abstract

Background: Pain description is fundamental to health care. The McGill Pain Questionnaire (MPQ) has been validated as a tool for the multidimensional measurement of pain; however, its use relies heavily on language proficiency. Although the MPQ has remained unchanged since its inception, the English language has evolved significantly since then. The advent of the internet and social media has allowed for the generation of a staggering amount of publicly available data, allowing linguistic analysis at a scale never seen before.

Objective: The aim of this study is to use social media data to examine the relevance of pain descriptors from the existing MPQ, identify novel contemporary English descriptors for pain among users of social media, and suggest a modification for a new MPQ for future validation and testing.

Methods: All posts from social media platforms from January 1, 2019, to December 31, 2019, were extracted. Artificial intelligence and emotion analytics algorithms (Crystalace and CrystalFeel) were used to measure the emotional properties of the text, including *sarcasm*, *anger*, *fear*, *sadness*, *joy*, and *valence*. Word2Vec was used to identify new pain descriptors associated with the original descriptors from the MPQ. Analysis of count and pain intensity formed the basis for proposing new pain descriptors and determining the order of pain descriptors within each subclass.

Results: A total of 118 new associated words were found via Word2Vec. Of these 118 words, 49 (41.5%) words had a count of at least 110, which corresponded to the count of the bottom 10% (8/78) of the original MPQ pain descriptors. The count and intensity of pain descriptors were used to formulate the inclusion criteria for a new pain questionnaire. For the suggested new pain questionnaire, 11 existing pain descriptors were removed, 13 new descriptors were added to existing subclasses, and a new *Psychological* subclass comprising 9 descriptors was added.

Conclusions: This study presents a novel methodology using social media data to identify new pain descriptors and can be repeated at regular intervals to ensure the relevance of pain questionnaires. The original MPQ contains several potentially outdated pain descriptors and is inadequate for reporting the psychological aspects of pain. Further research is needed to examine the reliability and validity of the revised MPQ.

(JMIR Form Res 2021;5(11):e31366) doi: [10.2196/31366](https://doi.org/10.2196/31366)

KEYWORDS

pain descriptors; social media; artificial intelligence; emotion analytics; McGill Pain Questionnaire

Introduction

Pain is “an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage” [1]. Although the experience of pain

is common, each person’s pain is unique and felt physically only by that person. Pain measurement is essential for diagnosis, monitoring of disease progression, and evaluation of treatment effectiveness [2]. This understanding and measurement of pain occur largely through verbal reporting by the person

experiencing the pain. Health care providers use this information to characterize and determine pain intensity, and in many instances, make judgments regarding treatment [3]. As verbalization of pain can sometimes be difficult, comprehensive questionnaires to elicit better descriptions of pain for diagnostic accuracy are important.

Various instruments have been developed for the assessment of pain. For acute pain, pain scales that focus on identifying pain location and intensity, such as the visual analog scale and numeric rating scale, are most commonly used [4]. Although shown to be inferior to both the visual analog scale and numeric rating scale, 4-point or 6-point verbal categorical rating scales using adjectives to describe different levels of pain have also been successfully used [5]. For children, pain scales that use images of happy and unhappy faces have been found to be appropriate [6]. All these instruments have been validated in clinical and research settings; however, their accuracy is dependent on timely recording.

Assessment of chronic pain is indisputably more complex. The long-term burden of pain plays a profound role in shaping an individual's physical and psychological state. In addition, the negative downstream effects of chronic pain can exacerbate the original pain condition through various pathways that remain poorly understood. Owing to these complexities, the above unidimensional instruments that only describe pain in terms of intensity may be too simplistic for meaningful clinical correlation.

Several authors have emphasized the need to recognize the multidimensional aspects of pain [7-9], and 6 dimensions have previously been described: physiological, sensory, affective, cognitive, behavioral, and sociocultural [7]. Health care professionals have acknowledged the varying qualities of pain and often depend on characteristic descriptions of pain to identify eventual diagnoses.

The McGill Pain Questionnaire (MPQ), created in 1971, is one of the most frequently cited instruments and has been validated for use in asymptomatic, symptomatic, and persistently symptomatic populations [10]. Its role in clinical research has been established, in large part because of the multidimensional nature of this instrument [7,11]. The original MPQ [12,13] comprises 78 pain descriptors broadly divided into 3 major classes: sensory, affective, and evaluative. These pain descriptors are further categorized into 20 subclasses, with the words within the respective subclasses ranked based on pain intensity.

The MPQ can be administered by an interviewer who reads the instructions to the patient and defines any words that are not understood [14]. It may also be self-administered. There is heavy reliance on language proficiency, a lack of which may limit the effectiveness of the instrument. Regional linguistic patterns may also be inadvertently incorporated into practice. In addition, language itself is constantly changing and evolving, with use and meanings constantly being updated.

The MPQ has remained unchanged since its inception, and the impact of modern language use on the relevance of MPQ pain descriptors remains unreported. During this time, the invention

of the internet and its exponential penetration have drastically reshaped our world. Social media, which comprises forums, blogs, business networks, social gaming, microblogs, photo-sharing platforms, and chat apps, has evolved dramatically alongside these developments. Furthermore, >50% of the global population was expected to access the internet in 2019, and the same figure was expected to use social media platforms [15]. This generates a staggering amount of publicly available social media data, allowing linguistic analysis at a scale that has never been seen before.

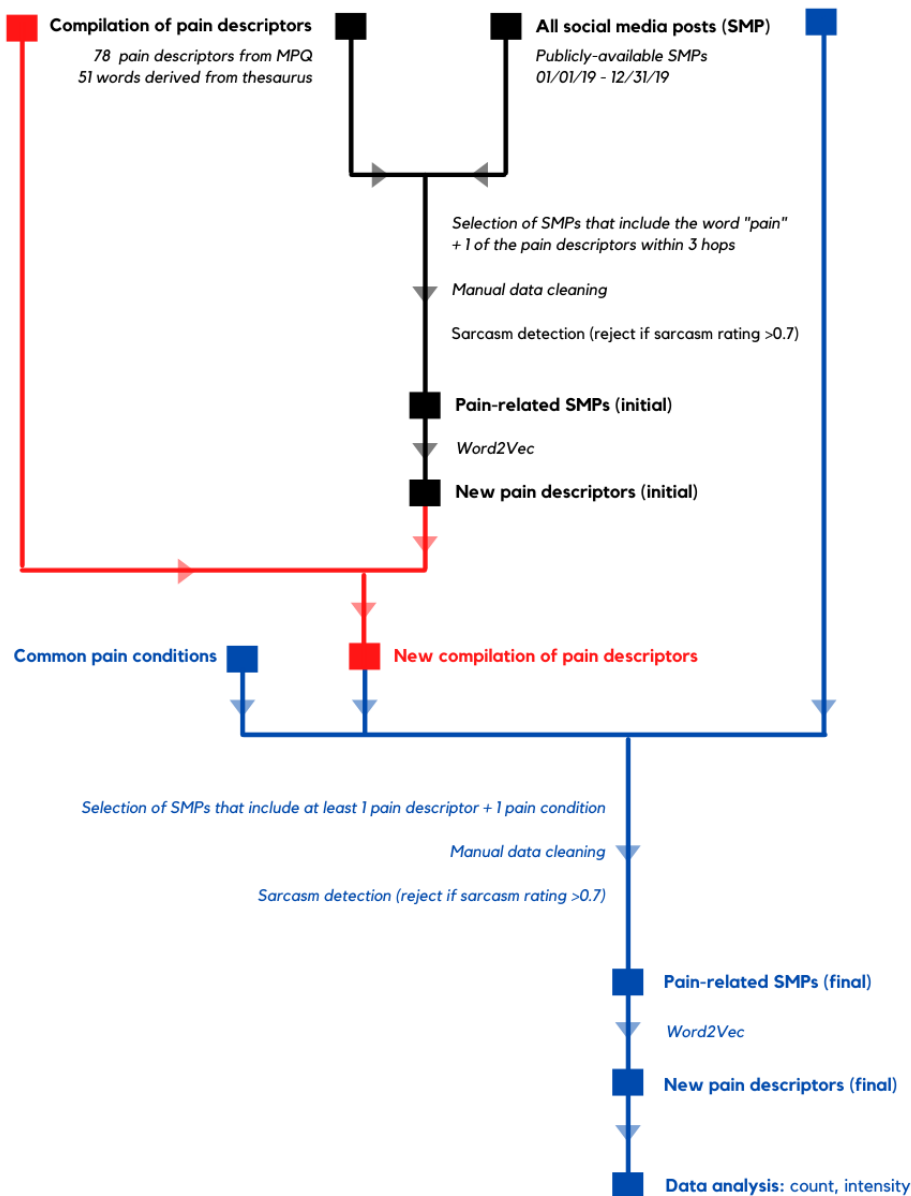
Natural language processing or computational linguistics, together with machine learning algorithms, have evolved substantially over the years to be able to analyze, learn, and understand the linguistic contexts of words, identify sentiments and emotions, and form neural network models [16]. These have enabled health researchers to use social media data to address a wide range of public health concerns, including surveillance and inference of developing trends in infectious diseases, targeting of public health strategies, evaluation of public health interventions, ascertaining public perceptions of nonmedical use of opioids, and even predicting disease status [17-21].

In relation to chronic pain, social media represents a snapshot of natural day-to-day colloquial language rather than formal communication. Furthermore, the nature of social media encourages users to capture their thoughts and ideas instantaneously. This is particularly important for accurate pain reporting, which is vulnerable to recall bias and relies heavily on timely reporting. Previous studies have observed extensive amounts of web-based conversations regarding pain [22], suggesting that social media may be a valid place to obtain a large amount of data regarding pain experience. Another study demonstrated that the characteristics of pain conditions could be discerned from social media posts [23], indicating that the depth of data may be sufficient for health care workers to further analyze and understand pain. Therefore, the objectives of this investigation are to use social media data to examine the relevance of pain descriptors from the existing MPQ, identify novel contemporary English descriptors for pain among users of social media, and suggest a modification for a new MPQ for future validation and testing.

Methods

Overview

This study used artificial intelligence and emotion analytics algorithms for the derivation and analysis of pain expression from social media platforms. The workflow was executed by a company specializing in linguistic and emotion analytics (INTNT.AI) and comprised 5 main steps conducted in an iterative process: (1) preliminary data gathering, (2) data cleaning, (3) Word2Vec (patent number US9037464B1; Google Inc), (4) final data gathering and cleaning, and (5) data analysis. The workflow is summarized in Figure 1. Preliminary data gathering was first performed to identify new pain descriptors through Word2Vec, which could then be used to maximize the search for relevant social media posts in the final data gathering.

Figure 1. Workflow for the derivation of new pain descriptors. SMP: social media post.

All posts from social media platforms over a 1-year period, from January 1, 2019, to December 31, 2019, were extracted. These data were acquired from a social listening platform (Meltwater) that aggregates and gives direct and official access to all accounts open to the public on Twitter, Facebook, Instagram, and YouTube.

Preliminary Data Gathering

A list comprising 78 pain descriptors from the MPQ and 51 additions yielded through the use of a mixture of web-based thesauruses was used to identify relevant social media posts (Multimedia Appendix 1). Word networks were created from social media posts using word graphs and path search algorithms that allow the linking of words to the root word of interest [24,25]. Words immediately next to the root word are described as one *hop* away, whereas subsequent layers of words are in increasing levels of hops.

Only social media posts containing the word *pain* and one of the targeted pain descriptors within 3 *hops* were included in this

study. The criterion of within 3 *hops* was decided based on pilot testing that showed good richness of social media posts that balanced relevance with the breadth of content.

Data Cleaning of the Social Media Posts

The selected social media posts were manually scrutinized and cleaned. Usernames, hyperlinks, and internet-specific symbols were removed. In addition, the content of the social media posts was evaluated for relevance. Social media posts that contained the pain descriptor in irrelevant contexts were removed.

The remaining posts were then input into a sarcasm detection machine (Crystalace, Institute of High Performance Computing, Agency for Science, Technology and Research), which is a support vector machine classifier trained with an affect-cognition-sociolinguistics feature model [26,27]. The Crystalace sarcasm detection method rated social media posts for sarcasm on a scale of 0-1, with 1 indicating maximum sarcasm. Social media posts with a sarcasm rating >0.7 were removed from the database.

space, and had a positive vector distance. A maximum of 20 words found to be most similar to each of the predetermined 129 keywords was selected for inclusion and further pruning. These were evaluated for relevance to pain descriptors. Entries with contrasting meanings to the keywords (eg, *hot vs cold*), entries that shared the same root word as the MPQ keywords, and entries irrelevant to pain description were removed.

The number of mentions, or count, of each pain descriptor within the final data set of selected social media posts was computed. Analysis of the counts for the 78 original MPQ keywords was conducted to determine the minimum threshold level for the inclusion of new associated words. The bottom 10% of the original MPQ pain descriptors were found to have counts of <110. Therefore, a count of 110 was set as the minimum threshold level for the prevalence of word use, and all words with a count of <110 were removed.

The intensities of all descriptors were also analyzed using natural language processing emotion analytics algorithms (CrystalFeel, Institute of High Performance Computing, Agency for Science, Technology and Research), which considered the entire sentence or paragraph in which each descriptor was found. The CrystalFeel algorithms allowed for the measurement of the emotional properties in text, including *anger, fear, sadness, joy, and valence*. Intensity was reflected on a scale of 0-1, with 1 implying maximum pain intensity. This formed the basis for determining the order of pain descriptors within each subclass.

Results

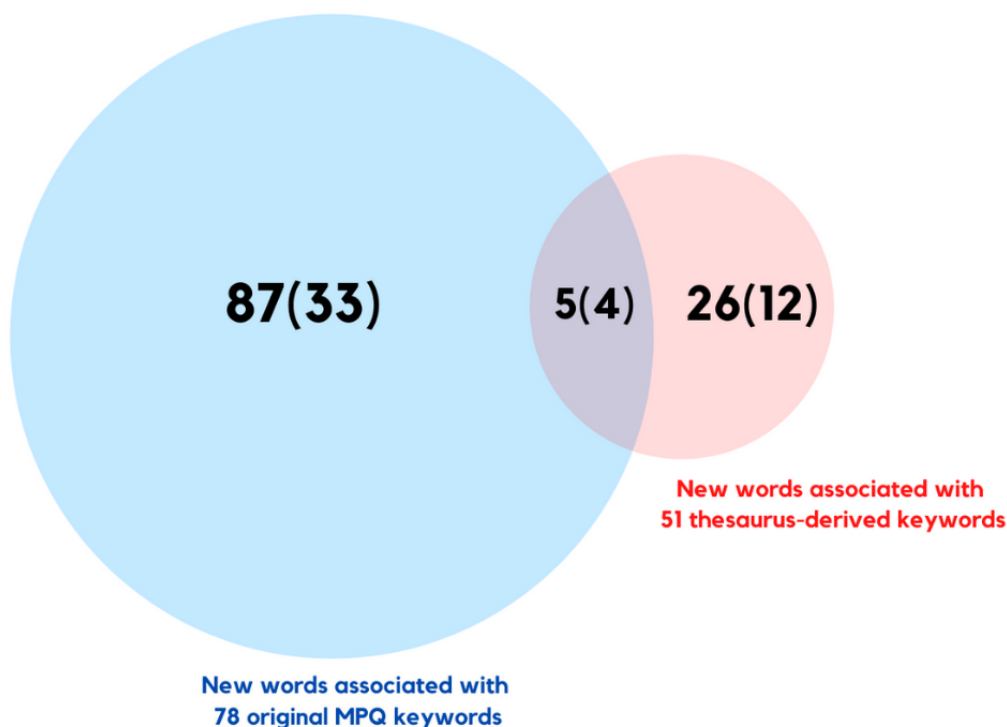
Sample of Social Media Posts

A total of 572,742 social media posts were obtained from a preliminary round of data gathering. Following manual evaluation of the 572,742 social media posts for relevance, 8310 (1.45%) social media posts were removed, whereas 7824 (1.37%) social media posts were removed after failing to meet the threshold criterion for sarcasm. Word2Vec identified 34 new pain descriptors that were used together with the original list of words to widen the search for relevant social media posts in the second round of data gathering. A total of 1,877,122 social media posts were identified in the second round of data gathering. After data cleaning and the additional inclusion criteria of containing at least one pain condition as well as one pain descriptor in the social media post, 11.55% (216,873/1,877,122) of social media posts remained for the final data analysis.

Identification of New Pain Descriptors Through Word2Vec

Using Word2Vec, a total of 118 new associated words were found for the 129 pain descriptor keywords defined in this study, following the removal of repetitions. Of these 118 words, 5 (4.2%) were associated with both the original MPQ and thesaurus-derived keywords, 87 (73.7%) were associated with MPQ keywords only, and 26 (22%) were associated with thesaurus-derived keywords only (Figure 3).

Figure 3. A total of 118 new words were found to be associated with the 78 original McGill Pain Questionnaire keywords and 51 thesaurus-derived keywords; bracketed values indicate the number of words with count ≥110. MPQ: McGill Pain Questionnaire.



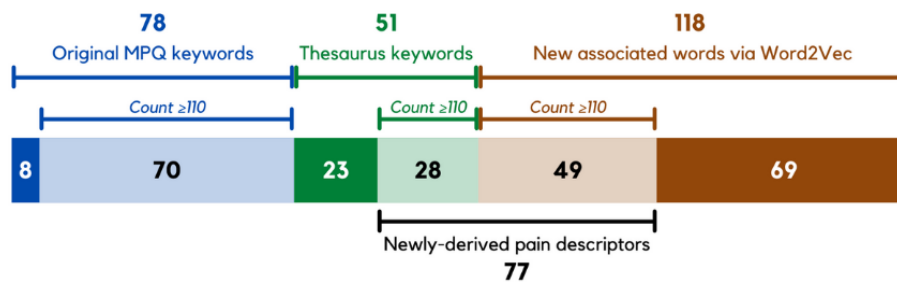
Count and Intensity of the Original MPQ Keywords and New Pain Descriptors

Of the 118 new associated words acquired through Word2Vec, 49 (41.5%) words were found to have a count of at least 110, meeting the minimum threshold level for the prevalence of word use.

28 (23.7%) thesaurus-derived keywords met the minimum threshold count of 110. These and the 49 out of 118 (41.5%)

new associated words derived through Word2Vec were combined into a single list for further evaluation, representing the final list of 77 newly derived pain descriptors (Figure 4). From this list, the pain descriptors that received the top 10 highest counts in descending order were *anxiety* (96909), *depression* (65223), *fear* (48165), *excruciating* (32094), *anger* (31506), *discomfort* (31333), *depress* (29822), *sadness* (25817), *low* (23027), and *fever* (21441).

Figure 4. Breakdown of original and newly derived pain descriptors. MPQ: McGill Pain Questionnaire.



Of the 78 original MPQ keywords, the pain descriptors that received the top 10 highest counts in descending order were *sharp* (26679), *hot* (25405), *tingling* (25405), *intense* (24269), *numb* (23964), *unbearable* (20929), *cold* (19099), *sore* (16127), *burning* (14204), and *itchy* (14204).

The intensity of the 78 original MPQ keywords ranged from 0.367 (for *flickering* and *jumping*) to 0.699 (for *terrifying*). For the combined list of 77 newly derived pain descriptors, the intensity ranged from 0.325 (for *scratching*) to 0.7467 (for *suicidal*).

The counts and intensities of existing and new words are presented in [Multimedia Appendix 3](#).

Suggested New Pain Questionnaire

Overview

Analyses of count and intensity of the pain descriptors, as well as manual analysis of the social media posts for the context in which the word was used, were used to recommend the inclusion of pain descriptors in a new pain questionnaire. The suggested changes, categorized by the MPQ subclasses, are summarized in [Table 1](#).

Table 1. Comparison of the original McGill Pain Questionnaire (MPQ) with the suggestion for a new pain questionnaire.

Subclass and pain descriptors from the original MPQ	Suggested words for a new pain questionnaire	Ranking reordered? (yes or no)
Temporal		
<ul style="list-style-type: none"> • Flickering • Quivering (removed) • Pulsing • Throbbing • Beating (removed) • Pounding 	<ul style="list-style-type: none"> • Flickering • Pulsing • Throbbing • Pounding 	<ul style="list-style-type: none"> • No
Spatial		
<ul style="list-style-type: none"> • Jumping • Flashing • Shooting 	<ul style="list-style-type: none"> • Jumping • Flashing • Shooting 	<ul style="list-style-type: none"> • No
Punctuate pressure		
<ul style="list-style-type: none"> • Pricking • Boring • Drilling • Stabbing • Lancinating (removed) 	<ul style="list-style-type: none"> • Pricking • Boring • Drilling • Stabbing • Puncturing (new) 	<ul style="list-style-type: none"> • No
Incisive pressure		
<ul style="list-style-type: none"> • Sharp • Cutting • Lacerating (removed) 	<ul style="list-style-type: none"> • Cutting • Sharp 	<ul style="list-style-type: none"> • Yes
Constrictive pressure		
<ul style="list-style-type: none"> • Pinching • Pressing • Gnawing • Cramping • Crushing 	<ul style="list-style-type: none"> • Pressing • Gnawing • Crushing • Pinching • Cramping 	<ul style="list-style-type: none"> • Yes
Traction pressure		
<ul style="list-style-type: none"> • Tugging • Pulling • Wrenching 	<ul style="list-style-type: none"> • Tugging • Contraction (new) • Pulling • Wrenching • Clenching (new) 	<ul style="list-style-type: none"> • No
Thermal		
<ul style="list-style-type: none"> • Hot • Burning • Scalding (removed) • Searing 	<ul style="list-style-type: none"> • Hot • Searing • Burning 	<ul style="list-style-type: none"> • Yes
Brightness		
<ul style="list-style-type: none"> • Tingling • Itchy • Smarting (removed) • Stinging 	<ul style="list-style-type: none"> • Scratching (new) • Tingling • Itchy • Stinging 	<ul style="list-style-type: none"> • No
Dullness		
<ul style="list-style-type: none"> • Dull • Sore • Hurting • Aching • Heavy 	<ul style="list-style-type: none"> • Sore • Aching • Dull • Hurting • Heavy 	<ul style="list-style-type: none"> • Yes
Sensory miscellaneous		

Subclass and pain descriptors from the original MPQ	Suggested words for a new pain questionnaire	Ranking reordered? (yes or no)
<ul style="list-style-type: none"> • Tender • Taut (removed) • Rasping (removed) • Splitting 	<ul style="list-style-type: none"> • Tender • Splitting 	<ul style="list-style-type: none"> • No
Tension		
<ul style="list-style-type: none"> • Tiring • Exhausting 	<ul style="list-style-type: none"> • Tiring • Straining (new) • Exhausting 	<ul style="list-style-type: none"> • No
Autonomic		
<ul style="list-style-type: none"> • Sickening • Suffocating 	<ul style="list-style-type: none"> • Sickening • Suffocating 	<ul style="list-style-type: none"> • No
Fear		
<ul style="list-style-type: none"> • Fearful • Frightful (removed) 	<ul style="list-style-type: none"> • Fearful • Horrendous (new) • Horrifying (new) 	<ul style="list-style-type: none"> • No
Punishment		
<ul style="list-style-type: none"> • Terrifying • Punishing (removed) • Grueling • Cruel • Vicious • Killing 	<ul style="list-style-type: none"> • Grueling • Cruel • Vicious • Killing • Terrifying 	<ul style="list-style-type: none"> • Yes
Affective-evaluative-sensory miscellaneous		
<ul style="list-style-type: none"> • Wretched • Blinding 	<ul style="list-style-type: none"> • Wretched • Blinding 	<ul style="list-style-type: none"> • No
Evaluative		
<ul style="list-style-type: none"> • Annoying • Troublesome • Miserable • Intense • Unbearable 	<ul style="list-style-type: none"> • Mild (new) • Troublesome • Intense • Annoying • Irritating (new) • Unbearable • Horrible (new) • Miserable • Excruciating (new) • Distressing (new) 	<ul style="list-style-type: none"> • Yes
Supplementary a		
<ul style="list-style-type: none"> • Spreading • Radiating • Penetrating • Piercing 	<ul style="list-style-type: none"> • Spreading • Radiating • Penetrating • Piercing 	<ul style="list-style-type: none"> • No
Supplementary b		
<ul style="list-style-type: none"> • Tight • Numb • Drawing (removed) • Squeezing • Tearing 	<ul style="list-style-type: none"> • Bruising (new) • Tight • Numb • Squeezing • Tearing 	<ul style="list-style-type: none"> • No
Supplementary c		

Subclass and pain descriptors from the original MPQ	Suggested words for a new pain questionnaire	Ranking reordered? (yes or no)
<ul style="list-style-type: none"> • Cool • Cold • Freezing 	<ul style="list-style-type: none"> • Cool • Cold • Freezing 	<ul style="list-style-type: none"> • No
Supplementary d		
<ul style="list-style-type: none"> • Nagging • Nauseating • Agonizing • Dreadful • Torturing 	<ul style="list-style-type: none"> • Nagging • Nauseating • Agonizing • Dreadful • Torturing 	<ul style="list-style-type: none"> • No
Psychological (new)		
___ ^a	<ul style="list-style-type: none"> • Worried (new) • Angry (new) • Fearful • Sad (new) • Depressed (new) • Nervous (new) • Anxious (new) • Feel hopeless (new) • Suicidal (new) 	<ul style="list-style-type: none"> • N/A^b

^aNot available.

^bN/A: not applicable.

Removal of Words

Of the 78 pain descriptors, 8 (10%) pain descriptors from the original MPQ were removed because of low use (count<110). The words removed were *quivering* (from the *Temporal* subclass), *lancinating* (from the *Punctuate Pressure* subclass), *lacerating* (from the *Incisive Pressure* subclass), *scalding* (from the *Thermal* subclass), *smarting* (from the *Brightness* subclass), *taut* and *rasping* (from the *Sensory miscellaneous* subclass), and *frightful* (from the *Fear* subclass).

In addition, of the 78 pain descriptors from the original MPQ, 3 (4%) pain descriptors were removed because, on manual analysis of the social media posts, the contexts in which they were used were found to have deviated from pain description. The word *beating* (from the *Temporal* subclass) was mainly used as a synonym for *overcoming pain* or the physical action of a beating instead of being used as a pain descriptor. *Punishing* (from the *Punishment* subclass) and *drawing* (from the *Supplementary b* subclass) were also removed as these words were more often used as verbs rather than adjectives.

Addition of Words

A total of 13 new associated words were added to the pre-existing subclasses. The added words were *puncturing* (to the *Punctuate Pressure* subclass); *contraction* and *clenching* (to the *Traction Pressure* subclass); *scratching* (to the *Brightness* subclass); *straining* (to the *Tension* subclass); *horrendous* and *horrifying* (to the *Fear* subclass); *mild*, *irritating*, *horrible*, *excruciating*, and *distressing* (to the *Evaluative* subclass); and *bruising* (to the *Supplementary b* subclass).

Creation of a Psychological Subclass

An entirely new *Psychological* subclass was suggested to acknowledge the nonphysical aspects of pain. This subclass

comprised 9 newly identified descriptors relevant to the description of the emotional state of a person. These were selected for their high counts, which ranged from 1139 to 65,223. Intensity was also taken into consideration during this selection to ensure good graduality of pain description, and this ranged from 0.47 to 0.7467.

The identified pain descriptors consisted of a mixture of nouns and adjectives. For consistency, the 9 selected descriptors were modified to fit into the sentence “My pain makes me ___.” This resulted in the final selection of *worried*, *angry*, *fearful*, *sad*, *depressed*, *nervous*, *anxious*, *feel hopeless*, and *suicidal* for this new subclass.

Reordering of Pain Descriptors to Reflect Pain Intensity

The pain descriptors of 6 of the 20 subclasses were reordered based on their emotional intensity to reflect decreasing pain intensity. These new rankings are shown in [Table 1](#).

Discussion

Principal Findings

This study provides insights into modern language use in the context of pain description. We found infrequent use and even a change in context for the use of several descriptors from the original MPQ, reflecting the evolution of language and suggesting limitations in the current MPQ. We also identified several new pain descriptors that can be used to update the MPQ, including the emergence of a possible *Psychological* subclass of pain descriptors that can capture the emotional and mental aspects of pain. These changes may improve the application of the MPQ in physical settings and digital telehealth platforms. An updated MPQ with relevant and appropriate pain descriptors may improve the ability of the MPQ to accurately characterize the pain experience, leading to more tailored diagnoses and

treatment plans. In addition, better awareness of modern language use can also influence the design of conversational artificial intelligence systems, commonly known as *chatbots*, which have the potential to decrease the need for physical encounters in health care facilities [31]. The linguistic styles of chatbots are thought to affect relationship building between patients and the respective chatbots [32].

The top 10 highest counts for the 78 original MPQ keywords were substantially lower than those for the new words, ranging from 14,204 to 26,679 and from 21,441 to 96,909, respectively. This suggests that words previously selected for the MPQ may no longer be as commonly used today.

Interestingly, 6 out of the top 10 newly identified pain descriptors that received the top 10 highest counts were relevant to the emotional or mental description of pain, namely *anxiety*, *depression*, *fear*, *anger*, *depress*, and *sadness*, which led to the emergence of a *Psychological* subclass.

Our study combined the big data afforded by social media posts with artificial intelligence and additional emotion analytics algorithms that allow for a broad analysis of social media posts with high speed and accuracy. The combination of pain research with this technology, originally designed to help businesses understand what their customers want through linguistic and contextual cues, helps to address a pressing health care need. As much as 40% of the population contends with chronic pain, and the extensive cumulative impact of chronic pain in the United States alone was estimated to exceed US \$500 billion annually [33].

The updated pain definition by the International Association for the Study of Pain highlights the subjective and emotional nature of pain. The literature reports various links between pain and mood or psychiatric disorders. For instance, pain and major depressive disorder often occur concurrently, appearing to mutually exacerbate the severity of the individual conditions [34-37]. Chronic pain has been found to increase the risk of depressive disorders and comorbidity; inversely, depression is also a risk factor for the later development of chronic pain [38-42]. Although not fully understood, several studies have reported shared biological pathways and neurotransmitters between both conditions [43-45]. The involvement of the limbic structures of the brain in both pain processing and depressed mood can be triggered by stress and indicates the sharing of neurobiological factors between pain and depression [46-49]. In addition, the pain experienced may be altered by emotions such as anger expression [50], anxiety, or the feeling of powerlessness, with 1 author suggesting that reducing the perception of powerlessness may reduce pain intensity [51].

Unfortunately, these psychological, emotional, and behavioral interactions with pain experience are rarely brought up in patient interviews. The existing MPQ focuses largely on the physical description of pain. Although its existing words allow the interviewer to infer the emotional toll of pain on the patient, there is lack of a dedicated segment that acknowledges the psychological burden of pain. The introduction of a new

Psychological subclass is a significant change from the existing MPQ, prompting patients to definitively select the resultant emotion caused by their pain. The patient's choice from the selection of *worried*, *angry*, *fearful*, *sad*, *depressed*, *nervous*, *anxious*, *feel hopeless*, and *suicidal* may also indicate the patient's primary coping mechanism in the face of the challenges of the pain condition. The behavior of a patient who feels *anger* may be starkly different from another who feels *fearful* and could aid future interventions or referrals by health care providers.

Limitations and Future Direction

This study has some limitations. Only social media posts in English were included in this study, and most social media posts originated from the Western hemisphere. Therefore, the application of these findings is largely limited to the United States and may not be generalizable to other English-speaking countries. Future work using similar artificial intelligence mechanisms stratified by geographical regions to address regional linguistic differences, and perhaps different languages, may be explored.

Similarly, the sample population in this study was limited to users of social media platforms. The preferred social media platform differs according to age, and in general, the overall population tends to skew young [52,53]. It is possible that the pain descriptors found are specific to this population only; therefore, the results may not be generalizable to older adults for whom the older MPQ terminologies may still apply. There may also be a lack of representation of people of lower socioeconomic status or those whose pain conditions severely compromise daily living [54].

Although the artificial intelligence and emotion analytics algorithms selected for the study had previously demonstrated good accuracy, we did not evaluate the validity of the newly identified pain descriptors. The complexity of human thought and expression makes further validation of the suggested questionnaire on a sample of the target patients necessary. Future research could involve a qualitative focus group to examine the face validity of the word descriptors and to ensure that patients feel that their pain is adequately described by the options available or to suggest other terms. Validation and reliability of the new pain questionnaire should be conducted, and the proposed reordering of words within each subclass should be tested.

Conclusions

The original MPQ is inadequate for reporting the psychological aspects of pain. Several descriptors from the original MPQ were also noted to have infrequent use or changes in context. This study used artificial intelligence and emotion analytics algorithms to identify contemporary vocabulary for pain description. The described methodology could be repeated at regular intervals to ensure the relevance of the pain questionnaires. Further research is needed to examine the reliability and validity of the revised MPQ.

Acknowledgments

The authors would like to express their gratitude to the team at INTNT.AI (Singapore) for their invaluable support and technical guidance: Mr Manuel Ho, Mr Xuyuan Kee, Dr Murphy Choy, and Dr Yang Yinping. This study was supported by a grant from the Faculty of Dentistry, National University of Singapore.

Authors' Contributions

MYT, CEG, and HHT conceived and designed the study, collected the data, contributed to data or analysis tools (obtained through a partnership with INTNT.AI), and performed the analysis. MYT and CEG were involved with writing the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Pain descriptors for the first round of data gathering.

[\[DOCX File , 15 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

List of common pain conditions.

[\[DOCX File , 14 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Counts and intensity of pain descriptors from the McGill Pain Questionnaire, web-based thesauruses, and identified through Word2Vec.

[\[DOCX File , 22 KB-Multimedia Appendix 3\]](#)

References

1. Raja SN, Carr DB, Cohen M, Finnerup NB, Flor H, Gibson S, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *Pain* 2020 Sep 01;161(9):1976-1982 [[FREE Full text](#)] [doi: [10.1097/j.pain.0000000000001939](https://doi.org/10.1097/j.pain.0000000000001939)] [Medline: [32694387](https://pubmed.ncbi.nlm.nih.gov/32694387/)]
2. Eliav E, Gracely R. Measuring and assessing pain. In: Sharav Y, Benoliel R, editors. *Orofacial Pain and Headache*. Amsterdam: Elsevier; 2008:45-56.
3. Dansie E, Turk D. Assessment of patients with chronic pain. *Br J Anaesth* 2013 Jul;111(1):19-25 [[FREE Full text](#)] [doi: [10.1093/bja/aet124](https://doi.org/10.1093/bja/aet124)] [Medline: [23794641](https://pubmed.ncbi.nlm.nih.gov/23794641/)]
4. Breivik H, Borchgrevink P, Allen S, Rosseland L, Romundstad L, Hals EK, et al. Assessment of pain. *Br J Anaesth* 2008 Jul;101(1):17-24 [[FREE Full text](#)] [doi: [10.1093/bja/aen103](https://doi.org/10.1093/bja/aen103)] [Medline: [18487245](https://pubmed.ncbi.nlm.nih.gov/18487245/)]
5. Breivik EK, Björnsson GA, Skovlund E. A comparison of pain rating scales by sampling from clinical trial data. *Clin J Pain* 2000 Mar;16(1):22-28. [doi: [10.1097/00002508-200003000-00005](https://doi.org/10.1097/00002508-200003000-00005)] [Medline: [10741815](https://pubmed.ncbi.nlm.nih.gov/10741815/)]
6. Hicks C, von Baeyer CL, Spafford P, van Korlaar I, Goodenough B. The Faces Pain Scale - revised: toward a common metric in pediatric pain measurement. *Pain* 2001;93:173-183. [doi: [10.1016/S0304-3959\(01\)00314-1](https://doi.org/10.1016/S0304-3959(01)00314-1)]
7. McGuire DB. Comprehensive and multidimensional assessment and measurement of pain. *J Pain Sympt Manag* 1992 Jul;7(5):312-319. [doi: [10.1016/0885-3924\(92\)90064-o](https://doi.org/10.1016/0885-3924(92)90064-o)]
8. Clark C, Ferrer-Brechner T, Janal M, Carroll D, Yang J. The dimensions of pain: a multidimensional scaling comparison of cancer patients and healthy volunteers. *Pain* 1989;37(1):23-32. [doi: [10.1016/0304-3959\(89\)90149-8](https://doi.org/10.1016/0304-3959(89)90149-8)]
9. Clark WC, Carroll JD, Yang JC, Janal MN. Multidimensional scaling reveals two dimensions of thermal pain. *Pain* 1987;28(2):277-278. [doi: [10.1016/0304-3959\(87\)90143-6](https://doi.org/10.1016/0304-3959(87)90143-6)]
10. Dworkin R, Turk D, Revicki D, Harding G, Coyne K, Peirce-Sandner S, et al. Development and initial validation of an expanded and revised version of the Short-form McGill Pain Questionnaire (SF-MPQ-2). *Pain* 2009 Jul;144(1-2):35-42. [doi: [10.1016/j.pain.2009.02.007](https://doi.org/10.1016/j.pain.2009.02.007)] [Medline: [19356853](https://pubmed.ncbi.nlm.nih.gov/19356853/)]
11. Flaherty SA. Pain measurement tools for clinical practice and research. *AANA J* 1996 Apr;64(2):133-140. [Medline: [9095685](https://pubmed.ncbi.nlm.nih.gov/9095685/)]
12. Melzack R, Torgerson W. On the language of pain. *Anesthesiology* 1971 Jan;34(1):50-59 [[FREE Full text](#)] [doi: [10.1097/00000542-197101000-00017](https://doi.org/10.1097/00000542-197101000-00017)] [Medline: [4924784](https://pubmed.ncbi.nlm.nih.gov/4924784/)]
13. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1(3):277-299. [doi: [10.1016/0304-3959\(75\)90044-5](https://doi.org/10.1016/0304-3959(75)90044-5)]
14. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ),

- Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken)* 2011 Nov 07;63(Suppl 11):240-252 [FREE Full text] [doi: [10.1002/acr.20543](https://doi.org/10.1002/acr.20543)] [Medline: [22588748](https://pubmed.ncbi.nlm.nih.gov/22588748/)]
15. Internet Live Stats. URL: <https://www.internetlivestats.com/> [accessed 2019-06-27]
 16. Hirschberg J, Manning CD. Advances in natural language processing. *Science* 2015 Jul 17;349(6245):261-266. [doi: [10.1126/science.aaa8685](https://doi.org/10.1126/science.aaa8685)] [Medline: [26185244](https://pubmed.ncbi.nlm.nih.gov/26185244/)]
 17. Subbiah IM, Hamilton E, Knoll M, Shanahan K, Meisel J. A big world made small: using social media to optimize patient care. *Am Soc Clin Oncol Educ Book* 2019 May;39:212-218. [doi: [10.1200/edbk.246643](https://doi.org/10.1200/edbk.246643)]
 18. McGloin AF, Eslami S. Digital and social media opportunities for dietary behaviour change. *Proc Nutr Soc* 2014 Oct 16;74(2):139-148. [doi: [10.1017/s0029665114001505](https://doi.org/10.1017/s0029665114001505)]
 19. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. *J Med Internet Res* 2014 Feb 11;16(2):e13 [FREE Full text] [doi: [10.2196/jmir.2912](https://doi.org/10.2196/jmir.2912)] [Medline: [24518354](https://pubmed.ncbi.nlm.nih.gov/24518354/)]
 20. Chew AM, Gunasekeran DV. Social media big data: the good, the bad, and the ugly (un)truths. *Front Big Data* 2021 Jun 1;4:623794 [FREE Full text] [doi: [10.3389/fdata.2021.623794](https://doi.org/10.3389/fdata.2021.623794)] [Medline: [34142082](https://pubmed.ncbi.nlm.nih.gov/34142082/)]
 21. Chan B, Lopez A, Sarkar U. The canary in the coal mine tweets: social media reveals public perceptions of non-medical use of opioids. *PLoS One* 2015 Aug 7;10(8):e0135072 [FREE Full text] [doi: [10.1371/journal.pone.0135072](https://doi.org/10.1371/journal.pone.0135072)] [Medline: [26252774](https://pubmed.ncbi.nlm.nih.gov/26252774/)]
 22. Reilly E, Sengupta R. Back pain, ankylosing spondylitis and social media usage; a descriptive analysis of current activity. *Rheumatol Int* 2020 Sep 19;40(9):1493-1499 [FREE Full text] [doi: [10.1007/s00296-020-04600-w](https://doi.org/10.1007/s00296-020-04600-w)] [Medline: [32430806](https://pubmed.ncbi.nlm.nih.gov/32430806/)]
 23. Ahlwardt K, Heavilin N, Gibbs J, Page J, Gerbert B, Tsoh JY. Tweeting about pain: comparing self-reported toothache experiences with those of backaches, earaches and headaches. *J Am Dent Assoc* 2014 Jul;145(7):737-743 [FREE Full text] [doi: [10.14219/jada.2014.30](https://doi.org/10.14219/jada.2014.30)] [Medline: [24982280](https://pubmed.ncbi.nlm.nih.gov/24982280/)]
 24. Wang D, Zou L, Pan W, Zhao D. Keyword graph: answering keyword search over large graphs. In: Zhou S, Zhang S, Karypis G, editors. *Advanced Data Mining and Applications*. Berlin, Heidelberg: Springer; 2012:635-649.
 25. Qiao M, Qin L, Cheng H, Yu JX, Tian W. Top-K nearest keyword search on large graphs. *Proc VLDB Endow* 2013 Aug 26;6(10):901-912. [doi: [10.14778/2536206.2536217](https://doi.org/10.14778/2536206.2536217)]
 26. Gupta R, Yang Y. CrystalNest at SemEval-2017 Task 4: using sarcasm detection for enhancing sentiment classification and quantification. In: *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*. 2017 Presented at: 11th International Workshop on Semantic Evaluation (SemEval-2017); August, 2017; Vancouver, Canada p. 626-633. [doi: [10.18653/v1/s17-2103](https://doi.org/10.18653/v1/s17-2103)]
 27. Gupta RK, Yang Y. CrystalFeel at SemEval-2018 Task 1: understanding and detecting emotion intensity using affective lexicons. In: *Proceedings of The 12th International Workshop on Semantic Evaluation*. 2018 Presented at: The 12th International Workshop on Semantic Evaluation; June, 2018; New Orleans, Louisiana p. 256-263. [doi: [10.18653/v1/s18-1038](https://doi.org/10.18653/v1/s18-1038)]
 28. Haiman J. *Talk Is Cheap: Sarcasm, Alienation, and the Evolution of Language*. Oxford, United Kingdom: Oxford University Press; 1998:1-232.
 29. Crystalfeel - multidimensional emotion intensity analysis from natural language. A*STAR's Institute of High Performance Computing. 2021. URL: <https://socialanalyticsplus.net/crystalfeel> [accessed 2021-05-31]
 30. Goldberg Y, Levy O. word2vec explained: deriving Mikolov et al.'s negative-sampling word-embedding method. arXiv. 2014. URL: <https://arxiv.org/abs/1402.3722> [accessed 2021-10-30]
 31. Li LW, Chew AM, Gunasekeran DV. Digital health for patients with chronic pain during the COVID-19 pandemic. *Br J Anaesth* 2020 Nov;125(5):657-660 [FREE Full text] [doi: [10.1016/j.bja.2020.08.003](https://doi.org/10.1016/j.bja.2020.08.003)] [Medline: [32863018](https://pubmed.ncbi.nlm.nih.gov/32863018/)]
 32. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. *ACM Trans Comput-Hum Interact* 2005 Jun;12(2):293-327. [doi: [10.1145/1067860.1067867](https://doi.org/10.1145/1067860.1067867)]
 33. Institute of Medicine, Board on Health Sciences Policy, Committee on Advancing Pain Research, Care, and Education. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*. Washington, DC: National Academies Press (US); 2011.
 34. Li J. Pain and depression comorbidity: a preclinical perspective. *Behav Brain Res* 2015 Jan 01;276:92-98 [FREE Full text] [doi: [10.1016/j.bbr.2014.04.042](https://doi.org/10.1016/j.bbr.2014.04.042)] [Medline: [24797835](https://pubmed.ncbi.nlm.nih.gov/24797835/)]
 35. Ohayon MM. Specific characteristics of the pain/depression association in the general population. *J Clin Psychiatry* 2004;65(Suppl 12):5-9. [Medline: [15315471](https://pubmed.ncbi.nlm.nih.gov/15315471/)]
 36. Geerlings SW, Beekman AT, Deeg DJ, Twisk JW, Van Tilburg W. Duration and severity of depression predict mortality in older adults in the community. *Psychol Med* 2002 May 20;32(4):609-618. [doi: [10.1017/s0033291702005585](https://doi.org/10.1017/s0033291702005585)] [Medline: [12102375](https://pubmed.ncbi.nlm.nih.gov/12102375/)]
 37. IsHak WW, Wen RY, Naghdechi L, Vanle B, Dang J, Knosp M, et al. Pain and depression: a systematic review. *Harv Rev Psychiatry* 2018;26(6):352-363. [doi: [10.1097/hrp.000000000000198](https://doi.org/10.1097/hrp.000000000000198)]
 38. Hilderink P, Burger H, Deeg D, Beekman A, Voshaar RC. The temporal relation between pain and depression: results from the longitudinal aging study Amsterdam. *Psychosom Med* 2012;74(9):945-951. [doi: [10.1097/PSY.0b013e3182733fdd](https://doi.org/10.1097/PSY.0b013e3182733fdd)] [Medline: [23115345](https://pubmed.ncbi.nlm.nih.gov/23115345/)]

39. Yamamoto Y, Hayashino Y, Akiba T, Akizawa T, Asano Y, Saito A, et al. Depressive symptoms predict the subsequent risk of bodily pain in dialysis patients: Japan Dialysis Outcomes and Practice Patterns Study. *Pain Med* 2009 Jul 01;10(5):883-889. [doi: [10.1111/j.1526-4637.2009.00661.x](https://doi.org/10.1111/j.1526-4637.2009.00661.x)] [Medline: [19682272](https://pubmed.ncbi.nlm.nih.gov/19682272/)]
40. Fillingim RB, Ohrbach R, Greenspan JD, Knott C, Diatchenko L, Dubner R, et al. Psychological factors associated with development of TMD: the OPPERA prospective cohort study. *J Pain* 2013 Dec;14(12 Suppl):75-90 [FREE Full text] [doi: [10.1016/j.jpain.2013.06.009](https://doi.org/10.1016/j.jpain.2013.06.009)] [Medline: [24275225](https://pubmed.ncbi.nlm.nih.gov/24275225/)]
41. Knaster P, Karlsson H, Estlander A, Kalso E. Psychiatric disorders as assessed with SCID in chronic pain patients: the anxiety disorders precede the onset of pain. *Gen Hosp Psychiatry* 2012 Jan;34(1):46-52. [doi: [10.1016/j.genhosppsych.2011.09.004](https://doi.org/10.1016/j.genhosppsych.2011.09.004)] [Medline: [22001549](https://pubmed.ncbi.nlm.nih.gov/22001549/)]
42. McBeth J, Jones K. Epidemiology of chronic musculoskeletal pain. *Best Pract Res Clin Rheumatol* 2007 Jun;21(3):403-425. [doi: [10.1016/j.berh.2007.03.003](https://doi.org/10.1016/j.berh.2007.03.003)] [Medline: [17602991](https://pubmed.ncbi.nlm.nih.gov/17602991/)]
43. Gallagher R, Verma S. Managing pain and comorbid depression: a public health challenge. *Semin Clin Neuropsychiatry* 1999 Jul;4(3):203-220. [Medline: [10498788](https://pubmed.ncbi.nlm.nih.gov/10498788/)]
44. Blier P, Abbott F. Putative mechanisms of action of antidepressant drugs in affective and anxiety disorders and pain. *J Psychiatry Neurosci* 2001 Jan;26(1):37-43 [FREE Full text] [Medline: [11212592](https://pubmed.ncbi.nlm.nih.gov/11212592/)]
45. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. *Arch Intern Med* 2003 Nov 10;163(20):2433-2445. [doi: [10.1001/archinte.163.20.2433](https://doi.org/10.1001/archinte.163.20.2433)] [Medline: [14609780](https://pubmed.ncbi.nlm.nih.gov/14609780/)]
46. Giesecke T, Gracely RH, Williams DA, Geisser ME, Petzke FW, Clauw DJ. The relationship between depression, clinical pain, and experimental pain in a chronic pain cohort. *Arthritis Rheum* 2005 May;52(5):1577-1584 [FREE Full text] [doi: [10.1002/art.21008](https://doi.org/10.1002/art.21008)] [Medline: [15880832](https://pubmed.ncbi.nlm.nih.gov/15880832/)]
47. Wager TD, Atlas LY, Lindquist MA, Roy M, Woo C, Kross E. An fMRI-based neurologic signature of physical pain. *N Engl J Med* 2013 Apr 11;368(15):1388-1397. [doi: [10.1056/nejmoa1204471](https://doi.org/10.1056/nejmoa1204471)]
48. Schmidt-Wilcke T, Clauw DJ. Fibromyalgia: from pathophysiology to therapy. *Nat Rev Rheumatol* 2011 Jul 19;7(9):518-527. [doi: [10.1038/nrrheum.2011.98](https://doi.org/10.1038/nrrheum.2011.98)] [Medline: [21769128](https://pubmed.ncbi.nlm.nih.gov/21769128/)]
49. Clauw DJ. Fibromyalgia: a clinical review. *J Am Med Assoc* 2014 Apr 16;311(15):1547-1555. [doi: [10.1001/jama.2014.3266](https://doi.org/10.1001/jama.2014.3266)] [Medline: [24737367](https://pubmed.ncbi.nlm.nih.gov/24737367/)]
50. Bruehl S, Chung OY, Burns JW. Anger expression and pain: an overview of findings and possible mechanisms. *J Behav Med* 2006 Dec 29;29(6):593-606. [doi: [10.1007/s10865-006-9060-9](https://doi.org/10.1007/s10865-006-9060-9)] [Medline: [16807797](https://pubmed.ncbi.nlm.nih.gov/16807797/)]
51. Walding MF. Pain, anxiety and powerlessness. *J Adv Nurs* 1991 Apr;16(4):388-397. [doi: [10.1111/j.1365-2648.1991.tb03427.x](https://doi.org/10.1111/j.1365-2648.1991.tb03427.x)] [Medline: [2061501](https://pubmed.ncbi.nlm.nih.gov/2061501/)]
52. Bolton RN, Parasuraman A, Hoefnagels A, Migchels N, Kabadayi S, Gruber T, et al. Understanding Generation Y and their use of social media: a review and research agenda. *J Serv Manag* 2013 Jun 14;24(3):245-267. [doi: [10.1108/09564231311326987](https://doi.org/10.1108/09564231311326987)]
53. Smith A, Anderson M. Social media use in 2018. Pew Research Center. 2018. URL: <https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/> [accessed 2021-10-30]
54. Caes L, Jones A, Jordan A. Engaging use of social media as a research tool to capture the daily life experiences of young people with chronic pain. *Evid Based Nurs* 2018 Oct 29;21(4):91-93. [doi: [10.1136/eb-2018-102980](https://doi.org/10.1136/eb-2018-102980)] [Medline: [30158175](https://pubmed.ncbi.nlm.nih.gov/30158175/)]

Abbreviations

MPQ: McGill Pain Questionnaire

Edited by G Eysenbach; submitted 18.06.21; peer-reviewed by D Gunasekeran, C Cheng, R Lee; comments to author 25.08.21; revised version received 06.09.21; accepted 18.09.21; published 25.11.21

Please cite as:

Tan MY, Goh CE, Tan HH

Contemporary English Pain Descriptors as Detected on Social Media Using Artificial Intelligence and Emotion Analytics Algorithms: Cross-sectional Study

JMIR Form Res 2021;5(11):e31366

URL: <https://formative.jmir.org/2021/11/e31366>

doi: [10.2196/31366](https://doi.org/10.2196/31366)

PMID:

©Ming Yi Tan, Charlene Enhui Goh, Hee Hon Tan. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium,

provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.