A Cross-Modal Study of Pain Across Communities in the United States

Arnav Aggarwal
University of Pennsylvania
Philadelphia, PA, USA
arnavagg@seas.upenn.edu

Sunny Rai
University of Pennsylvania
Philadelphia, PA, USA
sunnyrai@seas.upenn.edu

Salvatore Giorgi
University of Pennsylvania
Philadelphia, PA, USA
sgiorgi@sas.upenn.edu

Shreya Havaldar
University of Pennsylvania
Philadelphia, PA, USA
shreyah@seas.upenn.edu

Garrick Sherman
University of Pennsylvania
Philadelphia, PA, USA
garricktsherman@gmail.com

Juhi Mittal
University of Pennsylvania
Philadelphia, PA, USA
mittaljuhi98@gmail.com

Sharath Chandra Guntuku
University of Pennsylvania
Philadelphia, PA, USA
sharathg@seas.upenn.edu

ABSTRACT

Pain is one of the most prevalent reasons for seeking medical attention in the United States. Understanding how different communities report and express pain can aid in directing medical efforts and in advancing precision pain management. Using a large-scale self-report survey data set on pain from Gallup (2.5 million surveys) and social media posts from Twitter (1.8 million tweets), we investigate a) if Twitter posts could predict community-level pain and b) how expressions of pain differ across communities in the United States. Beyond observing an improvement of over 9% (in Pearson r) when using Twitter language over demographics to predict community-level pain, our study reveals that the discourse on pain varied significantly across communities in the United States. Evangelical Hubs frequently post about God, lessons from struggle, and prayers when expressing pain, whereas Working Class Country posts about regret and extreme endurance. Academic stresses, injuries, painkillers, and surgeries were the most commonly discussed pain themes in College Towns; Graying America discussed therapy, used emotional language around empathy and anger, and posted about chronic pain treatment; the African American South posted about struggles, patience, and faith when talking about pain. Our study demonstrates the efficacy of using Twitter to predict survey-based self-reports of pain across communities and has implications in aiding community-focused pain management interventions.

CCS CONCEPTS

• **Applied computing → Health informatics**.

KEYWORDS

pain, social media, natural language processing, American communities

WWW '23 Companion, April 30–May 04, 2023, Austin, TX, USA

ACM ISBN 978-1-4503-9419-2/23/04…$15.00
https://doi.org/10.1145/3543873.3587642

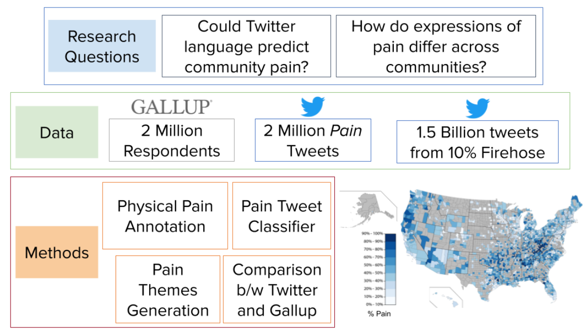

Figure 1: Overview of the study with a map of the percentage of individuals in each U.S. county reporting pain through Gallup

ACM Reference Format:
Arnav Aggarwal, Sunny Rai, Salvatore Giorgi, Shreya Havaldar, Garrick Sherman, Juhi Mittal, and Sharath Chandra Guntuku. 2023. A Cross-Modal Study of Pain Across Communities in the United States. In *Companion Proceedings of the ACM Web Conference 2023 (WWW '23 Companion), April 30–May 04, 2023, Austin, TX, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3543873.3587642

1 INTRODUCTION

Millions of Americans experience pain daily, impairing their ability to carry out essential routine activities [70]. The Declaration of Montreal (2011)—IASP [50], Article 1 states the right of all people to have access to pain management without *discrimination* and Article 2 states the right of all people to have *acknowledgment* of their pain. Pain is a highly subjective experience that is influenced by a variety of biological, psychological, and social factors [45, 60]. Patients become frustrated and feel stigmatized if their pain is trivialized, their symptoms are dismissed, or their discomfort is labeled as merely "psychological" by their healthcare professional [65]. Moreover, the opioid epidemic in the United States has resulted in restrictions on legal opioid drug accessibility, weakening the goal of inclusive pain management. Pitcher et al. [53] emphasize the need to study pain across heterogeneous populations to design effective policies for accessible pain management.

Pain measurements are often not collected or reported inaccurately [54, 64]. For instance, the commonly used Numeric Pain Rating Scale asks the patient, to make three pain ratings on a scale of 0 to 10, corresponding to current, best, and worst pain experienced over the past 24 hours and the average is taken to represent the level of pain. The setting in which these measurements are frequently performed (emergency room in a hospital) also renders any preventive intervention impossible [46]. There is no standardized definition of pain [8, 23], and how one perceives, expresses, and copes with it has been found to significantly vary [63]. In order to capture the "subjective" pain, Wideman et al. [68] recommends the use of qualitative pain narratives such as observing, talking, or listening to the patients over quantitative pain assessment scales.

Several studies have investigated the relationship between language on social media and public health, including mental health [14], psychological stress [31, 48], well-being [36], and vaccination rates [16]. While Twitter has been used to study specific examples of pain such as dental pain and back pain, along with the sentiment of users' tweeting about pain [22, 34, 62], the study of pain across various communities and the corresponding socio-cultural context remains unexplored. Additionally, while several definitions of pain have been suggested, there is limited work evaluating the relationship between self-reported pain and the expressions of pain. If automated methods could provide an understanding of communities, in terms of pain levels and their expression, those communities could be targeted for more thorough assessment and personalized pain management.

Pain is closely associated with socioeconomic variables such as age, sex, and race. The American Communities Project (ACP) is a non-spatial proximity-based county-level clustering that uses demographic, cultural, and socioeconomic characteristics such as race, income, education, and religion when clustering all counties in the U.S. into one of 15 communities such as Big Cities, African American South, and College towns [9]. Prior studies [33, 43] have extensively used these communities to validate the variation of health and behavioral outcomes across U.S. counties. For our study, we use Twitter to study the ecological expressions of pain in communities across the U.S. as described in ACP typology.
In summary, we examine two questions:

- **RQ1** Does Twitter language predict self-reported community-level pain above and beyond demographics?
- **RQ2** How do the expressions of pain on Twitter vary across ACP communities?

First, we identify tweets that contain expressions of physical pain. To do this, we train an automatic classifier to distinguish between expressions of emotion and physical pain on an annotated data set. We then apply this classifier on a large unannotated data set of tweets containing the keyword 'pain'. From the tweets returned by the classifier, we generate pain-related topics via Latent Dirichlet Allocation (LDA) [5]. Next, we extract the distribution of pain topics on a public county-level Twitter data set (County Tweet Lexical Bank) [28] to evaluate the predictive utility of Twitter on self-reported pain in Gallup surveys and to understand the variance in pain discourse across the U.S. communities.

2 BACKGROUND

Pain is closely associated with a multitude of socioeconomic variables including age, sex, and race along with external environmental factors such as opioid prescription. Lack of attention to the variables directly or indirectly obstructing pain management may reinforce socioeconomic disparities, limiting the development of particular groups. Here, we describe a few characteristics which have been found to significantly impact the experience of pain.

2.1 Age and Biological Sex

The human body tends to wear down with growing age which may lead to more pain. In a study of Americans, pain was shown to increase with age till one's mid-life and thereafter stabilize [8]. Interestingly, each successive generation in the U.S. reported more pain than the last, particularly for those without a bachelor's degree. No developed country other than the U.S. has such a cycle of increasing pain and excess midlife pain [8]. Similarly, women were reported to show greater pain sensitivity than men in 29 out of 34 pain variables [51].

2.2 Race and Ethnicity

Experience of pain is also associated with race and ethnicity and has biological, social, and psychological pathways [60]. Black Americans were found to be the most pain sensitive and faced the most severe joint pain and work limitations, followed by Hispanic Americans, Asian Americans, and Non-Hispanic White Americans [6, 18, 51]. Several reasons for the racial and ethnic differences in pain have been suggested including reports of under-treatment of minority populations [2]. Shavers et al. [60] also suggested that ethnic disparities might be caused by varying degrees of access to healthcare or disbelief towards a patient's pain-related experiences, behavior, and community.

2.3 Opioids for Pain Management

In contrast to other parts of the world where opioids are highly regulated, their widespread use and availability have been found to be another possible cause for the variance in the experience of pain in the U.S. The variance in pain could both be the consequence and the cause of the U.S. opioid epidemic. For instance, prolonged opioid use has been shown to increase pain in the long term even though it reduces pain in the shorter term [12].

We build on these associations between pain and demographic and societal factors by studying the naturalistic expressions around pain on Twitter across the different American communities as defined in the ACP typology.

3 DATA

In this study, we use both self-reported survey-based measures of pain and expressions of pain on social media (i.e., Twitter). Each data set is described below. The authors' university Institutional Review Board deemed this study exempt. We will make available the Twitter data collected for this study, in compliance with Twitter's Terms of Service as a resource to the research community to better understand the pain and the associated social aspects that influence health.

3.1 Self-reported Pain: Gallup Well Being Index

The Gallup-Sharecare Well-Being Index is a large national phone survey with 1,000 U.S. adults surveyed daily from 2008-2012 and 3,500 U.S. adults surveyed daily from 2013-2019 to ask the following binary question: *"Did you experience physical pain for a lot yesterday?"* The answer could be Yes/1 or No/0. Demographics (e.g., age, gender, and income) and U.S. county location information are also collected for each response. In total, the data set consists of $N = 2,541,688$ responses.

While Ward et al. [67] suggest a minimum of 300 responses are needed for stable county-level estimates, we decided to lower this threshold to 200 in order to increase the total number of counties used in our analysis ($N = 1,397$ for 300 vs $N = 1,641$ for 200). The pain response for all individuals within a county is then averaged. In order to compensate for disproportionate selection probabilities and non-response amongst participants, we used a set of post-stratification weights derived by Gallup through a standard iterative proportional fitting algorithm [21].

3.2 County Twitter Data

The County Tweet Lexical Bank (CTLB) is a publicly available data set of U.S. county-level linguistic features (i.e., 25,000 1grams and 2,000 LDA topic distributions) extracted from 10% Twitter sample from 2009-2015 [28] which has been shown to reliably predict socio-economic and well-being related parameters. This data set consists of approximately 1.5 billion tweets from 6 million U.S. county-mapped Twitter users. Twitter users are geolocated to U.S. counties via latitude and longitude coordinates in their tweets or via self-reported location information in their Twitter profile. Each user in the data set posted at least 30 tweets and each of the 2,041 U.S. counties in the data set contains at least 100 such users. See Giorgi et al. [28] for full details on the county mapping process, minimum tweet thresholds, linguistic feature extraction, etc. For our analysis, we took the intersection of counties with sufficient Gallup responses (as mentioned above), resulting in $N = 1,641$ counties. This data is used in both **RQ1** and **RQ2**.

3.3 Expressions of Pain

Another data set of 8,103,702 tweets was obtained from the Twitter Historical API that all contain at least one pain-related word (see Appendix A.1 for a full list of the pain-related words considered). These tweets were then mapped to corresponding counties in the U.S. based on a set of heuristic rules – a combination of latitude and longitude coordinates and locations in the user descriptions, i.e., the same process used to county map the CTLB (above) as originally developed by Schwartz et al. [58]. We did this to identify tweets posted from the U.S. and to avoid any cross-cultural confounds in expressions of pain. We removed tweets that were not from 1,641 intersecting counties with Gallup to maintain geographical consistency, which left us with 1,809,805 million tweets. A subset of this data is manually annotated and used to train a physical pain classifier and build pain-related LDA topics.

3.3.1 *Pain Tweets Annotation.*

We picked a random set of 3,000 tweets from the Twitter Expressions of Pain data set and asked two annotators to label if each tweet was an expression of pain or not. If annotated as pain (by both annotators), these tweets were further segregated into whether they were *emotional pain* or *physical pain*. An example of a physical pain tweet would be: "My knee has been paining since last week" while an example of an emotional pain tweet would be "Ever since she left, I have been in pain." The annotators were initially asked to annotate a common sample of 300 tweets. The agreement between two annotators was found to be 0.9 (Cohen's Kappa) on this common set indicating a uniform understanding of pain by both annotators. The rest of the tweets were then independently annotated by each annotator. Due to the metaphorical and colloquial language on Twitter, several tweets that contained the keyword 'pain' were in fact song lyrics, ads for pain relief balms, etc. For instance, several tweets referred to T-Pain, the American rapper. After annotation, 26.3% of the tweets were classified as pain, of which 22.2% were classified as emotional pain.

3.4 Pain Tweets Classifier

We created two task-specific classifiers, respectively for (a) detecting expressions of pain in tweets, and (b) identifying physical pain tweets among pain tweets. For both tasks, we extracted the following features:

- Bag-of-Words (1 grams)
- 100 topics generated on the Twitter Expressions of Pain data using Latent Dirichlet Allocation (LDA) [5] with an *alpha* of 5, *beta* of 0.01, and 1000 Gibbs sampling iterations through the MALLET package
- Linguistic Inquiry and Word Count (LIWC) categories [52]

3.4.1 *Detecting Pain Expressions.*

For the task of Pain detection, we evaluated the performance of the above feature sets on classification models namely Logistic Regression (LR) [10], Random Forest Classifier (RFC) [35], and Extra Trees Classifiers (ETC) [24]. The annotated tweets were split into 10 stratified folds and models were tested in a cross-validation framework. The results are provided in Appendix Table A1.

RFC and ETC on a concatenation of all features had the highest performance i.e. 0.86 (in terms of AuC). Considering the better precision, we picked the RFC model to identify the pain-expressing tweets from the unannotated tweet data set as described in Sec. 3. This resulted in 89% of the database being classified as pain-related tweets. We obtained a random sample of 100 tweets predicted as pain by the classifier and found that the estimates were accurate.

3.4.2 *Identifying Physical Pain.*

To identify tweets expressing physical pain, we trained another classifier using the 26.3% of tweets that were marked with the type of pain (emotional vs physical) by human annotators in Section 3.3.1. While emotional pain is also of interest, we wanted to limit the scope to physical pain to be consistent with the Gallup survey question. As can be seen in Table A2, the RFC model when trained on a combined feature set had the highest AuC and precision scores, and was thus applied on all tweets predicted as expressing pain in Sec. 3.4.1.

Upon heuristic observation, we set the threshold for a tweet to be classified as a physical pain-related tweet to be 0.44 which resulted in 98.6% of tweets from the Pain data set being classified as emotional pain tweets.

4 PAIN TOPICS GENERATION

To obtain the underlying themes in tweets identified as expressing physical pain, we used a process called Content Specific LDA (CSLDA) [71]. CSLDA is a standard LDA topic modeling algorithm used across thematically limited corpora (e.g., pain). As such, there is a text preprocessing pipeline that identifies words associated with a given theme and removes non-representative words (e.g., generic Twitter language). This method has been used to identify alcohol consumption topics [29], Black Lives Matter topics [26], and diabetes topics [30]. Once the text is preprocessed, then the standard LDA algorithm is used to automatically identify topics. Full details on the CSLDA pipeline can be found in the Supplement.

To build a set of pain topics, we chose to create sets of 25, 50, 75, and 100 topics from our corpus of physical pain tweets. We then evaluated the topics using both qualitative and quantitative methods. For the qualitative evaluation, three authors were asked to evaluate each topic set according to three criteria: (1) is there a breadth of themes across the topics, (2) do single topics contain single themes, and (3) is there minimal repetition of themes across a large number of topics. All three raters independently decided that the set of 75 topics best satisfied the three criteria.

Table 1: CSLDA Topic Evaluation: topic uniqueness and coherence scores for each set of pain topics. Bolded numbers show the topic set chosen through a qualitative evaluation process.

Number of Topics	Topic Uniqueness	Coherence
25	1.00	.65
50	.91	.67
75	**.79**	**.67**
100	.68	.65

For the quantitative evaluation, we calculate both topic uniqueness (TU) and coherence. TU is a measure of topic diversity and is inversely proportional to the number of times a set of representative keywords is repeated across the set of topics [49]. TU scores range from 0 to 1 and high TU scores indicate that the keywords are rarely repeated across the topics. Full details on TU can be found in the Supplement. Coherence measures the semantic similarity between words in the topic using Normalized Pointwise Mutual Information [61]. Coherence is measured for each topic and averaged across all topics. Scores range from 0 to 1, with a score of 1 representing a topic with high semantic similarity between words. We use the Gensim Python package to calculate the coherence scores [56]. All results are shown in Table 1.

5 ESTIMATING COMMUNITY PAIN WITH TWITTER LANGUAGE

Social media is popular for large-scale population evaluation as it offers a low-cost, non-intrusive alternative to traditional surveys with finer spatiotemporal scales [19, 25, 66]. Models trained on social media language can offer robust community estimates when sufficient responses are unavailable [27]. We thus evaluate the role of Twitter language in estimating community pain. Specifically, we ask if Twitter language can be used to predict U.S. county-level self-reports of pain.

We performed ridge regression in a 10-fold cross-validation setting to avoid overfitting and used the following features:

- Demographics: Log of the median income (from the U.S. Census), Education Index, % of the Gallup sample that identifies as: Black (since they had the highest pain scores), married, and female from Gallup data.
- Language Feature-1: Normalized frequency of 73 LIWC-2015 categories (e.g. Affect, Pronouns, Cognitive Processes etc.) [52].
- Language Feature-2: Frequency of 1grams obtained from the CTLB [28]
- Language Feature-3: Probabilities of the 75 pain topics generated in Section 4

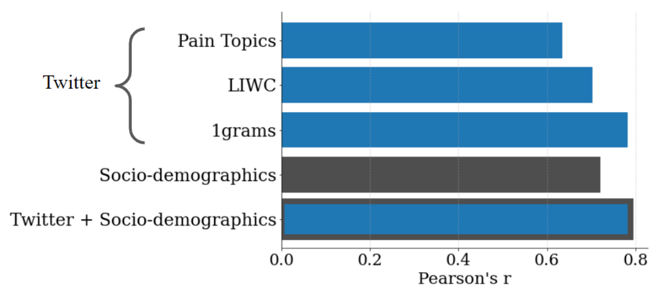

Figure 2: Performance of models predicting self-reported pain from the Gallup data. For each model, the figure shows the Pearson correlation between the predicted pain and actual pain. Twitter + Socio-Demographics Model is significantly better than Demographics only model (p < 0.001)

As can be seen in Figure 2, Twitter language feature (1grams) outperformed demographic variables at predicting community pain by over 9% (Pearson r of .78 vs .72), showing significantly increased predicting power over demographics. In fact, 1grams outperformed pain topics and LIWC categories. Providing the model with demographics and all language-related features obtained the best performance (r=0.79). Twitter language features themselves captured most of the explainable variance (r=0.78) with comparable performance to the model with demographics and all language features (δr=0.006). The results strongly support the use of social media language in measuring community-level pain. We believe that the performance of Twitter language (1grams) is due to its ability to capture a variety of parameters (*heart disease, life satisfaction, education*) [28].

6 VARIATION IN EXPRESSIONS OF PAIN ON TWITTER

The subset of tweets identified as expressions of pain by the classifier in Sec 3.4.1 was used to model themes related to pain and then applied on a large-scale county-level Twitter data set to then identify statistically significant differences across communities in the U.S.

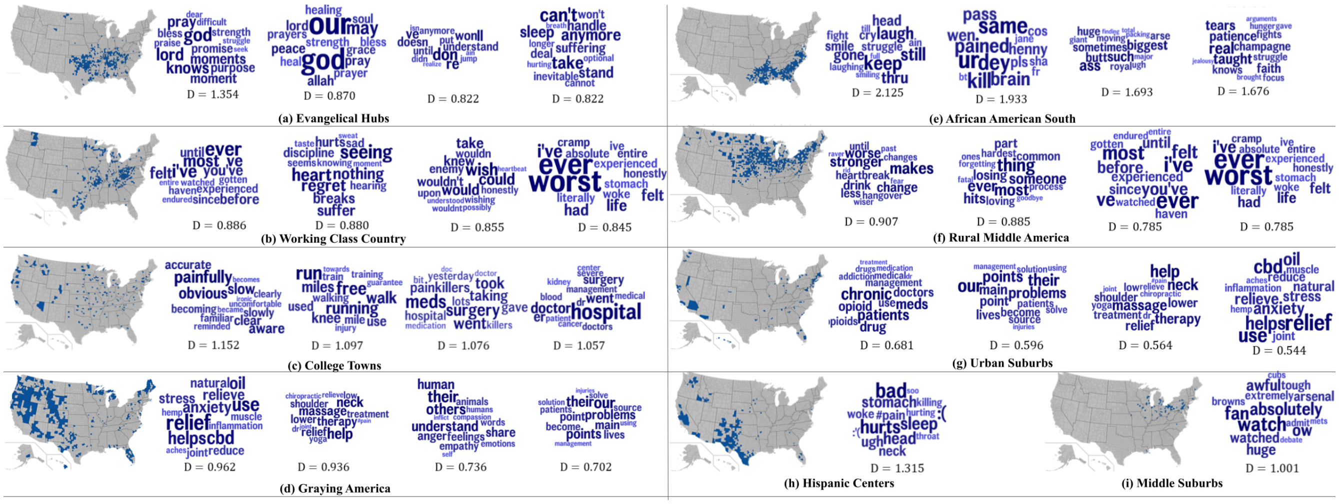

Figure 3: Top 4 pain topics positively correlated with each of the ACPs. We reported Cohen's D values (i.e., normalized mean differences) with $p < 0.05$ (as determined via a Logistic regression) after the Benajamini-Hochberg correction and controlled for SES (income and education). The list of all significant topics per ACP is shown in Supplmentary Table A4.

We extracted the probability of the occurrence of each of the 75 pain topics for every county in the CTLB. Topics were then used as input in a Logistic regression model with dummy variables for each community in ACP as outcomes. Effect sizes were calculated by Cohen's D values (i.e., the mean difference between the outcome groups divided by the pooled standard deviation) and significance was determined using the p-value returned from the Logistic regression with a significance threshold of .05, after adjusting for multiple comparisons using Benjamini-Hochberg's multi-test correction [4]. We did not consider the Native American Lands and Aging Farmlands communities since neither was well represented in the data (only 7 counties in the data set are in the Native American Lands classification while no counties are in the Aging Farmlands; see Table A3 for full counts).

As seen in Figure 3, Evangelical Hubs frequently post about God, lessons from struggle, and prayers when expressing their pain. The words such as *suffering, anymore, stand, handle, can't* are also observed indicating lacking pain management. Working Class Country discusses *regret, discipline, suffering* and qualifier words such as *worst, ever, absolute, literally* indicating the intensity of the experienced pain. Academic stressors, sports injuries, painkillers, and surgeries were the most commonly discussed pain themes in College Towns. It is also interesting to note that College towns frequently refer to medical terms such as *hospital, doctor, surgery, blood* when discussing pain. This is not observed in other counties even when the pain seems to be unbearable and impairing. Graying America discussed therapies (*massage, chiropractic, yoga*), used emotional and empathetic language when discussing pain indicating more acceptance towards pain, and posted alternatives (*cbd, oil, hemp*) to opioids for dealing with chronic pain. the African American South posted about struggles, patience, and faith.

Rural Middle America discussed *heartbreak, hangover, forgetting, losing*. We speculate that this could be either due to the misclassification of emotional pain tweets or somatic symptoms experienced when dealing with personal loss. Nevertheless, the experienced pain is described intensely (*worst, absolute, ever*). Urban Suburbs posted about *opioids, cbd oil,* and *therapy* when talking about pain. Pain topics in Urban Suburbs are similar to ones observed in Graying America (*massage, yoga, therapy, cbd, oil*). This could be reflective of chronic pains associated with old age, being experienced by the younger population. We found only one topic significantly associated with Hispanic Centers and Middle Suburbs. Hispanic Centres' topic is clearly about bodily pains (*neck, stomach, head, throat*). Middle Suburbs' pain topic is sports oriented and it is likely the case of misclassification of emotional pain tweets.

7 DISCUSSION

Our experiments demonstrate the efficacy of machine learning models built using language expressions of pain on Twitter for reliably predicting self-reports on the experience of pain from Gallup. This paves the way for using social media as an effective and low-cost tool to measure pain in a community, monitor pain levels over time, and understand potential reasons for pain across communities to provide access to personalized pain management and care. Additionally, not only can social media be used to measure pain, but it can also be used to distinguish between emotional and physical pain, further allowing for targeted public health programs [17].

The topics derived from CSLDA reveal the variation in the content of tweets about pain across different American Communities. For instance, Hispanic centers specifically discuss *stomach, throat, neck, body* related pain indicating infections whereas Urban Suburbs pain topics are very similar to Graying America which is a cause of concern. The awareness of medical support in College Towns is reassuring but also raises a question on the lack of such medical terms in other communities. Several of these factors could also be explained through mixed interactions. For instance, the higher prevalence of individuals with a postgraduate degree in College Towns could influence better access to care and lead to lower pain levels Nevertheless, these insights can potentially suggest a personalized cause and approach to pain and pain management among different communities [55] which could inform personalized interventions and informed public policies [15].

We also note pain topics due to misclassified tweets, indicating emotional pain or rhetorical pain associated with excitement/ enjoyment. For our study, we used a naive Bag of words approach to build an interpretable model. However, the classification performance can indeed be improved using state-of-the-art NLP models capable of capturing the underlying intended sense of the word *pain*.

The use of social media in healthcare could help government agencies target under-served communities and increase equity in healthcare, as has been seen during the COVID-19 pandemic [47]. Tracking levels and expressions of pain at the community level has a number of public health applications. Of note are the rising mortality rates across the U.S., which have coincided with increased reports of physical pain [7]. Another public health crisis, the U.S. opioid epidemic, has received considerable research attention, with over-prescribing by pharmaceutical companies and physicians believed to drive increasing mortality rates [44]. While this "supply" side of the opioid epidemic is certainly a factor in opioid availability, it ignores root causes such as physical and psychological determinants of health (i.e., the "demand" side), of which pain may be a contributing factor. Finally, socioeconomic hardships, lack of health care, and other social determinants of health have all been shown to increase the risk of physical pain [69]. This work fits in with several studies which attempt to use digital data sources to access regional stressors and their links to public health [1, 19, 25, 42].

7.1 Limitations and Future Work

In this work, we studied a cross-section sample of Twitter and Gallup data without considering the temporal variance associated with experiences of pain. Social media could be used to measure the intensity of pain and its variation over time as it has been used to measure sentiment during the COVID-19 pandemic [32, 57]. To formulate an even deeper understanding of pain and its cultural differences, it would be integral to understand how the expression of pain has varied over time if there have been particular changes due to notable events, especially among specific communities (e.g. the impact of George Floyd murder on African Americans [20]). Further, we only considered linguistic features when analyzing Twitter data in this study; exploring social network features could further enhance our understanding of the relationship between online social capital and the science of pain, which has been found to be a significant indicator among patient cohorts [3, 38]. Additionally, question wording in the Gallup surveys could have also introduced bias in our data, as self-reports of pain are subjective.

While we only used Twitter for our analysis, several studies have found that the social media platforms such as Facebook and more recently, TikTok, attract a larger audience and have richer health-related information [37]. However, it is also relatively difficult to collect data from these platforms due to API and privacy constraints. Community-specific social media platforms such as WeChat and Weibo would also limit bias. We limited our analysis to English tweets however, given the breadth of languages in the United States, it may be useful to extend our model to different languages, particularly non-Western Educated Industrial Rich Democratic (WEIRD) samples [39] to uncover the varying cultural norms around pain [11, 41]. Last but not the least, the barriers to internet access in under-served communities can limit data points and restrict effective healthcare monitoring and pain management in these areas.

7.2 Ethical Considerations and Broader Impact

Using social media for research has various ethical implications, particularly when the proposed technology stack may have a role to play in public health and policy making. It is therefore imperative to have safeguards that preserve and respect user privacy, agency, and consent when collecting public data which may divulge health stats and other personal information. In this study, we only considered county-level data having a certain number of users to prevent any user-level identifiable details from leaking through. While this data would be beneficial to the academic community and will be released in compliance with Twitter's Terms of Service, we will ensure to not make the user-level data public.

REFERENCES

[1] Rediet Abebe, Salvatore Giorgi, Anna Tedijanto, Anneke Buffone, and H. Andrew Schwartz. 2020. Quantifying Community Characteristics of Maternal Mortality Using Social Media. In *The World Wide Web Conference*.

[2] Mohammad Al-Harthy, Richard Ohrbach, A Michelotti, and Thomas List. 2016. The effect of culture on pain sensitivity. *Journal of oral rehabilitation* 43, 2 (2016), 81–88.

[3] Cecilia Åslund, Bengt Starrin, and Kent W Nilsson. 2010. Social capital in relation to depression, musculoskeletal pain, and psychosomatic symptoms: a cross-sectional study of a large population-based cohort of Swedish adolescents. *BMC Public Health* 10, 1 (2010), 1–10.

[4] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57, 1 (1995), 289–300. https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1995.tb02031.x

[5] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.

[6] Claudia M Campbell and Robert R Edwards. 2012. Ethnic differences in pain and pain management. *Pain management* 2, 3 (2012), 219–230.

[7] Anne Case and Angus Deaton. 2015. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proceedings of the National Academy of Sciences* 112, 49 (2015), 15078–15083.

[8] Anne Case, Angus Deaton, and Arthur A Stone. 2020. Decoding the mystery of American pain reveals a warning for the future. *Proceedings of the National Academy of Sciences* 117, 40 (2020), 24785–24789.

[9] Dante Chinni and Ari Pinkus. 2019. A New Portrait of Rural America. *American Communities Project* (2019).

[10] David R Cox. 1958. The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)* 20, 2 (1958), 215–232.

[11] Jesse Cui, Tingdan Zhang, Kokil Jaidka, Dandan Pang, Garrick Sherman, Vinit Jakhetiya, Lyle H Ungar, and Sharath Chandra Guntuku. 2022. Social Media Reveals Urban-Rural Differences in Stress across China. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 114–124.

[12] David C Currow, Jane Phillips, and Katherine Clark. 2016. Using opioids in general practice for chronic non-cancer pain: an overview of current evidence. *Medical Journal of Australia* 204, 8 (2016), 305–309.

[13] James Gimpel Dante Chinni. 2011. *Our Patchwork Nation: The Surprising Truth about the" Real" America.* Penguin, New York, NY, USA.

[14] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th annual ACM web science conference*. 47–56.

[15] Colette M DeMonte, William D DeMonte, and Beverly E Thorn. 2015. Future implications of eHealth interventions for chronic pain management in underserved populations. *Pain management* 5, 3 (2015), 207–214.

[16] Matthew R DeVerna, Francesco Pierri, Bao Tran Truong, John Bollenbacher, David Axelrod, Niklas Loynes, Christopher Torres-Lugo, Kai-Cheng Yang, Filippo Menczer, and John Bryden. 2021. CoVaxxy: A Collection of English-Language Twitter Posts About COVID-19 Vaccines.. In *ICWSM*. 992–999.

[17] Adam G Dunn, Kenneth D Mandl, and Enrico Coiera. 2018. Social media interventions for precision public health: promises and risks. *NPJ digital medicine* 1, 1 (2018), 1–4.

[18] Christopher L Edwards, Roger B Fillingim, and Francis Keefe. 2001. Race, ethnicity and pain. *Pain* 94, 2 (2001), 133–137.

[19] Johannes C Eichstaedt, H Andrew Schwartz, Margaret L Kern, Gregory Park, Darwin R Labarthe, Raina M Merchant, Sneha Jha, Megha Agrawal, Lukasz A Dziurzynski, Maarten Sap, Christopher Weeg, Emily E Larson, Lyle H Ungar, and Martin EP Seligman. 2015. Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science* 26 (2015), 159–169. Issue 2.

[20] Johannes C Eichstaedt, Garrick T Sherman, Salvatore Giorgi, Steven O Roberts, Megan E Reynolds, Lyle H Ungar, and Sharath Chandra Guntuku. 2021. The emotional and mental health impact of the murder of George Floyd on the US population. *Proceedings of the National Academy of Sciences* 118, 39 (2021), e2109139118.

[21] Stephen E Fienberg et al. 1970. An iterative procedure for estimation in contingency tables. *The Annals of Mathematical Statistics* 41, 3 (1970), 907–917.

[22] Krzysztof Fiok, Waldemar Karwowski, Edgar Gutierrez, Maham Saeidi, Awad M Aljuaid, Mohammad Reza Davahli, Redha Taiar, Tadeusz Marek, and Ben D Sawyer. 2021. A study of the effects of the COVID-19 pandemic on the experience of back pain reported on Twitter® in the United States: a natural language processing approach. *International Journal of Environmental Research and Public Health* 18, 09 (2021), 4543.

[23] Mary Moore Free. 2002. Cross-cultural conceptions of pain and pain control. In *Baylor University Medical Center Proceedings*, Vol. 15. Taylor & Francis, 143–145.

[24] Pierre Geurts, Damien Ernst, and Louis Wehenkel. 2006. Extremely randomized trees. *Machine learning* 63, 1 (2006), 3–42.

[25] Joseph Gibbons, Robert Malouf, Brian Spitzberg, Lourdes Martinez, Bruce Appleyard, Caroline Thompson, Atsushi Nara, and Ming-Hsiang Tsou. 2019. Twitter-based measures of neighborhood sentiment as predictors of residential population health. *PloS one* 14, 7 (2019), e0219550.

[26] Salvatore Giorgi, Sharath Chandra Guntuku, Mckenzie Himelein-Wachowiak, Amy Kwarteng, Sy Hwang, Muhammad Rahman, and Brenda Curtis. 2022. Twitter Corpus of the #BlackLivesMatter Movement and Counter Protests: 2013 to 2021. *Proceedings of the International AAAI Conference on Web and Social Media* 16, 1 (May 2022), 1228–1235. https://ojs.aaai.org/index.php/ICWSM/article/view/19373

[27] Salvatore Giorgi, Khoa Le Nguyen, Johannes C Eichstaedt, Margaret L Kern, David B Yaden, Michal Kosinski, Martin EP Seligman, Lyle H Ungar, H Andrew Schwartz, and Gregory Park. 2021. Regional personality assessment through social media language. *Journal of Personality* (2021).

[28] Salvatore Giorgi, Daniel Preotiuc-Pietro, Anneke Buffone, Daniel Rieman, Lyle H. Ungar, and H. Andrew Schwartz. 2018. The Remarkable Benefit of User-Level Aggregation for Lexical-based Population-Level Predictions. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*.

[29] Salvatore Giorgi, David B Yaden, Johannes C Eichstaedt, Robert D Ashford, Anneke EK Buffone, H Andrew Schwartz, Lyle H Ungar, and Brenda Curtis. 2020. Cultural Differences in Tweeting about Drinking Across the US. *International Journal of Environmental Research and Public Health* 17, 4 (2020), 1125.

[30] Heather Griffis, David A Asch, H Andrew Schwartz, Lyle Ungar, Alison M Buttenheim, Frances K Barg, Nandita Mitra, Raina M Merchant, et al. 2020. Using social media to track geographic variability in language about diabetes: Infodemiology analysis. *JMIR diabetes* 5, 1 (2020), e14431.

[31] Sharath Chandra Guntuku, Anneke Buffone, Kokil Jaidka, Johannes C Eichstaedt, and Lyle H Ungar. 2019. Understanding and measuring psychological stress using social media. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 214–225.

[32] Sharath Chandra Guntuku, Garrick Sherman, Daniel C Stokes, Anish K Agarwal, Emily Seltzer, Raina M Merchant, and Lyle H Ungar. 2020. Tracking mental health and symptom mentions on Twitter during COVID-19. *Journal of general internal medicine* 35, 9 (2020), 2798–2800.

[33] Laura Hawks, Candace Cosgrove, Mathew Neiman, Brita Roy, Christopher Wildeman, Sean Coady, and Emily A Wang. 2020. Five-Year Mortality among Americans Incarcerated in Privatized Versus Public Prisons: the Mortality Disparities in American Communities Project. *Journal of General Internal Medicine* (2020), 1–4.

[34] N Heaivilin, B Gerbert, JE Page, and JL Gibbs. 2011. Public health surveillance of dental pain via Twitter. *Journal of dental research* 90, 9 (2011), 1047–1051.

[35] Tin Kam Ho. 1995. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, Vol. 1. IEEE, 278–282.

[36] Kokil Jaidka, Salvatore Giorgi, H Andrew Schwartz, Margaret L Kern, Lyle H Ungar, and Johannes C Eichstaedt. 2020. Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences* 117, 19 (2020), 10165–10171.

[37] Kokil Jaidka, Sharath Guntuku, and Lyle Ungar. 2018. Facebook versus Twitter: Differences in self-disclosure and trait prediction. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 12.

[38] Kenneth Jay and Lars L Andersen. 2018. Can high social capital at the workplace buffer against stress and musculoskeletal pain?: Cross-sectional study. *Medicine* 97, 12 (2018).

[39] Dan Jones. 2010. A WEIRD view of human nature skews psychologists' studies.

[40] Dan Jurafsky, Victor Chahuneau, Bryan R Routledge, and Noah A Smith. 2014. Narrative framing of consumer sentiment in online restaurant reviews. *First Monday* (2014).

[41] Mingyang Li, Louis Hickman, Louis Tay, Lyle Ungar, and Sharath Chandra Guntuku. 2020. Studying politeness across cultures using English Twitter and Mandarin Weibo. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–15.

[42] Yuchen Li, Ayaz Hyder, Lauren T Southerland, Gretchen Hammond, Adam Porr, and Harvey J Miller. 2020. 311 service requests as indicators of neighborhood distress and opioid use disorder. *Scientific reports* 10, 1 (2020), 1–11.

[43] John Logan. 2006. Population displacement and post-Katrina politics: The New Orleans mayoral race, 2006. *American Communities Project* (2006).

[44] Bertha K Madras. 2017. The surge of opioid use, addiction, and overdoses: responsibility and response of the US health care system. *JAMA psychiatry* 74, 5 (2017), 441–442.

[45] Samantha M Meints, Alejandro Cortes, Calia A Morais, and Robert R Edwards. 2019. Racial and ethnic differences in the experience and treatment of noncancer pain. *Pain management* 9, 3 (2019), 317–334.

[46] Ronald Melzack. 1990. The tragedy of needless pain. *Scientific American* 262, 2 (1990), 27–33.

[47] Raina M Merchant, Eugenia C South, and Nicole Lurie. 2021. Public health messaging in an era of social media. *Jama* 325, 3 (2021), 223–224.

[48] Lewis Mitchell, Morgan R Frank, Kameron Decker Harris, Peter Sheridan Dodds, and Christopher M Danforth. 2013. The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PloS one* 8, 5 (2013), e64417.

[49] Feng Nan, Ran Ding, Ramesh Nallapati, and Bing Xiang. 2019. Topic modeling with wasserstein autoencoders. *arXiv preprint arXiv:1907.12374* (2019).

[50] International Pain Summit of the International Association for the Study of Pain et al. 2011. Declaration of Montréal: declaration that access to pain management is a fundamental human right. *Journal of pain & palliative care pharmacotherapy* 25, 1 (2011), 29–31.

[51] Cara Ostrom, Eric Bair, William Maixner, Ronald Dubner, Roger B Fillingim, Richard Ohrbach, Gary D Slade, and Joel D Greenspan. 2017. Demographic predictors of pain sensitivity: results from the OPPERA study. *The Journal of Pain* 18, 3 (2017), 295–307.

[52] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015.* Technical Report.

[53] Mark H Pitcher, Michael Von Korff, M Catherine Bushnell, and Linda Porter. 2019. Prevalence and profile of high-impact chronic pain in the United States. *The Journal of Pain* 20, 2 (2019), 146–160.

[54] Stanley Rachman and Arnold Arntz. 1991. The overprediction and underprediction of pain. *Clinical Psychology Review* 11, 4 (1991), 339–355.

[55] Fatemeh Ramezani, Bijan Riazi Farzad, and Atousa Janzadeh. 2021. Personalized Pain Medicine: Turning Theory into Policy. *Archives of Neuroscience* 8, 4 (2021).

[56] Radim Rehurek and Petr Sojka. 2010. Software framework for topic modelling with large corpora. In *In Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks.* Citeseer.

[57] Koustuv Saha, John Torous, Eric D Caine, Munmun De Choudhury, et al. 2020. Psychosocial effects of the COVID-19 pandemic: large-scale quasi-experimental study on social media. *Journal of medical internet research* 22, 11 (2020), e22600.

[58] Hansen Schwartz, Johannes Eichstaedt, Margaret Kern, Lukasz Dziurzynski, Richard Lucas, Megha Agrawal, Gregory Park, Shrinidhi Lakshmikanth, Sneha Jha, Martin Seligman, et al. 2013. Characterizing geographic variation in well-being using tweets. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 7. 583–591.

[59] H Andrew Schwartz, Salvatore Giorgi, Maarten Sap, Patrick Crutchley, Lyle Ungar, and Johannes Eichstaedt. 2017. Dlatk: Differential language analysis toolkit. In *Proceedings of the 2017 conference on empirical methods in natural language processing: System demonstrations.* 55–60.

[60] Vickie L Shavers, Alexis Bakos, and Vanessa B Sheppard. 2010. Race, ethnicity, and pain among the US adult population. *Journal of health care for the poor and underserved* 21, 1 (2010), 177–220.

[61] Shaheen Syed and Marco Spruit. 2017. Full-text or abstract? examining topic coherence scores using latent dirichlet allocation. In *2017 IEEE International conference on data science and advanced analytics (DSAA).* IEEE, 165–174.

[62] Patrick J Tighe, Ryan C Goldsmith, Michael Gravenstein, H Russell Bernard, and Roger B Fillingim. 2015. The painful tweet: text, sentiment, and community structure analyses of tweets pertaining to pain. *Journal of medical Internet research* 17, 4 (2015), e84.

[63] Sacide Yildizeli Topcu. 2018. Relations among pain, pain beliefs, and psychological well-being in patients with chronic pain. *Pain Management Nursing* 19, 6 (2018), 637–644.

[64] Dennis C Turk and Ronald Melzack. 2011. The measurement of pain and the assessment of people experiencing pain. (2011).

[65] Emily O Wakefield, William T Zempsky, Rebecca M Puhl, and Mark D Litt. 2018. Conceptualizing pain-related stigma in adolescent chronic pain: A literature review and preliminary focus group findings. *Pain reports* 3, Suppl 1 (2018).

[66] Tao Wang, Markus Brede, Antonella Ianni, and Emmanouil Mentzakis. 2017. Detecting and characterizing eating-disorder communities on social media. In *Proceedings of the Tenth ACM International conference on web search and data mining*. 91–100.

[67] George Ward, Jan-Emmanuel De Neve, Lyle H Ungar, and Johannes C Eichstaedt. 2021. (Un) happiness and voting in US presidential elections. *Journal of personality and social psychology* 120, 2 (2021), 370.

[68] Timothy H Wideman, Robert R Edwards, David M Walton, Marc O Martel, Anne Hudon, and David A Seminowicz. 2019. The multimodal assessment model of pain: a novel framework for further integrating the subjective pain experience within research and practice. *The Clinical journal of pain* 35, 3 (2019), 212.

[69] Steven H Woolf and Heidi Schoomaker. 2019. Life expectancy and mortality rates in the United States, 1959-2017. *Jama* 322, 20 (2019), 1996–2016.

[70] R Jason Yong, Peter M Mullins, and Neil Bhattacharyya. 2022. Prevalence of chronic pain among adults in the United States. *Pain* 163, 2 (2022), e328–e332.

[71] Mohammadzaman Zamani, H. Andrew Schwartz, Johannes Eichstaedt, Sharath Chandra Guntuku, Adithya Virinchipuram Ganesan, Sean Clouston, and Salvatore Giorgi. 2020. Understanding Weekly COVID-19 Concerns through Dynamic Content-Specific LDA Topic Modeling. In *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*. Association for Computational Linguistics, Online, 193–198. https://www.aclweb.org/anthology/2020.nlpcss-1.21

A APPENDIX

A.1 Pain related keywords

The following keywords were used to identify tweets related to pain in Section 3.3: pain, headache ache, heartache, headache, toothache, stomache, stomachache, earache, backache, toothache.

A.2 Classification Models for Pain

Table A1: Evaluation Metrics for Pain Classification

Model	Features	AuC	F1	Precision	Recall
LR	All	0.83	**0.71**	0.71	**0.71**
	1grams	0.81	0.70	0.72	0.68
	LIWC	0.81	0.66	0.74	0.63
	Topics	0.82	0.70	0.79	0.66
RFC	All	**0.86**	0.64	0.85	0.61
	1grams	0.84	0.65	**0.90**	0.62
	LIWC	0.81	0.66	0.74	0.63
	Topics	0.82	0.70	0.79	0.66
ETC	All	**0.86**	0.66	0.84	0.62
	1grams	0.85	0.67	0.86	0.63
	LIWC	0.81	0.66	0.74	0.63
	Topics	0.82	0.70	0.79	0.66

Table A2: Evaluation Metrics for Pain Type (Physical vs Emotional Pain)

Model	Features	AuC	F1	Precision	Recall
LR	All	0.71	**0.64**	0.63	**0.64**
	1grams	0.68	0.63	0.63	0.63
	LIWC	0.71	0.57	0.64	0.57
	Topics	**0.76**	0.61	0.65	0.60
RFC	All	**0.76**	0.48	**0.73**	0.52
	1grams	0.72	0.47	0.66	0.52
	LIWC	0.71	0.57	0.64	0.57
	Topics	**0.76**	0.61	0.65	0.60
ETC	All	**0.76**	0.49	0.72	0.53
	1grams	0.72	0.49	0.64	0.53
	LIWC	0.71	0.57	0.64	0.57
	Topics	**0.76**	0.61	0.65	0.60

Table A3: Counties per ACP in our dataset.

ACP	Number of Counties
African American South	157
Aging Farmlands	0
Big Cities	46
College Towns	140
Evangelical Hubs	198
Exurbs	199
Graying America	135
Hispanic Centers	59
LDS Enclaves	14
Middle Suburbs	76
Military Posts	67
Native American Lands	7
Rural Middle America	312
Urban Suburbs	100
Working Class Country	131

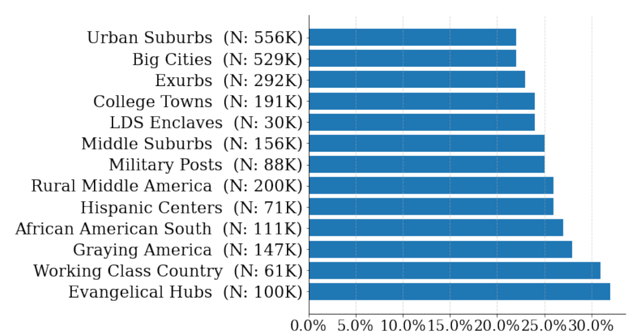

Figure 4: Percentage of people who reported pain through Gallup across different American Communities, as defined by [13]

A.3 Gallup Pain Measures across ACPs

Table A3 shows the counts of counties per ACP. Two ACPs–Native American Lands and Aging Farmlands–had less than 10 counties and were excluded from our analyses.

A.4 Pain LDA Topics

A.4.1 Content Specific LDA. Here we discuss the specifics of the CSLDA pipeline used to create the pain topics. For full details, please see Zamani et al. [71].

In order to find which words are associated with pain, we create two corpora: one of pain tweets and one of non-pain tweets. To do this, we select 1 million random U.S. county-mapped tweets which

Table A4: (Supplementary): Top 10 Topics significantly associated with each American Community. Effect sizes are reported as Cohen's D values. All topics are significant at $p < 0.001$ (via the Logistic regression model) after Benjamini-Hochberg multi-test correction. Six ACPs did not have any significant topic associations and are excluded from the table.

Topic ID	Top Words in the Topic	Cohen's d
African American South		
7	keep, still, laugh, thru, gone, head, smile, cry, struggle, fight	2.125
20	ur, same, dey, kill, pained, brain, pass, wen, henny, pls	1.933
29	ass, such, biggest, butt, sometimes, huge, arse, royal, moving, ugh	1.693
24	real, taught, tears, patience, faith, champagne, knows, fights, struggle, focus	1.676
42	take, pills, killers, meds, these, took, taking, pill, medicine, pop	1.489
64	god, lord, knows, pray, moments, moment, purpose, promise, bless, strength	1.291
46	stomach, having, these, chest, getting, sharp, bad, :(, ugh, sick	1.101
16	bad, hurts, :(, sleep, head, stomach, ugh, neck, #pain, woke	1.065
49	cause, pleasure, same, heal, self, meant, passion, feels, mans, wounds	0.925
58	bring, let, rain, drugs, could, wash, joy, brings, yesterday, shame	0.870
College Towns		
69	painfully, obvious, slow, aware, clear, accurate, slowly, becoming, familiar, clearly	1.152
28	run, free, running, walk, miles, knee, use, used, mile, walking	1.097
3	meds, surgery, went, took, taking, painkillers, gave, hospital, lots, yesterday	1.076
21	hospital, doctor, er, went, surgery, dr, medical, severe, blood, management	1.057
2	knee, shoulder, surgery, nerve, neck, hip, leg, muscle, injury, lower	1.038
74	very, many, different, times, also, ve, both, lot, which, growing	1.034
19	years, ago, had, still, after, months, since, few, almost, lost	0.987
36	trump, suffering, caused, political, india, election, inflict, economic, govt, politics	0.894
53	watch, fan, absolutely, ow, awful, huge, watched, arsenal, tough, extremely	0.876
4	days, few, these, past, weeks, months, years, hours, times, took	0.833
Evangelical Hubs		
64	god, lord, knows, pray, moments, moment, purpose, promise, bless, strength	1.354
71	our, god, may, lord, pray, allah, peace, heal, bless, strength	0.870
15	can't, take, anymore, sleep, stand, handle, suffering, deal, won't, inevitable	0.822
70	don, re, ll, ve, won, doesn, understand, until, put, anymore	0.822
11	even, imagine, can't, feeling, words, cannot, kind, describe, joy, explain	0.798
33	wish, could, would, take, knew, wouldn't, enemy, wouldn, upon, honestly	0.765
35	i've, had, ve, felt, before, myself, lot, times, changed, many	0.753
14	forever, temporary, worth, long, term, lasts, end, beauty, remember, glory	0.736
65	let, hold, forget, past, memories, remember, ends, caused, forgive, changes	0.724
56	heart, deep, broken, inside, full, soul, heal, carry, bones, apart	0.711
Graying America		
1	relief, use, cbd, helps, oil, anxiety, relieve, stress, natural, reduce	0.962
45	help, neck, massage, therapy, relief, lower, shoulder, treatment, low, yoga	0.936
59	their, others, understand, human, share, anger, feelings, empathy, animals, words	0.736
23	our, their, points, problems, main, point, lives, become, source, patients	0.702
25	chronic, patients, meds, drug, opioid, doctors, use, opioids, management, medical	0.619
52	their, our, those, children, ones, lives, lost, families, loved, many	0.613
41	chronic, #chronicpain, disease, health, issues, living, illness, severe, depression, cancer	0.612
67	life, suffering, world, joy, death, choose, happiness, fear, living, full	0.580
73	tattoo, getting, tolerance, painless, tattoos, low, skin, quick, scale, level	0.572
66	learn, change, strength, growth, lesson, growing, experience, process, grow, part	0.526
Hispanic Centers		
16	bad, hurts, :(, sleep, head, stomach, ugh, neck, #pain, woke	1.315
Middle Suburbs		
53	watch, fan, absolutely, ow, awful, huge, watched, arsenal, tough, extremely	1.001
Rural Middle America		
12	makes, stronger, worse, change, drink, less, heartbreak, until, hangover, changes	0.907
34	thing, most, ever, someone, hits, losing, part, common, loving, hardest	0.885
0	ever, i've, most, ve, felt, you've, before, until, since, experienced	0.785
10	worst, ever, i've, felt, had, life, stomach, experienced, literally, woke	0.785
27	seeing, heart, regret, nothing, suffer, breaks, hurts, discipline, sad, hearing	0.783
72	other, each, any, enough, deserve, different, kind, strong, side, type	0.749
48	period, birth, giving, labor, child, cramps, during, kidney, control, mother	0.749
37	after, legs, gym, workout, gain, sore, #nopainnogain, yesterday, leg, working	0.720
9	am, physically, myself, mentally, emotionally, pained, sick, kind, grateful, setting	0.689
53	watch, fan, absolutely, ow, awful, huge, watched, arsenal, tough, extremely	0.685
Urban Suburbs		
25	chronic, patients, meds, drug, opioid, doctors, use, opioids, management, medical	0.681
23	our, their, points, problems, main, point, lives, become, source, patients	0.596
45	help, neck, massage, therapy, relief, lower, shoulder, treatment, low, yoga	0.564
1	relief, use, cbd, helps, oil, anxiety, relieve, stress, natural, reduce	0.544
41	chronic, #chronicpain, disease, health, issues, living, illness, severe, depression, cancer	0.533
21	hospital, doctor, er, went, surgery, dr, medical, severe, blood, management	0.505
Working Class Country		
0	ever, i've, most, ve, felt, you've, before, until, since, experienced	0.886
27	seeing, heart, regret, nothing, suffer, breaks, hurts, discipline, sad, hearing	0.880
33	wish, could, would, take, knew, wouldn't, enemy, wouldn, upon, honestly	0.855
10	worst, ever, i've, felt, had, life, stomach, experienced, literally, woke	0.845
34	thing, most, ever, someone, hits, losing, part, common, loving, hardest	0.827
39	someone, again, numb, alone, can't, once, feels, remember, gone, goodbye	0.808
12	makes, stronger, worse, change, drink, less, heartbreak, until, hangover, changes	0.803
11	even, imagine, can't, feeling, words, cannot, kind, describe, joy, explain	0.799
13	death, die, slow, hope, painfully, dying, slowly, kill, dies, suffer	0.796
50	being, without, able, move, walk, sick, crying, decided, barely, tear	0.765

have been classified as related to physical pain by our classifier. We also select 1 million random U.S. county-mapped tweets which do *not* mention the keyword pain. We then combine the two corpora into a single corpus consisting of 2 million tweets and assign each tweet with a binary label of 1 for pain tweets and 0 otherwise. Next, we tokenize the tweets using a tokenizer designed for social media text [59] which produces 1,752,144 distinct 1grams. Since this number is larger than the number of tweets in our corpus (2 million), we remove any 1gram which has been used by less than 0.01% of the tweets (i.e., 1grams contained in less than 20,000 tweets). This resulted in a final feature set of 7,248 1grams.

Next, using the binary pain/not pain outcome, we calculated a weighted log odds ratio using an Informative Dirichlet prior [40]. This method estimates the difference in word frequency across the two corpora (pain and non-pain tweets) using a prior which shrinks each pain word frequency towards known frequencies in the matched sample. We then selected the 2,500 1grams most associated with the binary pain label. Finally, we filtered the 1,809,805 U.S. county-mapped pain tweets (i.e., tweets classified as mentioning physical pain) to contain only these 2,500 1grams. We then ran the standard LDA algorithm over this data set, using the Mallet Java package and all standard defaults except α (a prior on the number of topics per document) which we set to 2 since tweets are shorter than standard documents used in this setting.

A.4.2 Topic Uniqueness. The topic uniqueness (TU) metric considers a set of L keywords across the K topics. Specifically, TU measures how often these K keywords are repeated across each of the topics. Typically TU scores are bounded between 1 (high uniqueness) and $1/K$ (low uniqueness. Given that we are comparing TU across topic sets of varying size (i.e., varying K), we normalize all TU scores to be between 0 and 1. Nan et al. [49] set $L = 10$ which we increase to $L = 30$ in order to give a more conservative estimate, since as L increases the probability of a given word appearing in more than one topic also increases, thus driving down TU.