frontiers | Research Topics

Advances and challenges in pain measurement, 2nd Edition

Edited by
Kenneth Craig, Kai Karos and Loren Martin

Published in
Frontiers in Pain Research

ISSN 1664-8714
ISBN 978-2-8325-3328-4
DOI 10.3389/978-2-8325-3328-4

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Advances and challenges in pain measurement, 2nd Edition

Topic editors

Kenneth Craig — University of British Columbia, Canada
Kai Karos — Open University of the Netherlands, Netherlands
Loren Martin — University of Toronto, Canada

Citation

Craig, K., Karos, K., Martin, L., eds. (2023). *Advances and challenges in pain measurement, 2nd Edition*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3328-4

Publisher's note: In this 2nd edition, the following articles have been added: Bucolo M, Rance M, Nees F, Ruttorf M, Stella G, Monarca N, Andoh J and Flor H (2022) Cortical networks underlying successful control of nociceptive processing using real-time fMRI. *Front. Pain Res*. 3:969867. doi: 10.3389/fpain.2022.969867 ; Sanmugananthan VV, Cheng JC, Hemington KS, Rogachov A, Osborne NR, Bosma RL, Kim JA, Inman RD and Davis KD (2023) Can we characterize A-P/IAP behavioural phenotypes in people with chronic pain?. *Front. Pain Res*. 4:1057659. doi: 10.3389/fpain.2023.1057659 ; and Kuhlmann L, Olesen SS and Drewes AM (2022) Assessment of visceral pain with special reference to chronic pancreatitis. *Front. Pain Res*. 3:1067103. doi: 10.3389/fpain.2022.1067103.

Table of contents

frontiers in Pain Research

ORIGINAL RESEARCH
published: 12 August 2021
doi: 10.3389/fpain.2021.692567

Beyond Average: Providers' Assessments of Indices for Measuring Pain Intensity in Patients With Chronic Pain

Roberta E. Goldman[1], Joan E. Broderick[2], Doerte U. Junghaenel[2], Alicia Bolton[2], Marcella May[2], Stefan Schneider[2†] and Arthur A. Stone[2†]*

[1] *Department of Family Medicine, Warren Alpert Medical School of Brown University, Providence, RI, United States,* [2] *Dornsife Center for Self-Report Science, University of Southern California, Los Angeles, CA, United States*

OPEN ACCESS

Edited by:
Kenneth Craig, University of British Columbia, Canada

Reviewed by:
Yang Wang, University of Maryland, United States
Kai Karos, Centre for the Psychology of Learning and Experimental Psychopathology, KU Leuven, Belgium

***Correspondence:**
Roberta E. Goldman
Roberta_Goldman@brown.edu

† These authors share senior authorship

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *08 April 2021*
Accepted: *09 July 2021*
Published: *12 August 2021*

Citation:
Goldman RE, Broderick JE, Junghaenel DU, Bolton A, May M, Schneider S and Stone AA (2021) Beyond Average: Providers' Assessments of Indices for Measuring Pain Intensity in Patients With Chronic Pain. Front. Pain Res. 2:692567. doi: 10.3389/fpain.2021.692567

Introduction: Effective clinical care for chronic pain requires accurate, comprehensive, meaningful pain assessment. This study investigated healthcare providers' perspectives on seven pain measurement indices for capturing pain intensity.

Methods: Semi-structured telephone interviews were conducted with a purposeful sample from four US regions of 20 healthcare providers who treat patients with chronic pain. The qualitative interview guide included open-ended questions to address perspectives on pain measurement, and included quantitative ratings of the importance of seven indices [*average pain*, *worst pain*, *least pain*, *time in no/low pain*, *time in high pain*, *fluctuating pain*, *unpredictable pain*]. Qualitative interview data were read, coded and analyzed for themes and final interpretation. Standard quantitative methods were used to analyze index importance ratings.

Results: Despite concerns regarding 10-point visual analog and numeric rating scales, almost all providers used them. Providers most commonly asked about *average pain*, although they expressed misgivings about patient reporting and the index's informational value. Some supplemented *average* with *worst* and *least pain*, and most believed pain intensity is best understood within the context of patient functioning. *Worst pain* received the highest mean importance rating (7.60), *average pain* the second lowest rating (5.65), and *unpredictable pain* the lowest rating (5.20).

Discussion: Assessing *average pain* intensity obviates obtaining clinical insight into daily contextual factors relating to pain and functioning. Pain index use, together with timing, functionality and disability, may be most effective for understanding the meaning to patients of high pain, how pain affects their life, how life affects their pain, and how pain changes and responds to treatment.

Keywords: pain intensity, pain measurement, mixed-methods research, qualitative research, provider interviews, chronic pain

INTRODUCTION

Effective clinical care for chronic pain requires accurate, comprehensive, and meaningful pain assessment. It is widely acknowledged that patients' pain experiences are multidimensional, including sensory, affective, and perceptual aspects (1–3). Of the different dimensions of pain assessment, pain intensity is a primary focus in clinical care and pain management to indicate the magnitude of pain, and is meant to describe pain level or intensity (4). Self-reports of pain intensity are typically collected during patient encounters and also represent the primary outcome in most clinical trials of pain disorders (1, 5, 6). Although patient self-report pain ratings using a 0–10 numeric rating scale or a 100-mm visual analog scale are commonly used in clinical practice, improving the degree to which pain assessments provide clinically useful information can facilitate optimal patient care.

Many instruments are available to measure pain intensity. They vary by type of response options, descriptors used to anchor pain ratings (e.g., "pain as bad as one can imagine"), and the reporting period specified (e.g., pain over the past week, past month) (5). A common feature of most pain measures is their focus on *average* level of pain over a period of time. However, a fundamental quality of the pain experience is that pain does not remain at the same level all of the time. Prominent recommendations for core outcome measures in chronic pain clinical trials emphasize the measurement of specific features of pain intensity over time, such as pain maxima, minima, and frequency (1), as secondary outcomes. Additionally, temporal patterns of pain (e.g., episodic, chronic recurrent, constant but fluctuating in intensity) have been described as important to classifying chronic pain (2).

Over the past decades, real-time data collection methods involving Experience Sampling or Ecological Momentary Assessment (EMA) have received increasing attention in pain research. Using EMA, patients rate their momentary pain intensity multiple times per day in their natural environment, which makes it possible to capture temporal features of patients' pain intensity in great detail (7–10). While assessments of specific aspects of pain intensity other than average pain are beginning to be acknowledged in research on chronic pain, to date, it is unclear which temporal indices of patients' pain intensity should be assessed to achieve the greatest utility. Of various pain indices, the *worst* (highest) and *least* (lowest) pain over time have received substantial attention in empirical research, and have been recommended as outcomes in clinical trials (1, 11–14). Additionally, empirical studies suggest that the *amount of time* patients spend in *low* pain or *high* pain represent distinctive features of the pain experience (15–17). Evidence from observational research and clinical trials also highlights the importance of examining pain *fluctuation*, which has been linked to psychosocial outcomes and assay sensitivity (18–23). Finally, studies have shown that the *unpredictability* of shifts in pain [e.g., whether pain occurs after a specific trigger or without warning] is associated with central nervous system performance and functional outcomes (24–26).

These findings suggest that, from an empirical perspective, alternative measures of pain intensity may augment understanding of patients' pain experience and how pain relates to functioning in daily life. However, we know very little about the applied clinical relevance of such assessments, that is, the extent to which they would also augment the information available to clinicians in routine pain practice outside of the research context. This is an important gap in the existing literature because the benefits of utilizing measures that capture alternative aspects of pain intensity levels in patient care depend upon whether they fit the needs and perspectives of those providing medical care to patients.

The present study aims to address this gap. Incorporating stakeholders such as healthcare providers in research to evaluate outcome measures has been strongly promoted by policy makers and regulatory agencies (27–31). In the present mixed-methods paper (which uses data from a larger study that included providers, patients, and regulators), we aimed to investigate providers' perspectives on and ratings of the utility of measures focusing on alternative aspects of pain intensity when evaluating treatment outcomes in chronic pain care. We included a quantitative rating exercise within a qualitative individual interview of healthcare providers. The primary research questions guiding the healthcare provider interviews were: How do providers evaluate the utility of pain intensity assessment in clinical practice? Which aspects of pain intensity are most useful to providers in managing their patients' chronic pain? How do providers value assessments that capture specific aspects of patients' pain levels in addition to (or as alternative to) *average pain* level?

MATERIALS AND METHODS

Eligibility, Recruitment, and Providers

Healthcare providers were recruited for interviews through the American Academy of Pain Medicine (AAPM) mailing list. Providers were selected randomly from the list based on geographic region by targeting zip codes, representing 13 US states and four geographic regions—Northeast, Midwest, South and West. A total of 81 males (56%) and 64 females (44%) across the country were sent participation invitation letters by postal mail. An initial batch of 100 invitation letters was sent; due to low initial response, we used the same method to send a second batch of 45 letters to a new set of providers. All providers were purposively selected based on sex and geographic region. Follow-up phone calls were made to anyone who did not respond to the letter. Eligibility included ability to read and speak English, willingness to provide verbal informed consent, and work role including more than 8 h per week of seeing patients with chronic pain. A threshold of 8 h per week was selected to allow for inclusion of providers who were not exclusively focused on treating patients with chronic pain and treat patients outside of pain specialty settings. The process we used is concordant with purposive sample creation (32) whereby a small sample is selected that includes the diverse characteristics desired in the sample, and recruitment and data collection ceases when data saturation is achieved (33, 34).

Data Collection

Procedures

The study was approved by the University of Southern California Institutional Review Board (UP-15-00228) and informed consent was obtained from all enrolled providers. Participants were sent a reminder email 2 days prior to their scheduled interview with an informed consent information sheet, and a pain index sheet containing seven pain indices and definitions: patients' average pain, worst pain, least pain, the amount of time patients spend in no pain or low pain, the amount of time patients spend in high pain, the extent to which pain fluctuates, and the unpredictability of shifts in pain (**Table 1**). The indices were selected based on the literature of basic temporal and distributional characteristics of pain that are commonly derived from EMA and other diary methods (7, 10, 35). We note that this list is by no means exhaustive, and more complex temporal features of pain such as the dominance in duration of high vs. low pain states (8) or the autocorrelation of pain intensity states (7) that have been examined as EMA-derived pain outcomes are not considered here. Interviews were conducted by the first author (REG). The semi-structured interviews were audio recorded, lasted between 30 and 45 min, and were professionally transcribed. The initial monetary incentive offered to providers was a $150 gift card, which was later increased to $200 to enhance participant recruitment.

Interview

The interview question guide explored how healthcare providers typically collect information about their patients' pain levels; how they view each of the seven pain indices; and which indices might be most useful in their work with chronic pain patients, and why. Core questions were asked of all participants, supplemented by spontaneous probes and follow-up questions. Open-ended questions were followed by structured questions to explore providers' perspectives on and experiences with the seven different pain indices. During this latter part, the interviewer asked participants to talk about each index in terms of the most important/useful pain outcomes of pain treatment, and the most important/useful to them in their work with patients.

Next, the pain measurement concepts sheet was used for rank ordering and rating tasks intended to elucidate the subjective usefulness of each of these indices to providers for characterizing patients' pain (Results for the rank ordering task are presented elsewhere) (10). For the rating task, participants rated each of the indices independently for importance for measuring treatment response, where 0 = no importance and 10 = extremely important. Providers read their ratings aloud for the interviewer to document, and explained in their own words how they made their decisions.

Data Analysis

Standard quantitative methods were used to analyze the ratings of each of the indices. A repeated-measures ANOVA with one within-subjects factor (importance ratings) was performed to test the omnibus null hypothesis that all pain indices were rated as equally important. Pairwise *post-hoc* comparisons with Benjamini-Hochberg correction (36) to control for inflation of Type 1 error due to multiple (i.e., 21) comparisons were subsequently performed to test for differences in the mean importance ratings between individual indices.

Qualitative interview data were analyzed in iterative fashion, beginning as the transcripts became available and continued through and beyond data collection. In this way, the researchers were able to recognize when they reached data saturation such that no new content or concepts were appearing in the interview data, and data collection should stop. This process resulted in our ceasing data collection after 20 interviews were completed and analyzed.

First, the immersion/crystallization technique (37) for data analysis was used, which involved repeated readings of the transcripts with careful note-taking and team discussions about emerging patterns and themes. We constructed a saturation grid to track patterns as they emerged and to determine when no new information was obtained (33, 34). This process was

TABLE 1 | Pain indices and definitions presented to providers during the interviews.

Pain index	Definition/Explanation
Average pain intensity over a week	If we take many ratings of a patient's pain intensity during a week, add them up and then divide by the number of ratings, this would give us an average of a patient's pain during that week.
Level of pain intensity when it is at its worst during a week	If we take many ratings of a patient's pain intensity during a week, we could see what a patient's *highest* pain level was. This would indicate the level of pain intensity when it was at its worst.
Level of pain intensity when it is at its least during a week	If we take many ratings of a patient's pain intensity during a week, we could see what a patient's *lowest* pain level was. This would indicate the level of pain intensity when it was at its least.
Amount of time patient spends with no or low pain during a week	This refers to how much of the time during the week a patient didn't feel any or felt very little pain. That is, if we were to take many ratings of a patient's pain intensity, we could figure out the amount of time during a week that a patient had no pain or almost no pain.
Amount of time patient spends in high pain during a week	If we were to take many ratings of a patient's pain intensity during the week, we could figure out the amount of time when a patient had ratings of pain intensity at very high levels.
How much pain intensity fluctuates or changes during a week	If we take many ratings of a patient's pain intensity during a week, we can get a sense of how much a patient's pain intensity varies from moment-to-moment or day-to-day over the week. That is, whether the intensity is more or less constant or how much a patient's pain fluctuates [that is, goes up and down].
Amount of unpredictability of pain levels during a week	This refers to the degree to which a patient's pain intensity changes for reasons that the patient can't identify. If a patient doesn't know when and why his/her pain changes, then a patient's pain levels are unpredictable.

supplemented with template organizing style analysis (38) where a codebook and coding dictionary were created based on topics and themes identified through individual immersion and discussion among the project team members. This was followed by independent line-by-line coding by two team members using NVivo software (39, 40). Inter-rater reliability was assessed as the coders repeatedly met throughout the coding process to compare and refine their use of codes. Transcripts and code reports were then read again, with discussions among team members to consider alternative interpretations of the data, reconcile conflicting interpretations, and to come to final presentation of results (41–43). COREQ guidelines for reporting qualitative research were consulted during the preparation of this article.

RESULTS

In this paper, we present findings from analysis of qualitative data from provider interviews, as well as quantitative results of a pain index importance rating exercise providers completed during the interview.

Participant Characteristics

The 20 provider participants were drawn from four broad regions of the US: Northeast ($n = 5$), South ($n = 7$), Midwest ($n = 2$), and West ($n = 6$). There were 15 MDs, 2 NPs, 1 PA, 1 PhD Psychologist, and 1 PhD Pharmacologist/Toxicologist; 13 males and 7 females, aged 31–65, with mean age of 43.8. Years in practice ranged from 1 to 30, with mean years of 14. The sample size used in this study is consistent with qualitative research design, and our iterative analysis process ensured that saturation was reached (42, 44).

Providers' Perceptions of the Validity of Standardized Pain Rating Scales

Regardless of which type of index providers in this study favored for use in routine clinical care to measure patients' perceptions of their pain, almost all asked patients to report their level of pain using a 10-point visual analog or numeric rating scale. Nevertheless, providers described multiple problems with this method. Many providers stated that, over time, even if other indicators demonstrated that the patient's pain had improved (e.g., increased function), some patients persisted in reporting their pain intensity at a consistent, high level on the scale. Providers attributed this inertia in pain reporting to a patient's long-established self-identification as a person with a high level of pain.

> "And so, they really don't seem to move a lot on the number itself. And part of that is something I of course don't at all understand. But I think that it really has become just more of, 'I'm an 8.' It's just one of those things."

Other problems providers cited were patients' lack of literacy regarding use of scales, the idiosyncrasy with which the points on the scale are viewed from patient to patient, and patients' reluctance at times to even designate a point on the scale.

> "I think there's a big problem with the scale. A lot of patients just don't understand what it means. Some people, they are in terrible pain, but they will still give you a lower number, and others may not seem to be in such terrible pain, but they always have higher numbers."

> "[Patients] get frustrated when you ask them, 'What is the lowest pain?'. Sometimes people say, 'Well, the lowest pain is I don't have pain sometimes at all. And sometimes I have it but it's really bad.' I think it's a very difficult question to answer."

> "They say the pain is higher than it really is, or they say the pain is a 12 or 15. They walk comfortably into the office, and they'll be sitting there breathing normally. So I find a lot of patients, no matter how hard I try to put it in context, don't really understand."

Providers in this study observed that patients can more easily recall high pain than low pain. Therefore, participants believed that when patients are asked to recall their pain over a period of time up to the present, patients most often focus on the higher pain levels they experienced, regardless of what percentage of the time they endured high pain. In addition, many providers felt that patients tend to "catastrophize" their pain.

> "So if they tell me that they have chronic 15- and 12-out-of-10 pain usually, and they're sitting there comfortably in front of me, then I'll kind of dig into it a little deeper and say, 'Well, what's the worst pain you ever felt?"

Another variable inherent in the use of pain scales for interpreting patients' pain levels is the reason that individual patients are seeing the provider at that time.

> "It varies on what they're here for. Are they here to get pain medications? Then they're going to be a 10 all the time. Are they here to get a procedure? Then they may be a little bit lower on the scale. It varies on what I see that they're looking for. So you can ask them where their pain is, but you haven't figured out how to put a meter on that yet."

Some providers in this study noted that when patients are in pain at the moment of their medical visit, it can be difficult for them to focus on how the pain was different in the preceding period of time: "[Patients are] just trying to make it through the next hour until 'I get my pills." Further, many providers explained that patients with chronic pain over time come to relate to their pain as a significant element of their identity, which consciously or unconsciously, they become reluctant to relinquish for a variety of reasons.

> "All of this pain and how they relate to it has become part of their story that they tell themselves.... If they have been self-identifying as a pain patient in some way for a long period of time, I think simply that having to let that go and move on, in and of itself, is anxiety inducing. And so even if they're doing well they want to hedge their bets a little bit and they want to say, 'Okay, I'm feeling a little bit better but I am not about ready to say that I'm all the way better or I'm getting better cause what if this goes away, what if this is only temporary? I've been burned in the past.'"

Providers' Strategies for Assessing Patients' Pain Intensity

Several themes emerged in response to the question about how providers typically collect information about their patients' pain levels. A few providers asserted that they do not start their pain and treatment efficacy assessment by asking about intensity, and instead ask questions such as, "Did your pain get better?" Most providers, however, stated they typically begin by asking patients what their level of pain is at the current moment. Most then ask what their patients' pain has been *on average*, over a past period of time, sometimes unspecified though usually the past 7 or 30 days. Some ask for pain levels on different specific days, or weekends vs. weekdays. Many said they end their inquiry about pain intensity there, although some next proceed to one or more additional pain indices, most often *worst pain* and/or *least pain* over the specified period. As one provider explained:

> "When I ask people about low pain first before their worst pain, they don't even answer low pain. They would answer the worst pain. It's just because I think they think if they say the lowest pain first I would not ask about their worst pain. And they would not get the treatment they deserve or whatever.... Now what I've started doing is I ask the worst pain first."

One provider who predominantly provides injections and other pain-relieving procedures explained:

> "Unfortunately, we try to boil everybody down into a little pot and it never works with pain because it's so multidimensional. But when we measure pain as one point, the FDA decided it wasn't pain intensity that was important. It was pain relief. So that a patient could say, 'I feel relief' rather than 'My pain is this.'"

Many participants stated that pain ratings alone are not sufficient for understanding the patients' experiences of pain. Some ask patients for descriptive words about the pain, and most ask about function and ability/disability in addition to pain ratings. Some providers explained that juxtaposing what a patient was doing at the time of having *worst pain* in the past 7 days with what the patient was doing at the time of having *least pain* is critical for assessing whether the treatment is working. Many provided examples such as if a patient's *least pain* occurs in conjunction with lying on a comfortable couch and *worst pain* occurs when doing a physical task, pain ratings are placed within the context of daily life and can inform treatment decisions. While overall, *fluctuating pain* was rated by providers as among the least useful indices, some said they used this index specifically to ask about context and activity, and then to educate patients about managing their high and low pain levels throughout the day.

Some providers said that when patients with chronic pain succumb to fluctuating pain by avoiding normal daily activities that increase the pain, they do themselves a disservice, and besides treatment "[it takes] a little bit of education. Because we know that it's going to fluctuate. Sometimes catastrophizing, and just fear that the pain's going to get worse if they do anything, and anytime it fluctuates a little bit, [they believe] it's getting worse." Providers emphasized that however pain intensity is identified, effective treatment is predicated on their own good communication and listening skills.

> "Something that I view more about the population [patients with chronic pain] overall is that they don't feel heard and they don't feel believed by people. And whether that's their peers, or they're walking around hurting and they don't have a broken arm, or they're not in a wheelchair. Their life is very difficult, but they *look fine*. And that's a very frustrating experience for them. So to me, [patients' pain reporting] is really about communication, like a way to say, 'Things are really bad'.... It's one of the only ways they think they have to express how bad things are for them because they feel very misunderstood and not heard."

Some providers said they try to enhance the usefulness of the pain measurement scales by regularly engaging in educating patients about what the scales mean.

> "When people think 10, I don't take them at face value, sometimes. I'll say, 'So you barely got out of bed this morning 'cause you're at a 10?' 'Oh, well, okay, maybe it's an 8.' So I really educate them on the numbers and the specificity of it. 'When you say it's the worst pain of your life, explain to me.' So we have a lot of education in my practice and I really reinforce the patients to be involved and make it a team effort."

Regardless of the pain indices providers preferred, most used these in an effort to gauge how the patient's pain has changed over time and in response to treatment. As one provider stated: "So it's something that we can look and see where they were at. It's not very reproducible *between* people, but for the *same person* it might be indicative of how they're doing today vs. how they've done in the past." Ultimately, if the patient is receiving treatment for the pain, the goal is to "Look what pain does to disrupt life, and what is that treatment doing for that."

Providers' Perspectives on the Pain Indices

Understanding how healthcare providers viewed the importance of the seven pain indices was an essential component of data collection. The individual provider ratings for each pain index are displayed in box-and-whisker plots in **Figure 1**. **Table 2** provides summary statistics. Descriptively, *worst pain* received the highest mean importance rating. *Unpredictable pain* received the lowest rating. *Average pain* received the second lowest importance rating. It is notable that all seven pain indices received mean importance ratings above the midpoint of 5 on the 0–10 scale, suggesting that all indices were deemed somewhat important by providers. In addition, providers varied substantially in the importance ratings of each index, with standard deviations approaching or exceeding 2 scale points, suggesting there was limited consensus among providers about which indices are most and least important.

In statistical analyses, a one-way repeated-measures ANOVA yielded a significant omnibus F-test, [$F(6, 14) = 3.11$, $p = 0.007$], indicating significant differences in the mean importance ratings. *Post-hoc* pairwise comparisons showed that *worst pain* was rated significantly more important than *time in no/low pain* ($d = 0.64$, $p = 0.040$) and *unpredictable pain* ($d = 0.94$, $p = 0.01$). In

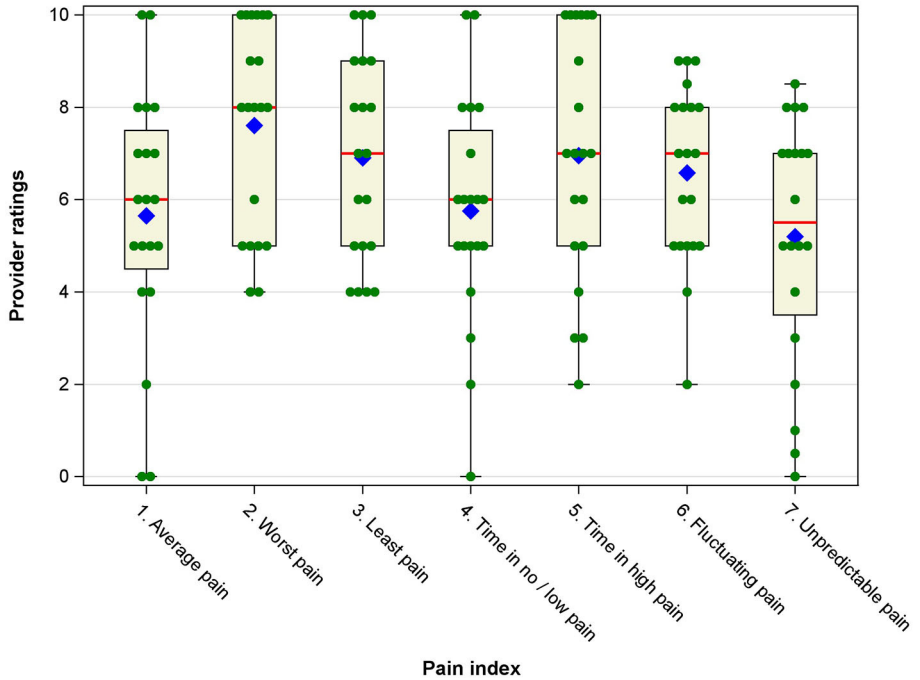


FIGURE 1 | Box-and-whisker plots of provider ratings of importance/usefulness of the pain intensity indices. Blue diamonds represent the mean, red vertical lines represent the median, boxes represent the 25th and 75th percentile, whiskers represent the range of ratings, and green filled circles represent individual provider ratings for each pain index.

TABLE 2 | Mean [SD] of provider ratings of importance/usefulness of individual pain intensity indices.

Pain index	Mean [*SD*]
Worst pain	7.60 [2.23]
Time in high pain	6.95 [2.67]
Least pain	6.90 [2.20]
Fluctuating pain	6.58 [1.93]
Time in no pain/low pain	5.75 [2.45]
Average pain	5.65 [2.76]
Unpredictable pain	5.20 [2.66]

Pain indices are displayed in order of mean importance ratings from most important to least important.

addition, *least pain* and *fluctuating pain* were rated significantly more important than *unpredictable pain* ($d = 0.64$, $p = 0.040$, and $d = 0.59$, $p = 0.048$, respectively).

In the qualitative interview component, providers claimed to use *average pain* exclusively or at least most often because they knew it to be the most commonly-used index in clinical care and clinical trials. Some admitted they had never considered other ways of measuring pain intensity until the six additional indices were outlined during the study interview. Despite their consistent use of *average pain*, providers described numerous problems with it which were reflected in the ratings. They stated that patients misconstrue the meaning of average to include the level of pain experienced most frequently (i.e., mode), rather than the arithmetic mean. Others asserted that since high pain is more memorable than low pain, the reported average will be pushed artificially higher. One provider explained how patients become irritated by the request to report average since patients see their pain as unique, not "average." Other providers said patients insisted that their pain was far worse than "average." These misconstrued ways of responding about *average pain* would corrupt the meaning and interpretation of this index if providers assumed patients were referencing the *average pain* level over the prescribed period.

Some providers explained that recall of specific pain levels during the designated time period was a problem even if patients knew how to calculate average. One provider tried to mitigate this problem by having patients keep pain logs: "And then we would go ahead and take their score, average it by the number of readings, and then we say, 'See, your average pain is 5.' [And the patient would respond], 'Oh no, it's got to be an 8'. So we stopped doing that." Some providers who acknowledged the inadequacy of *average pain* ratings still felt that seeing how patients' reported averages went up or down over time is useful, no matter how the patients conceptualize the concept of average, and so the index is still in common use. As a provider claimed, "I've found most value in the average because I think it's taking out jagged edges. I find the average is something that is going to give you a better curve with less disturbance in it." Some claimed that *average pain* may still be the best indicator of pain intensity over time, but not in the way it is currently used: "I think it would be

education on the patients' part, my part, making sure we're all on the same page."

Many providers said they valued the *least pain,* and to a lesser extent *time in low pain,* indices for the information these provide about how medication or behavioral treatment is working, and for the success and relief good values imply for the patient. A few providers stated this preference in terms of "putting a positive spin" on the patient's experience of chronic pain. In contrast, some noted that if a patient already has considerable time in low pain there is not much for the provider to do to help the patient, so the index is less useful. Many providers in the study stated that these two indices are difficult to use with patients. They attributed this evaluation to factors they have observed in their practice: patients' inherent bias toward remembering more clearly their worst pain and time in high pain; patients' tendencies to "catastrophize" their pain; the centrality in patients' minds of lowered functional capacity due to those times in high pain; and patients' reluctance to admit or talk about any lessening of their pain for fear that providers would become distracted from or not take seriously their reports of accompanying periods of high pain. Providers said they believed that the patients' primary goal is to keep their providers' "attention on the pain." In addition, at the time of the interviews there was increasing media attention across the US about the burgeoning prescription opioid crisis, and providers speculated that the resulting environment of heightened pressure to decrease opioid prescribing may impact patients' urgency to justify continuation of medication: "In some cases they feel that if they don't continue to rate their pain high, maybe you're going to say, 'You don't need all this medication."'

> "They are pretty wise about the number they *need to give*, for it to be noteworthy enough to a provider. So, that individual may be more likely to report a 7, an 8, a 9.... I have many people that will say, 'My 8's like anybody else's 20.' Everyone thinks that theirs is the worst."

> "I realize that there's that fear that if they say they're doing better, 'Oh good, then it's time to reduce their pain meds.'... So they absolutely do come in with the worst pain, 9, and they're not looking like they're about to die... I suspect that's what they're concerned about, that we're going to take their pain medication away and they're going to be miserable and not able to work or function or have a good quality of life."

Some providers offered that this reaction could backfire as patients who persist in reporting inflated pain levels may lead providers to reduce or completely de-prescribe seemingly ineffective medication.

> "But when you sit down and say, 'Look, this is not working for you. You've been seeing me for months and every time you come in here your pain's a 9 or a 10. That tells me that what we're doing is not working and now we need to reassess whether this is actually helping you. And being that it's so high all the time, I can't keep you on this medication."'

Given providers' views that patients are averse to reporting *least pain* and *time in no or low pain*, some explained that they avoided these indices in clinical practice even though they themselves felt they were good indicators of pain intensity. Providers who supplemented *average pain* with other indices said they do use *time in no or low pain* to ascertain how a patient's ability to function ("what they can do") has changed since the previous medical visit. Some said they felt that a patient's reporting of low pain is extremely significant since it is so much less memorable than high pain. However, given the difficulties of having patients focus on low pain, more providers used *worst pain and time in high pain* to understand what a patient can and cannot do, and they gave these two indices high importance ratings.

> "If we are able to reduce the amount of time in high pain, I think that would be a useful measure, and even if we're not able to reduce the average pain but are able to reduce the amount of time in high pain I think that would give patients a better quality of life. And I think that would be a useful thing to follow, and especially if we were to show medical necessity for our treatment, that we're reducing the time in high pain and it's improving quality of life."

Providers asserted that *unpredictable pain* is especially debilitating for patients, impedes patients in planning activities, and has a high emotional toll. However, providers explained they rated the index as least useful since they are unable to adequately treat these unexplainable onsets of pain.

Inextricability of Pain Intensity and Function

The importance of different pain indices for understanding patient functioning arose spontaneously throughout the interviews, with most providers stating that, ultimately, it is functionality that patients value and seek. Providers emphasized the importance of understanding the direct effect of pain intensity on the patient's functioning, and interpreting the meaning of each measurement in relation to what the person was doing at the time the measurement refers to. "It's more important to identify when the pain is at its worst and *what's going on at that point*, and when the pain is at its least *and what's going on at that point*. Sure, you can average those numbers, but I'm not quite sure if patients would think about it like that." Providers emphasized that highlighting function is particularly critical for patients with chronic pain, because these patients will likely live with some level of pain into the future.

> "When I talk to my patients about outcomes, I tell them that we are trying to improve function. We may not make it go away completely because most of the pains are chronic, they don't go away, but we are trying to improve the quality of life and improve the function. That's the goal."

> "I have a little graphic that I show people. You're trying to make life feel bigger so pain feels smaller by comparison. Your pain may not change at all, and that's just the truth."

Providers, therefore, emphasized that the most reasonable treatment goal is to increase functioning, which necessitates educating patients so as to minimize their tendency to give in to the pain and decrease their activity.

> "If the person has a memory of their pain coming down, that suggests that they're learning from what I'm trying to teach them. What I'm also trying to frame for them is that pain goes up and down. Increases and decreases in pain really have not a lot of meaning with respect to anything in a chronic pain patient being wrong. So, [I tell patients] 'You should continue with your activity program. Continue with your therapy. Yes, your pain is gonna be from time to time worse, but that doesn't mean you're causing harm."'

DISCUSSION

Accurate assessment of pain intensity is a basic necessity for gauging change in pain levels, providing adequate treatment, and communicating with patients about their pain (2, 45). Researchers have become increasingly interested in understanding pain intensity as a dynamic phenomenon (7, 8, 20). In fact, the ability to quantify, predict, and possibly influence dynamic aspects inherent in the ebb and flow of pain in patients' daily lives has been described as a paradigm shift in pain research (46). However, less is known about the extent to which assessments capturing specific temporal aspects of pain would augment the information available to clinicians in routine pain practice. This interview study with providers who care for patients with chronic pain found that, not surprisingly, *average pain* continues to be medical providers' most commonly-used index. Despite common usage, providers did not provide quantitatively high ratings of the importance or usefulness of the *average pain* index. Their reasons included patient confusion about the meaning of average and patients' inability to accurately recall pain levels.

Many participants complement their use of the *average pain* index with questions assessing *worst pain* and/or *least pain*, or *time in high pain* or *low/no pain*. There was little consensus among providers about which index is most useful or important, although overall, *worst pain* was rated highest among the seven indices, and *unpredictable pain* was rated lowest. As others have found (1, 2), *worst* and *least pain* are considered useful to better understand temporal fluctuations or to calculate an average. Interestingly, while *least pain* and *time in low/no pain* were believed to be important, providers found it challenging to focus patients' attention on these and so they may not be feasible indices to use in routine pain assessment.

Our finding that *worst pain* was rated as most important is interesting in view of the US Food and Drug Administration's (FDA) recommendation to use *worst pain* ratings as the primary outcome in drug clinical trials (14). It is possible that providers were aware of the FDA recommendation when making their importance ratings. Regardless, the perceived importance of *worst pain* was supported by this study, especially when coupled with information about patients' activities to better understand potential contributors to pain exacerbations.

The *fluctuating pain* index was rated as only moderately important. However, increasing empirical evidence supports the idea that identifying pain level variations may be an important clinical target. For example, momentary pain fluctuations have been found to relate to affective distress and activity limitations (47), and individuals with greater pain variability have shown higher depression levels and lower self-efficacy for pain management (20). Pain variability may hold promise for informing clinicians about potential barriers to successful adjustment and management (18, 21, 23, 48). Our participants recognized that unpredictable pain can be extremely distressing for patients. However, they were reluctant to ask about it because of their overall goal to control the pain, which is difficult for unpredictable shifts in pain when the reasons are not known.

Providers in our study also emphasized that the importance of different aspects of pain intensity must be understood in the context of its impact on patient functioning. This is in line with recommendations for "core outcome measures" for chronic pain by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) (1). A survey of patient stakeholders showed that patients considered a variety of functioning domains [e.g., emotional well-being, enjoyment of life, fatigue] as highly important for evaluating the consequences of their chronic pain (49). Different aspects of pain intensity may have interactive or cumulative effects on specific facets of patient functioning. For example, in a recent study, we found that pain variability, worst pain levels, and the time chronic pain patients spent at high levels of pain uniquely related to patient physical and social functioning above the effects of average pain (9). The present study supports the importance of recognizing the pain-functioning linkage from a clinical pain management perspective.

Finally, patients' ability and willingness to properly use the pain rating scale was a consistent provider concern in our study. Prior qualitative (50) and quantitative (51) research with chronic pain patients showed that the ostensibly simple task of completing standardized pain ratings is often approached idiosyncratically. The task to provide recall pain ratings over extended periods of time further adds to the complexity of obtaining accurate pain summary ratings (52). Pain rating trainings (53), as well as clearer instructions and more precise descriptions of scale anchors and recall periods (52), might improve pain rating accuracy. Whether these could be implemented in routine clinical care should be explored.

This study has several limitations. Our sample consisted predominantly of MDs, and the results may be different across different professional backgrounds or areas of specialty. Even though the invitation letters were sent through the AAPM mailing list, invitations to participate in the study were unsolicited, and only 14% of providers responded, which may have biased the results due to self-selection effects. Nevertheless, our sample was geographically diverse and robust in that we were able to stop recruiting interviewees after having interviewed 20 providers because our iterative data analysis process allowed us to identify that we had reached data saturation. Furthermore, clinicians generally had been treating patients for a considerable amount of time (average = 14 years). Providers who are newer to the field may not hold the same views. Additionally, even though our sample size was consistent with prior qualitative work, it should be considered small for the quantitative analyses. Larger samples could examine the

hierarchy of preferences for different pain measures and enable subgroup analyses to compare preferences based on clinicians' professional background, years in practice or area of specialty. Along similar lines, our sample consisted predominantly of health care providers with prescriptive authority; an interesting direction for future research would be to compare the preferences between providers with and without prescriptive authority. Finally, in future research, it would be valuable to compare the perspectives of healthcare providers with those of patients with chronic pain. Understanding how patients' perspectives might relate to providers' views could be particularly valuable when it comes to assessments of pain intensity because prior research has shown that patient and provider ratings of patient pain intensity do not necessarily correspond with one another (54–56). We note that we had originally attempted to compare the views of providers and patients as part of this study. Unfortunately, the patient interviews did not provide sufficiently detailed and nuanced information to pursue meaningful qualitative analysis in this group. It is well-possible that interview scripts that are specifically tailored to patients and their personal experiences with pain in daily life (rather than probing patients for their opinions about specific pain measures, as was attempted here to maximize comparability between interview scripts for patients and providers) would have yielded richer qualitative patient data.

CONCLUSIONS

The main goal of the present study was to examine whether specific aspects of patients' pain intensity other than average pain would be viewed as useful by providers. Most providers in our study agreed that inquiring about multiple aspects of pain intensity could augment patient evaluation in clinically relevant ways. They described how additional indices beyond or instead of *average pain* (particularly *worst pain* and *least pain*) would constitute a more effective strategy for pain measurement. Providers also mentioned the benefit of including contextual information about timing, function, and disability for enhancing understanding of patients' responses to treatment and for understanding the meaning to patients of high pain, how pain affects their life, how life affects their pain, and how pain changes and responds to treatment. Provider preferences are just one important aspect in a comprehensive effort to identify the relevance of alternative pain intensity measures. Future studies should therefore test the usefulness of soliciting different types of pain intensity information directly in clinic settings to evaluate the practical gains for routine care. Additionally, more research is needed to evaluate whether different aspects of pain intensity are differentially impacted by treatment, and whether assessment of multiple aspects of pain intensity could contribute to treatments that are more closely tailored to the needs of individual patients.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Southern California Institutional Review Board (UP-15-00228). The ethics committee waived the requirement of written informed consent for participation.

AUTHOR CONTRIBUTIONS

RG conducted interviews, data analysis, and took the lead on drafting the manuscript. JB contributed to study conceptualization and to drafting the manuscript. DJ contributed to study planning and to drafting the manuscript. AB conducted interviews, data coding and qualitative data analysis, and contributed to drafting the manuscript. MM conducted interviews, data coding and qualitative data analysis, and contributed to drafting the manuscript. SS contributed to study conceptualization, conducted data analysis, and contributed to drafting the manuscript. AS conceived of the presented idea, supervised the conduct of the study, and contributed to drafting the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Institutes of Health, National Institute of Arthritis and Musculoskeletal and Skin Diseases [NIAMS] R01 AR066200.

REFERENCES

1. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. (2005) 113:9–19. doi: 10.1016/j.pain.2004.09.012
2. Fillingim RB, Loeser JD, Baron R, Edwards RR. Assessment of chronic pain: domains, methods, and mechanisms. *J Pain*. (2016) 17:T10–20. doi: 10.1016/j.jpain.2015.08.010
3. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, editors. *Handbook of Pain Assessment*. 2nd ed. New York, NY: Guilford Press (2001). p. 15–34.
4. Jensen MP, Karoly P. Measurement of cancer pain via patient self-report. In: Chapman CR, Foley KM, editors. *Current and Emerging Issues in Cancer Pain: Research and Practice*. New York, NY: Raven Press (1993). p. 193–218.
5. Litcher-Kelly L, Martino SA, Broderick JE, Stone AA. A systematic review of measures used to assess Chronic Musculoskeletal Pain in clinical and randomized controlled clinical trials. *J Pain*. (2007) 8:906–13. doi: 10.1016/j.jpain.2007.06.009
6. Smith SM, Hunsinger M, McKeown A, Parkhurst M, Allen R, Kopko S, et al. Quality of pain intensity assessment reporting: ACTTION systematic review and recommendations. *J Pain*. (2015) 16:299–305. doi: 10.1016/j.jpain.2015.01.004

7. Mun CJ, Suk HW, Davis MC, Karoly P, Finan P, Tennen H, et al. Investigating intraindividual pain variability: methods, applications, issues, and directions. *Pain*. (2019) 160:2415–29. doi: 10.1097/j.pain.0000000000001626
8. Schneider S, Junghaenel DU, Ono M, Stone AA. Temporal dynamics of pain: an application of regime-switching models to ecological momentary assessments in patients with rheumatic diseases. *Pain*. (2018) 159:1346–58. doi: 10.1097/j.pain.0000000000001215
9. Schneider S, Junghaenel DU, Broderick JE, Ono M, May M, Stone AA. II. Indices of pain intensity derived from ecological momentary assessments and their relationships with physical, emotional, and social functioning: an individual patient data meta-analysis. *J Pain*. (2021) 22:371–85. doi: 10.1016/j.jpain.2020.10.002
10. Stone AA, Broderick JE, Goldman R, Junghaenel DU, Bolton A, May M, et al. I. Indices of pain intensity derived from ecological momentary assessments: rationale and stakeholder preferences. *J Pain*. (2021) 22:359–70. doi: 10.1016/j.jpain.2020.08.003
11. Lang EV, Tan G, Amihai I, Jensen MP. Analyzing acute procedural pain in clinical trials. *Pain*. (2014) 155:1365–73. doi: 10.1016/j.pain.2014.04.013
12. Atkinson TM, Mendoza TR, Sit L, Passik S, Scher HI, Cleeland C, et al. The brief pain inventory and its "pain at its worst in the last 24 hours" item: clinical trial endpoint considerations. *Pain Med*. (2010) 11:337–46. doi: 10.1111/j.1526-4637.2009.00774.x
13. Farrar JT, Pritchett YL, Robinson M, Prakash A, Chappell A. The clinical importance of changes in the 0 to 10 numeric rating scale for worst, least, and average pain intensity: analyses of data from clinical trials of duloxetine in pain disorders. *J Pain*. (2010) 11:109–18. doi: 10.1016/j.jpain.2009.06.007
14. U.S. Department of Health and Human Services Food and Drug Administration. *Guidance for Industry Analgesic Indications: Developing Drug and Biological Products*. (2014). Available online at: http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm384691.pdf (accessed: April 26, 2019).
15. Krueger AB, Stone AA. Assessment of pain: a community-based diary survey in the USA. *Lancet*. (2008) 371:1519–25. doi: 10.1016/S0140-6736(08)60656-X
16. Smith WR, Bauserman RL, Ballas SK, McCarthy WF, Steinberg MH, Swerdlow PS, et al. Climatic and geographic temporal patterns of pain in the Multicenter Study of Hydroxyurea. *Pain*. (2009) 146:91–8. doi: 10.1016/j.pain.2009.07.008
17. Schneider S, Stone AA. Distinguishing between frequency and intensity of health-related symptoms from diary assessments. *J Psychosom Res*. (2014) 77:205–12. doi: 10.1016/j.jpsychores.2014.07.006
18. Allen KD. The value of measuring variability in osteoarthritis pain. *J Rheumatol*. (2007) 34:2132–3.
19. Peters ML, Sorbi MJ, Kruise DA, Kerssens JJ, Verhaak PF, Bensing JM. Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain. *Pain*. (2000) 84:181–92. doi: 10.1016/S0304-3959(99)00206-7
20. Schneider S, Junghaenel DU, Keefe FJ, Schwartz JE, Stone AA, Broderick JE. Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: Associations with psychological variables. *Pain*. (2012) 153:813–22. doi: 10.1016/j.pain.2012.01.001
21. Jamison RN, Brown GK. Validation of hourly pain intensity profiles with chronic pain patients. *Pain*. (1991) 45:123–8. doi: 10.1016/0304-3959(91)90176-X
22. Farrar JT, Troxel AB, Haynes K, Gilron I, Kerns RD, Katz NP, et al. Effect of variability in the 7-day baseline pain diary on the assay sensitivity of neuropathic pain randomized clinical trials: an ACTTION study. *Pain*. (2014) 155:1622–31. doi: 10.1016/j.pain.2014.05.009
23. Harris RE, Williams DA, McLean SA, Sen A, Hufford M, Gendreau RM, et al. Characterization and consequences of pain variability in individuals with fibromyalgia. *Arthritis Rheum*. (2005) 52:3670–4. doi: 10.1002/art.21407
24. Allen K, Bosworth H, Coffman C, Jeffreys A, Oddone E, Yancy W. Associations of frequent predictable and unpredictable pain with functional and psychological outcomes. *Osteoarthritis Cartilage*. (2013) 21:S263–4. doi: 10.1016/j.joca.2013.02.550
25. Moseley GL, Brhyn L, Ilowiecki M, Solstad K, Hodges PW. The threat of predictable and unpredictable pain: differential effects on central nervous system processing? *Aust J Physiother*. (2003) 49:263–7. doi: 10.1016/S0004-9514(14)60142-2
26. Meulders A, Vansteenwegen D, Vlaeyen JW. Women, but not men, report increasingly more pain during repeated (un) predictable painful electrocutaneous stimulation: evidence for mediation by fear of pain. *Pain*. (2012) 153:1030–41. doi: 10.1016/j.pain.2012.02.005
27. Frank L, Basch E, Selby JV. The PCORI perspective on patient-centered outcomes research. *JAMA*. (2014) 312:1513–4. doi: 10.1001/jama.2014.11100
28. Kirwan JR, de Wit M. Patients as partners: building on the experience of Outcome Measures in Rheumatology (OMERACT). *Arthritis Rheumatol*. (2016) 68:1334–6. doi: 10.1002/art.39678
29. de Wit M, Abma T, Koelewijn-van Loon M, Collins S, Kirwan J. Involving patient research partners has a significant impact on outcomes research: a responsive evaluation of the international OMERACT conferences. *BMJ Open*. (2013) 3:e002241. doi: 10.1136/bmjopen-2012-002241
30. Williamson P, Altman D, Blazeby J, Clarke MJ, Gargon E. The COMET (Core Outcome Measures in Effectiveness Trials) initiative. *Trials*. (2011) 12:A70. doi: 10.1186/1745-6215-12-S1-A70
31. Hoang A, Goodman SM, Navarro-Millán IY, Mandl LA, Figgie MP, Bostrom MP, et al. Patients and surgeons provide endorsement of core domains for total joint replacement clinical trials. *Arthritis Res Ther*. (2017) 19:1–6. doi: 10.1186/s13075-017-1476-9
32. Kuzel A. Sampling in qualitative inquiry. In: Crabtree B, Miller W, editors. *Doing Qualitative Research, 2nd Edn*. Thousand Oaks: Sage Publications (1999). p. 33–45.
33. Fusch PI, Ness LR. Are we there yet? Data saturation in qualitative research. *Qual Rep*. (2015) 20:1408–16. doi: 10.46743/2160-3715/2015.2281
34. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant*. (2018) 52:1893–907. doi: 10.1007/s11135-017-0574-8
35. Stone AA, Broderick JE, Schneider S, Schwartz JE. Expanding options for developing outcome measures from momentary assessment data. *Psychosom Med*. (2012) 74:387–97. doi: 10.1097/PSY.0b013e3182571faa
36. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med*. (1990) 9:811–8. doi: 10.1002/sim.4780090710
37. Borkan J. Immersion/Crystallization. In: Crabtree B, Miller W, editors. *Doing Qualitative Research, 2nd Edn*. Thousand Oaks, CA: Sage Publications (1999). p. 179–94.
38. Crabtree BF, Miller WL. Using codes and code manuals: a template organizing style of interpretation. In: Crabtree BF, Miller WL, editors. *Doing Qualitative Research in Primary Care: Multiple Strategies, 2nd Edn*. Newbury Park, CA: Sage Publications (1999). p. 163–77.
39. NVivo qualitative data analysis software QSR International Pty Ltd. Version 12 (2018).
40. Bernard HR, Wutich A, Ryan GW. *Analyzing Qualitative Data: Systematic Approaches*. Thousand Oaks, CA: SAGE Publications (2017).
41. Stake R. *Qualitative Research: Studying How Things Work*. New York, NY: The Guildford Press (2010).
42. Patton MQ. *Qualitative Research and Evaluation Methods*. 4th ed. Thousand Oaks, CA: SAGE Publications (2015).
43. Rubin HJ, Rubin IS. *Qualitative Interviewing: The Art of Hearing Data*. 2nd ed. Thousand Oaks, CA: SAGE Publications (2005). doi: 10.4135/9781452226651
44. Toye F, Seers K, Allcock N, Briggs M, Carr E, Andrews J, et al. Patients' experiences of chronic non-malignant musculoskeletal pain: a qualitative systematic review. *Br J Gen Pract*. (2013) 63:e829–41. doi: 10.3399/bjgp13X675412
45. Cedraschi C, Delézay S, Marty M, Berenbaum F, Bouhassira D, Henrotin Y, et al. "Let's talk about OA pain": a qualitative analysis of the perceptions of people suffering from OA. Towards the development of a specific pain OA-related questionnaire, the Osteoarthritis Symptom Inventory Scale (OASIS). *PLoS ONE*. (2013) 8:e79988. doi: 10.1371/journal.pone.0079988
46. Winger JG, Vilardaga JCP, Keefe FJ. Indices of pain variability: a paradigm shift. *Pain*. (2019) 160:2411–2. doi: 10.1097/j.pain.0000000000001627
47. Litcher-Kelly L, Stone AA, Broderick JE, Schwartz JE. Associations among pain intensity, sensory characteristics, affective qualities, and

activity limitations in patients with chronic pain: a momentary, within-person perspective. *J Pain*. (2004) 5:433–9. doi: 10.1016/j.jpain.2004.07.005
48. Hutchings A, Calloway M, Choy E, Hooper M, Hunter DJ, Jordan JM, et al. The Longitudinal Examination of Arthritis Pain (LEAP) study: relationships between weekly fluctuations in patient-rated joint pain and other health outcomes. *J Rheumatol*. (2007) 34:2291–300.
49. Turk DC, Dworkin RH, Revicki D, Harding G, Burke LB, Cella D, et al. Identifying important outcome domains for chronic pain clinical trials: an IMMPACT survey of people with pain. *Pain*. (2008) 137:276–85. doi: 10.1016/j.pain.2007.09.002
50. Robinson-Papp J, George MC, Dorfman D, Simpson DM. Barriers to chronic pain measurement: a qualitative study of patient perspectives. *Pain Med*. (2015) 16:1256–64. doi: 10.1111/pme.12717
51. Williams ACD, Davies HTO, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. *Pain*. (2000) 85:457–63. doi: 10.1016/S0304-3959(99)00299-7
52. Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. *Pain*. (2008) 139:146–57. doi: 10.1016/j.pain.2008.03.024
53. Smith SM, Amtmann D, Askew RL, Gewandter JS, Hunsinger M, Jensen MP, et al. Pain intensity rating training: results from an exploratory study of the ACTTION PROTECCT system. *Pain*. (2016) 157:1056–64. doi: 10.1097/j.pain.0000000000000502
54. Prkachin KM, Solomon PE, Ross J. Underestimation of pain by health-care providers: towards a model of the process of inferring pain in others. *Can J Nurs Res Arch*. (2007) 39:88–106.
55. Seers T, Derry S, Seers K, Moore RA. Professionals underestimate patients' pain: a comprehensive review. *Pain*. (2018) 159:811–8. doi: 10.1097/j.pain.0000000000001165
56. Junghaenel DU, Broderick JE. Validation of the MPI patient profiles by partners and healthcare providers. *Pain*. (2009) 144:130–8. doi: 10.1016/j.pain.2009.03.025

Conflict of Interest: AS is a Senior Scientist with the Gallup Organization and a consultant with IQVIA and Adelphi Values, Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pain Research

ORIGINAL RESEARCH
published: 18 August 2021
doi: 10.3389/fpain.2021.723520

Exploration of Hospital Inpatients' Use of the Verbal Rating Scale of Pain

Luke Bosdet[1], *Katie Herron*[2] *and Amanda C. de C. Williams*[1]*

[1] *Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom,*
[2] *Pain Medicine Department, Liverpool University Hospitals NHS Foundation Trust, Liverpool, United Kingdom*

Background: Assessment of pain largely relies on self-report. Hospitals routinely use pain scales, such as the Verbal Rating Scale (VRS), to record patients' pain, but such scales are unidimensional, concatenating pain intensity and other dimensions of pain with significant loss of clinical information. This study explored how inpatients understand and use the VRS in a hospital setting.

Methods: Forty five participants were interviewed, with data analysed by thematic analysis, and completed a task concerned with the VRS and communication of other dimensions of pain.

Results: Participants anchored their pain experience in the physical properties of pain, its tolerability, and its impact on functioning. Their relationship to analgesic medication, personal coping styles, and experiences of staff all influenced how they used the VRS to communicate their pain.

Conclusion: Participants grounded and explained their pain in semantically similar but idiosyncratic ways. The VRS was used to combine pain intensity with multiple other elements of pain and often as a way to request analgesic medication. Pain scores need to be explored and elaborated by patient and staff, content of which will imply access to non-pharmacological resources to manage pain.

Keywords: pain measurement, pain assessment, pain communication, scale interpretation, analgesics

OPEN ACCESS

Edited by:
Loren Martin,
University of Toronto, Canada

Reviewed by:
Gabrielle Pagé,
Université de Montréal, Canada
Manon Ranger,
University of British Columbia, Canada

***Correspondence:**
Amanda C. de C. Williams
amanda.williams@ucl.ac.uk

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *10 June 2021*
Accepted: *26 July 2021*
Published: *18 August 2021*

Citation:
Bosdet L, Herron K and Williams ACdC (2021) Exploration of Hospital Inpatients' Use of the Verbal Rating Scale of Pain. Front. Pain Res. 2:723520. doi: 10.3389/fpain.2021.723520

INTRODUCTION

Both the original (1) and updated (2) definitions of pain make clear that the relationship between identifiable physical damage or pathology and the magnitude of pain is variable; pain cannot be directly observed or reliably estimated by clinicians. As a result, the preferred method of assessing pain in verbally competent patients is to use patient self-report. Thus, pain is "whatever the experiencing person says it is, existing whenever the experiencing person says it does" (3).

There are multiple methods used to assess pain; the most common are the numerical rating scale (4), the verbal rating scale (4), the visual analogue scale (5), and the McGill Pain Questionnaire (6, 7). None of the pain rating scales give instructions to indicate what pain phenomena are to be included, nor how they should be translated into the scale metric (8). They have different performance characteristics (9), with the former two more reliable than the latter (10). As they are accessible to

verbal enquiry and response, they are therefore better suited to hospital settings (11). Despite superior reliability and validity of the numerical rating scale (12, 13), even in older patients (14), use of the verbal rating scale (VRS) is common. The VRS requires patients to rate their pain using ordinally-arranged adjectives describing pain intensity (e.g., no pain, mild pain etc.). Scores are assigned to the adjectives (e.g., no pain = 0, mild pain = 1 etc.), and are often treated as an interval or ratio scale (5) to enable quantitative description of pain and calculation of change with treatment.

However, treating verbal measures of pain in this way is controversial. Ordinal verbal categories provide no information about the distance between points on the scale (5) that would allow interval-level scoring, and the assumption that those distances are consistent across people in pain is untested. People vary considerably in how they convert their pain into verbal categories (15), and pain ratings are confounded by psychological and decisional processes that do not fit the linear structure necessary for equidistance (16). Further, single ratings do not separate the constructs of pain intensity, distress, and interference, when these are likely to be variably associated and idiosyncratically represented in a single term (17, 18). Of particular importance is the lack of separation of the sensory and affective components of pain (19, 20). Last, as noted by Fordyce (21), pain ratings are behaviours, so it is important to consider the context in which they are provided and the implications of the rating for both patient and receiver.

It can be advantageous to address pain intensity and pain distress separately in relation to clinical intervention (22), to avoid giving analgesic drugs for high pain ratings that in fact represent emotional distress (19). It is therefore helpful to understand how patients use the VRS to communicate their pain-related needs, and to use this to inform clinicians' responses. This study followed Uher's methodological guidelines (8) to explore how hospital inpatients translated their experience of pain into the VRS categories and how they communicated their pain needs to medical and nursing staff during routine pain assessments.

MATERIALS AND METHODS

Procedure

Participants were recruited from adult inpatient wards in a central London hospital. NHS ethical approval was obtained (ID: 16/YH/0417) and as part of this process, an external "expert by experience" was consulted whose advice was used to make changes in the information sheets and protocols.

The researcher (LB) obtained an honorary contract with the hospital's specialist pain team who liaised with ward managers across the hospital for permission for ward staff to be approached about the study. Five wards agreed. The researcher then explained the study to the nurse-in-charge for that shift and obtained permission to collect data; the nurse-in-charge was asked to identify, and ask staff to approach, suitable patients based on the inclusion criteria: (a) over 16 years of age, (b) able to communicate effectively in English, and (c) with capacity to consent and take part. Eligible patients were then approached by the researcher. Data were collected across a period of 4 months in 2016–2017, with the process of asking permission and identifying patients repeated each day of data collection and on each ward.

Forty-five participants took part in the study. We intended to recruit patients equally across three groups: acute pain, chronic pain (longer than 3 months), and chronic with acute pain. However, participants' descriptions of their pain did not fit well into these groups so data on pain chronicity are provided. The study consisted of two parts: a semi-structured interview and a personal pain scale task. Both parts were conducted at the participant's bedside with their informed consent; we did not want to limit recruitment to patients who could walk to a private room, and few private spaces were available. Interviews were audio-recorded. Verbatim instructions are provided in **Supplementary Material**.

This hospital used a five-point VRS as the routine pain assessment for adults, with the categories of *No Pain*, *Mild*, *Moderate*, *Severe*, and *Very Severe* pain. The VRS was required by the hospital to be completed at the same time as other routine observations, usually every 4 h. Participant characteristics were recorded as they appeared in their medical notes: age, gender, ethnicity, and primary diagnosis (i.e., the reason for admission to hospital). For the sake of simplicity, comorbid diagnoses were not recorded. Participants were asked verbally about the length of time they had experienced pain.

Interview Protocol

A semi-structured interview was developed to understand how inpatients used the VRS and the process by which they made their pain ratings. The interview started by asking participants to rate their current pain. The following questions broadly covered: (a) how participants understood the VRS categories, (b) how they selected a category, (c) how pain affected their emotions, (d) how they coped with their pain, (e) what they thought of the VRS, and (f) what else they would want to communicate to the hospital staff about their pain. Interviews consisted of nine core questions and the interviewer had the option to ask follow-up questions to elaborate or clarify on the above aims. Example questions from the interview protocol include: (a) For you, what are the main differences between mild and moderate pain? (b) What else would you like to tell the nurse or doctor about your pain? The full interview schedule and introduction can be found in the **Supplementary Material**.

The interview data were analysed using Thematic Analysis (23). The analysis was grounded in a critical realist epistemology in which the experience of pain was recognised as real and located in the body, but recognising that each individual constructed the experience in personal ways both in relation to him or herself and in communicating with others. This epistemological standpoint was chosen as it validated the participants' experiences as authentic, but recognised that communicating the experience is influenced by both individual differences and social processes. The iterative steps recommended by Braun et al. (23) were followed.

Transcription and Immersion of the Data

Each of the interviews was transcribed using Express Scribe Transcription Software. A total of 27 interviews were transcribed

by the first author and the remaining 18 by a volunteer which were checked by the researcher against the audiotape for accuracy. The interviews were transcribed in accordance with recommendations in Barker et al. (24): verbatim speech content, but without information about the tone, loudness, speed etc. of speech. Aside from transcribing, all interview transcripts were re-read before beginning coding so the researcher would be familiar with the data.

Generating Initial Codes

The transcripts were uploaded into Nvivo, qualitative analysis software. The first author worked systematically through each of the transcripts, coding each unit of meaning found, and keeping as close to the original meaning as possible without implying any higher categorisation. All data were coded, without making assumptions of relevance to the research question to protect against the loss of potential themes or sub-themes at later stages.

Searching for Themes

The first author systematically worked through the codes of meaning to merge codes based on meta-level meanings from the explicit content of what the participant reported, rather than implicit or implied meaning. Previous theory also partly informed the type of codes that were chosen, in particular, that the pain experience can be divided into sensory, affective, and cognitive elements (4). For example, text coded as "stabbing," "throbbing," and "nagging" were coded under "Quality of Pain." We also began to focus on the research aims and discarded some codes that were irrelevant to the study. For example, a participant who identified as an alcoholic was anxious that they would not be able to stop drinking.

Reviewing and Redefining Themes

We examined the developed themes against Patton's (25) criteria of internal homogeneity and external heterogeneity, in other words, whether the codes were sufficiently similar to constitute a wider theme, and whether the theme was different enough from other themes to be considered separately. For example, "Quality of Pain" was later absorbed into a broader theme of "Physical Properties of Pain." This stage also involved credibility checks, described in the section below. Through this process the themes and subthemes evolved over several iterations before settling on the themes described in the **Results** section.

Quality Evaluation

In keeping with guidelines for qualitative research by Elliot et al. (26), we included: (a) a "reflexive statement" reporting the researcher's theoretical and personal orientation; (b) a wide range of participants, described in terms of their pain and length of hospital stay, to improve the likelihood of developing a broad understanding of the phenomenon; (c) multiple participant quotations to illustrate each theme; and (d) credibility checks by analytical auditing and testimonial validity. Analytical auditing required another researcher to code five randomly selected transcripts, blind to the first coder's decisions, for comparison on development of initial themes. Testimonial validity, in the form of "synthesised member checking" (27), involved asking the original participants for feedback on the accuracy of the analysis.

Reflexive Statement

A reflexive position was taken in order to make more transparent the researcher's biases in analysis and interpretation. The first author and lead researcher is a male in his early thirties who was training in clinical psychology. He is from a working class family that generally considered post-modern epistemologies as irrelevant, in reaction to which he developed an interest in constructionism, but with a strong preference for pragmatism. His training in cognitive behavioural therapy and systemic approaches both emphasised splitting experience into different elements and sequences, while also recognising the often bidirectional nature of cause and effect. He was drawn to the topic of pain assessment mainly through dissatisfaction with what he perceived as oversimplification, as well as a desire to produce research with real world application.

Personal Pain Scale Task

The purpose of this task was better understanding of how the VRS and elaborations of it described their experience of pain. Each participant was asked to elaborate their own personal pain scale using a horizontal line centred on a landscape A4 page as a template.

The general instructions to participants were to develop a scale that represented their pain. Participants were initially asked to record the VRS categories (*No Pain*, *Mild*, *Moderate*, *Severe* and *Very Severe*) on the line, spaced as made best sense to them, and then to add any terms they wished, located on the line. All terms were measured from *No Pain* (i.e., the left end of the line) and recorded in centimetres. Where participants did not indicate the exact position of a category on the line (e.g., they just wrote *Mild* above a section of the line), the position was calculated by the midpoint of the written word. During the task, participants were asked to "think out loud" and audio recorded in order to understand the method of development. The "thinking out loud" data were originally planned to be analysed in accordance with the method described under "Interview Protocol." However, these data did not add any new substantial information in addition to the interview data, so were not included in this study.

We first examined whether participants placed the VRS categories in sufficiently similar positions to be considered a shared category, then examined the distances between categories, equidistance in particular, and finally we examined participants' additions and modifications to their scales.

RESULTS

Participants

Forty-five participants (**Table 1**) completed the semi-structured interview and, of these, 29 agreed to complete the Personal Scale. Participants had a total of 25 different diagnoses, with the most common being Coxarthrosis ($n = 9$), Crohn's disease ($n = 6$), and fractures ($n = 5$).

TABLE 1 | Participant sample characteristics.

Characteristic	Frequency
Gender	Male = 10; female = 35
Age	*M*: 50 (SD = 18); range: 19–81
Pain chronicity	*Mdn:* 6 years; range: 1 day−40 years
	N for pain < 1 year = 10
Ethnicity	White British: 28 (62%)
	White other: 5 (11%)
	Black or Black British: 4 (9%)
	Asian or British Asian: 1 (2%)
	Other: 1 (2%)
	Not Stated or Missing: 6 (13%)
Diagnostic category	Arthritis related disorders and problems: 17 (38%)
	Gastrointestinal problems: 17 (38%)
	Tumour related disorders: 3 (7%)
	Injuries and other disorders: 6 (13%)
	Missing: 2 (4%)
Recruitment wards	Orthopaedics: 21 (47%)
	Gastroenterology: 14 (31%)
	Oncology: 7 (16%)
	Short stay surgery: 3 (7%)

Semi-Structured Interview

Analysis of the qualitative data from the semi-structured interviewed produced eight themes with three subthemes. These were grouped in three clusters: (a) how the pain experience was anchored, (b) relationship with analgesic drugs, and (c) relationship with staff. **Figure 1** displays a map of the themes and relationships between them. The themes are explored below, highlighting similarities and differences between participants.

Cluster 1: How Pain Experience Was Anchored

This cluster of themes pertains to how participants operationalised their pain in order to anchor the VRS categories, and included the physical properties of pain, how pain impacted on their function, their ability to endure pain, and how they coped with pain.

Theme: Physical Properties of Pain

Unsurprisingly, many participants ($n = 25$) made reference to the physical sensations of pain when demarcating categories of the pain measure. This included the amount of pain, number of pains, the longevity, constancy, and qualities of pain. Generally, as the number of these properties increased, reported pain severity worsened. However, the precedence and concatenation of these properties varied across participants. For example, pain longevity and constancy were sometimes given more prominence than the amount of pain.

P14: I go back to the comparison with the broken leg and gastritis … Obviously, that hurt more than that … But this ultimately hurts more than that did because it's there all the time …

Similarly, some participants commented on how the number of pains had an additive effect on pain ratings.

P42: I don't just think of one pain I think of all my pain… and then amalgamate it according to how much, how much pain I'm in… if only one thing is hurting, then it will be a lower score than if my joints are very sore and I've got my pancreas kicking off, my bowels cramping…

Subtheme: Comparison to Other Pains

Many participants compared current pain with other experiences of pain for their pain categories ($n = 17$), and with hypothetical pains. The time frame of these comparisons also varied, from the previous day in hospital to distant occasions. There were references to "everyday" or "normal" pains, as well as more exceptional pain from the past.

P28: I'd compare [the current pain] to my kidney stones, I compare all my pain now to the worst pain I've ever experienced …

P12: … that would be stabbing pain I think. I mean I assume what you'd feel if you'd been shot …

Comparisons to other pains also had emotional meaning.

P45: today I'm feeling pretty good … but I feel a lot better because I was previously in quite severe pain.

Theme: Interference With Activities

The majority of participants ($n = 34$), and the most prevalent theme in this cluster, described pain severity by referencing how much pain interfered with important activities. This includes mental activities, such as concentration and maintaining attention, as well as conversing with others, sleep, movement, and coping strategies. Participants described both what they could and could not do to delineate the severity of their pain.

P44: I know [the pain is] there but I can also forget about it and focus on something else … I know I'm hurting but I know, I can do something else, you know, read, listen to something, the pain is not getting in the way of something else that I'm doing, that would be mild for me.

P2: 'Mild', I can have a conversation with someone and completely focus on that conversation. 'Moderate', my mind will start focusing slightly on the pain and I will lose the conversation slightly, or miss parts of what that person is saying, my concentration won't be as good. 'Severe', I wouldn't be able to have a conversation.

Participants reported using a wide range of coping strategies, the most common being focusing away from pain ($n = 10$), interacting with other people ($n = 6$), and physical activities such as going for walks ($n = 6$). With greater pain, participants reported being unable to use these strategies due to insufficient physical and mental resources, and hoped by reporting higher levels of pain to be given analgesics to help cope with it.

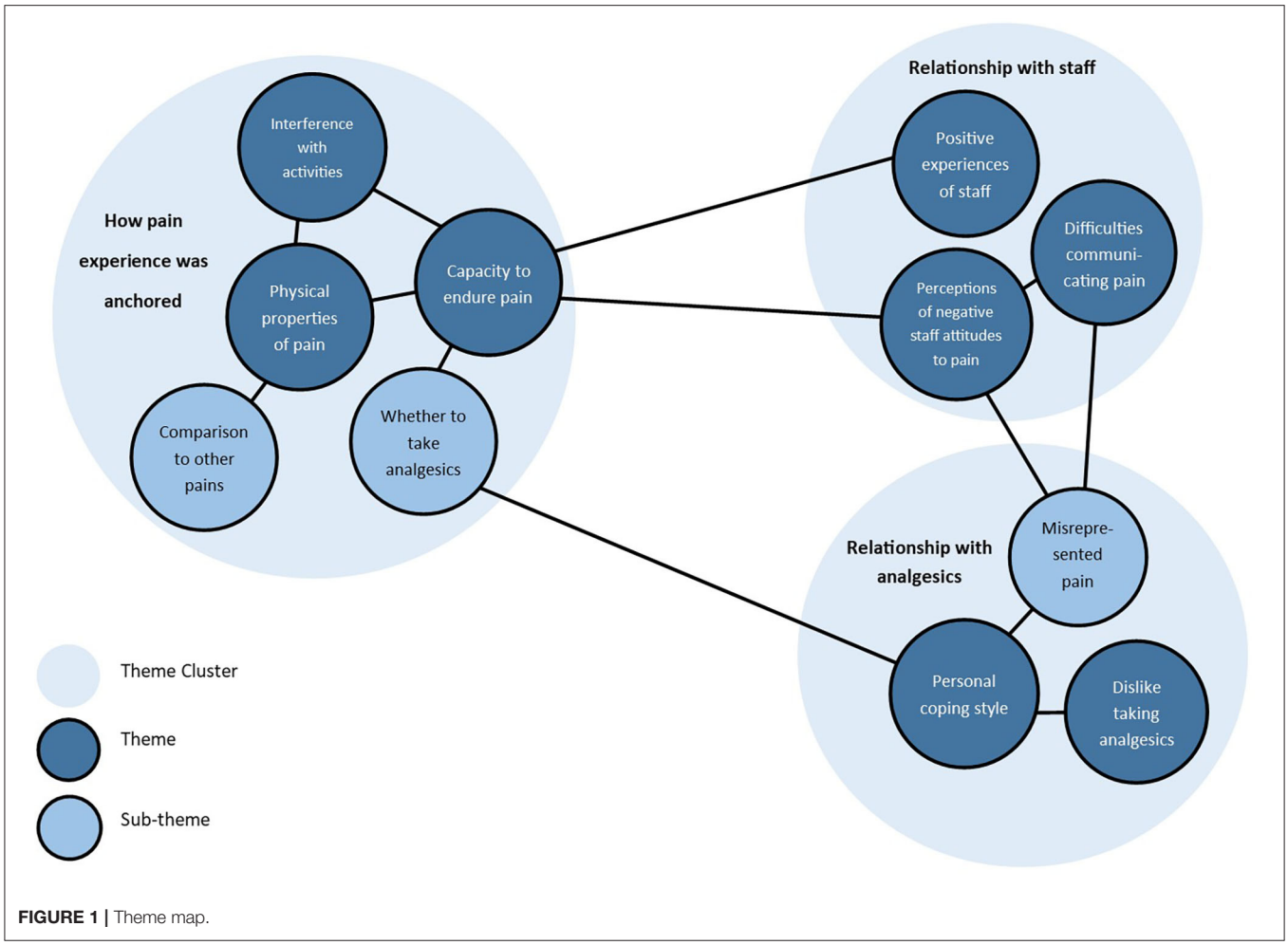


FIGURE 1 | Theme map.

P42: Normally I'm very good at distraction, mindfulness, that sort of thing … and if I can't use them, all I want is my medication.

Theme: Capacity to Endure Pain

In addition to the physical qualities of pain and how it interfered with activities, participants also spoke about the tolerability of pain ($n = 15$). As pain became less bearable, severity of pain ratings increased.

P1: Mild is something you can actually deal with …

P28: [Moderate pain is] probably stuck in bed but [I] can tolerate it …

The *Very Severe* category was often described more elaborately compared to the other categories of the VRS. The words used often represented the limits of capacity, such as unbearable ($n = 3$), agony ($n = 4$), and excruciating ($n = 1$). Some participants reserved the *Very Severe* category for only the worst occasions and used it rarely ($n = 7$).

P29: Oh, very severe is all-consuming, you can't think of anything, and when it gets like that yes I will, I do start crying and screaming … it is hell.

P10: whilst I'm in [very severe pain] I actually wish to die which is like, shocked me because normally I never do …

Many participants commented on the emotional impact of being in pain. This included feeling low ($n = 14$), angry ($n = 7$), and anxious ($n = 4$) as a result of pain. The hospital environment also contributed to these emotions, with some participants stating that their usual coping mechanisms were constrained by the ward environment.

P29: So yeah, [pain] controls everything with my emotions … When I'm having a bad time it turns me into a nasty, snappy, aggressive, horrible person and that's not who I am.

Some participants described how emotions in turn affected how tolerable the pain was. Generally, negative moods exacerbated pain and reduced capacity to tolerate pain.

P38: if you're getting a bit anxious and down with the pain then it's getting up to that severe level and you're having to ask for pain medication …

Subtheme: Whether to Take Analgesics

One subtheme of this theme addressed whether participants would use analgesics ($n = 27$). In this sense, the VRS was used as a communication to nurses that the patient required analgesics. Some participants described a threshold at which they would begin to consider analgesics, mostly *Moderate* ($n = 7$) or *Severe* ($n = 4$). This consideration was related to the "Personal Coping Theme" in the "Relationship to Analgesics" cluster, in that the participant's approach to managing pain affected when they would use analgesics.

P31: moderate pain is something that you kind of live with. Severe pain I guess you'd ring the call bell and say can I have [analgesic] please.

This subtheme was also expressed as the effects of analgesics, in that pain became more tolerable.

P38: I've always got a pain but [analgesics] will bring it down to a manageable level.

Cluster 2: Relationship to Analgesics

The second cluster concerns participants' mixed relationship with analgesics: welcoming help to cope with pain when other coping methods were not enough, but disliking the sense of dependence on analgesics or concerns about possible long-term effects of use. The need for analgesics strongly influenced how the VRS was used.

Theme: Dislike Taking Analgesics

Although all participants who reported pain said that they took some form of analgesic, many described aspects of analgesics that they disliked ($n = 10$), for reasons including side-effects, the build-up of tolerance, and fears of long-term damage.

P2: I wonder, if everyone actually understood the severity of use, overusing painkillers and what it does to their body, if they would necessarily do that all the time.

For those with chronic pain ($n = 36$), especially those with Crohn's Disease ($n = 6$) there was often conflict between adequate analgesia and sedation that impaired everyday life.

P8: it may dissociate me from the pain but it doesn't help the pain itself … and I don't rate dissociation as help because I still want to be able to do what I want to do.

P2: it's got rid of my pain but I haven't gained anything from that, I've still lost my day.

Another reason for disliking analgesics was the fear that analgesia prevented patients from checking their pain levels ($n = 3$), a concern they addressed by periodically stopping or refusing analgesics.

P23: I need to know how bad the pain is, so if I'm junked up with painkillers I don't know.

P43: I am the type of person who from time-to-time will stop taking painkillers in order just to check [my pain]

Theme: Personal Coping Style

Participants varied in their strategies for managing their pain and reporting pain levels to staff. Some participants had an uncomplicated approach to reporting their pain, preferring to give accurate responses when asked, and describing a straightforward relationship between their pain and the use of analgesics.

P19: When it's there it's there, I always say it … I won't try to hide it.

P13: I'm in pain and I don't want to have a conversation about it, they're here and they know what to do and that's it.

Subtheme: Misrepresented Pain

The VRS was widely understood as a way to communicate need for analgesia, but some participants described deliberately over- or under-reporting pain in order to influence the offer of analgesics. Many ($n = 20$) described under-reporting pain in the belief that they had a higher pain tolerance ($n = 12$), or in order to avoid making a negative impression on staff by appearing "soft," a "nuisance" or a "wimp." Two participants described how these attitudes developed from their families of origin.

P11: I think potentially it could be cultural or generational as to why I don't think it's the done thing to say that I'm in pain … I grew up single parent family, mother who was extremely hard working and never complained a day … so it would for me feel wrong, I feel as though I'm moaning if I'm complaining …

Some participants described a preference for handling pain using their own emotional coping methods, so under-reported pain in order to avoid discussions about analgesics.

P8: I know that painkillers at that point aren't going to help, and my own techniques are going to be far superior so it's a lot easier to say I'm in no pain and get on with what I do.

Deliberately over-reporting pain was much less frequently described ($n = 4$); participants described this as goal-orientated, most commonly to take control of when and what analgesics they received.

P30: because by the time they actually go get the pain relief, they were only going to give me moderate pain relief like, it would have already turned into severe

P42: I can feel when my pain is progressing, and I like to pre-empt it before it gets to, before it gets too high. Because when it gets too high, it's then very very difficult to get back down again … So I might give a slightly higher pain score.

Cluster 3: Relationship With Staff

The themes in this cluster concern using the measure as a communication tool in an ongoing relationship with hospital staff. Participants discussed the difficulties of communicating their pain, as well as the positive effects of attentive staff.

Theme: Perceptions of Negative Staff Attitudes to Pain

Many participants described negative experiences with staff about their pain ($n = 20$), often suggesting disapproval of the use of analgesics or the report of pain. For example, participants related that some staff did not act on requests for analgesics, failed to pass on key information to other staff, or in one case directly refused to give prescribed analgesics. Several participants also described fears of being negatively evaluated by staff when asking for analgesics.

P26: sometimes in the morning the doctors go 'I gather you had a really good night' and you're like well, no, I told them I was in severe pain and that, so I don't think things get passed to the doctors unless they're really serious things.

Int: And do you think pain is taken seriously?

P26: Not really, no …

Several participants described the problems caused by staff members' assumptions about what indicated pain ($n = 5$); this was a particularly prominent concern for participants with chronic pain problems, who noted that they do not always display their pain.

P42: [The staff] criteria for severe is in tears, can't really communicate, asking for medication, and being kind of, having a face of, pulling a face … Making noises, that sort of thing, and if you're completely absent of that and you give an answer of severe then, I've had plenty of times where someone has said, but you look, you don't look like you're in severe pain, or they've kind of raised an eyebrow to sort of say, oh, oh yeah, course …

P8: You can't have pain if you're smiling, that would be a very good [laughs] assumption, if you're doing a crossword and listening to music you can't be in pain, when in fact that's exactly what I do when I am in pain.

Many participants reported that staff used incorrect presentation of the VRS, using numbers instead of categories, or recording their own estimated pain levels without asking the participant.

P42: quite often people will write down a score, but they haven't asked you. They haven't asked you what your pain is … I was finding that I was getting marks of, that said no pain, or moderate pain, or low pain … which isn't, isn't right

P8: my pain [has been] assessed in at least five different ways … I've been nought to four, one way, and nought to four the other way. Er, one to ten, ten to one, and the mild, moderate, severe but, again, on the ward I've never been asked until you said it if my pain was very severe. That's the first time I realised that was on the scale is when you said it …

Theme: Difficulties Communicating Pain

Many participants remarked on the difficulties of communicating pain to staff, with or without the VRS ($n = 27$). On the VRS some participants struggled to distinguish between adjacent categories ($n = 5$). Participants also described the difficulty of converting the pain experience into scale categories.

P14: I would just tell [staff asking on the VRS] I was completely unable to give an answer because I find the entire thing ridiculous … I don't think you can quantify pain when pain can mean so many different things …

Two participants reflected on how difficult it was for staff to understand pain using medical knowledge and training; others commented more on the inadequacy of the scale in portraying pain. There was, nevertheless, some recognition of the subjective nature of pain and the difficulty for staff in understanding how people used the pain scale.

P30: you think you know what pain is, like from what they teach in University, but it's nothing like that when you experience it yourself.

P28: so my pain to someone else's pain is going to be completely different, the way we rate it, so how is a nurse going to then be able to perceive that in terms on prescribing pain medication?

Theme: Positive Experiences of Staff

The final theme of this cluster consists of how patients used the VRS in relation to positive experiences of relationships with staff ($n = 10$). Participants described how consistent and responsive care for their pain enabled them to report their pain needs more easily. For a few participants, this helped them overcome their usual stoic style which served as a barrier to requesting analgesics.

P11: virtually everybody who I've come into contact with will ask me are you in pain? And they don't just ask are you in pain, they're asking using the scale, so you're getting used to the idea that it's not going to be a shock to say to somebody you're in pain

P15: people ask you, they ask very regular, that come and check on you, and they, they've very positive to you, you know, calling on the bell et cetera so you feel well cared … I wouldn't feel negative about saying well I am in pain.

Another positive experience of staff was their demonstration that they observed non-verbal signs indicating pain.

P29: But they know me well enough here that they can gauge my pain levels against what I'm doing …

Two participants described how the attentiveness of staff made them feel more reassured and relaxed, which helped them deal with their pain.

P17: I think that they know exactly what's going on with me, and, you know, where I should be and … I feel very sort of calm and relaxed about it …

Summary of Themes

How the VRS was used varied by participant across three main areas. Participants reified the categories in semantically similar but idiosyncratic ways. This included grounding the category demarcations using physical sensations, impact on functioning, and levels of tolerance. However, these demarcations also interacted with emotional state and current needs, such as sleep. The main use of the VRS reflected its use as communication, mainly expressing a need for analgesics. Individual participants' relationships to pain and analgesics played a key role in this communication, and positive and negative experiences of staff responses influenced this communication, enabling participants to communicate their pain needs or discouraging them from doing so.

Personal Scale Task

Of the 45 participants interviewed, 29 (64%) agreed to complete the personal scale task, and 16 participants declined or were unable (e.g., due to poor eyesight, fatigue after the interview section). The 21 participants who recorded all five original VRS categories were included in the following analyses. The positions of all terms were normalised, such that 0 and 100 represented the two ends of the horizontal line. For example, *Severe* placed 18 cm from the left on a 26.8 cm line would be recorded as 67.2. This section first determined where categories were positioned by participants on the scale, whether the categories were positioned similarly by participants, and then tested the assumption that categories were equidistant.

Figure 2 displays box plots for all four categories, *Mild* ($M = 11.7$, SD $= 6.2$), *Moderate* ($M = 33.4$, SD $= 11.3$), *Severe* ($M = 63.9$, SD $= 14.6$), and *Very Severe* ($M = 84.6$, SD $= 15.6$). All categories except *Very Severe* met assumptions for normal distribution. *Very Severe* was found to be significantly negatively skewed (z score $= -3.34$) and leptokurtic (z score $= 3.16$). Two scores in the *Very Severe* category were outliers (see **Figure 2**) with z scores > -2. Since the nature of this study was exploratory and did not assume normal distributions, these scores were retained and non-parametric tests used: a Kruskal–Wallis test and follow-up planned comparisons with Mann–Whitney tests. The four category positions were significantly different, $H_{(3)} = 69.79$, $p < 0.001$. *Mild* was significantly different from *Moderate*, $U = 20$, $z = -5.04$, $p < 0.001$; *Moderate* was significantly different from *Severe*, $U = 22$, $z = -4.99$, $p < 0.001$; and *Severe* was significantly different from *Very Severe*, $U = 62$, $z = -3.98$, $p < 0.001$.

To test the assumption of equidistance, the distance between each placed category on the scale was calculated for each participant who had recorded all five categories ($n = 21$). This created four distances: (1) *No Pain* to *Mild*, (2) *Mild* to *Moderate*, (3) *Moderate* to *Severe*, and (4) *Severe* to *Very Severe*. Distances were again normalised to 0 to 100 for comparison; for example, a distance *Mild* to *Moderate* of 6.6 cm on a 26.8 cm scale was recorded as 24.6.

No Pain to *Mild* was the smallest distance ($M = 11.7$, SD $= 6.2$), while *Moderate* to *Severe* was the largest ($M = 30.5$, SD $= 10.1$). *Mild* to *Moderate* ($M = 21.7$, SD $= 9.6$), and *Severe* to *Very Severe* ($M = 20.7$, SD $= 8.6$) were of similar size.

All four distances met assumptions for normality, so a one-way ANOVA was used to test the assumption of equidistance between adjacent categories. The overall result indicated significant differences: $F_{(3, 80)} = 16.08$, $p < 0.001$, and all but two *post-hoc* comparisons were statistically significant (see **Table 2**) using a Bonferroni adjusted alpha level of 0.008 (0.05/6). Overall, the assumption that there are equal distances between pain categories was not supported. In particular, there is a large difference between *Moderate* and *Severe*.

Additions to and Modifications of the Scale

Of the 29 personal scales elaborated by participants, four had no changes or additions to the VRS. Four participants chose to expand the VRS categories but did not add any new ones. Sixteen participants added their own categories to the VRS, and two created a completely new set of categories. Three participants made major structural changes to the scale. Overall, every scale was unique in representing the participant's relationship with pain. Some representative examples of each type of change are displayed below (**Figure 3**).

Figure 3A shows P22's scale (*Very Severe* has been shortened to "*Very*"). This participant had a very short recent experience of pain and chose not to make any additions or changes to the scale. In contrast, P42 (**Figure 3B**) reported a longer experience of pain and had used of the scale over many years. This participant's personal scale was superimposed over the *Severe* and *Very Severe* categories in the form of a numerical scale, converted back to VRS terms when answering medical staff. P11 (**Figure 3C**) also chose to expand on the existing categories, but by adding interventions that might be required and personal experience or evaluation of that pain. P20 (**Figure 3D**) altered the scale completely by adding a y axis of "Intensity/Heat" to represent the partial independence of these aspects in their nerve pain; they also used the two-dimensional space to map different pain locations, as pain often varied across their body. Last, P14 (**Figure 3E**) replaced the VRS categories that did not describe their experience of pain with their own descriptions of feelings and experiences of pain.

DISCUSSION

This study explored how hospital inpatients understood and used a VRS pain scale, presented routinely for monitoring. Overall, participants described a rich variety of meanings in their communication of pain, and reporting pain was heavily influenced not only by social and emotional factors but also specifically by participants' perceived need for analgesics and likelihood of the staff providing them. A large proportion of the interviews were spent discussing analgesic medication, despite there only being one question in the interview protocol.

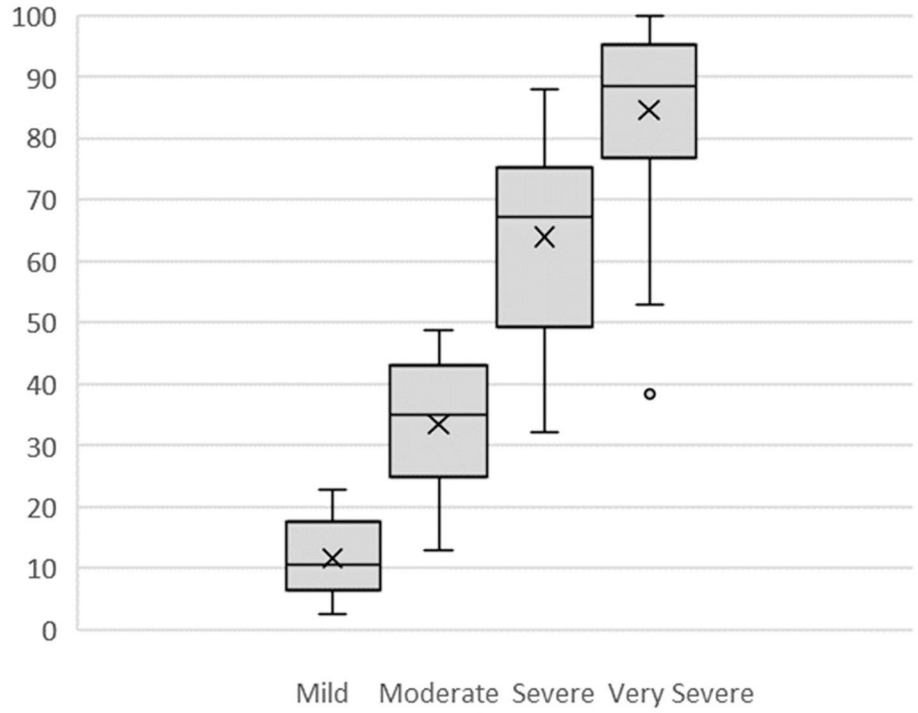


FIGURE 2 | Box Plots of VRS mild, moderate, severe, and very severe numerical values assigned by participants. X is mean; —— is median.

TABLE 2 | Category distance comparisons.

Comparison	Statistics (*t* test, *p*-value, effect size)
No pain to mild and mild to Moderate ($M = 11.73$) ($M = 21.71$)	$t(20) = -3.92$, $p = 0.001$, $d = 1.23$
No pain to mild and moderate to severe ($M = 11.73$) ($M = 30.48$)	$t(20) = -7.08$, $p < 0.001$, $d = 2.24$
No pain to mild and severe to very severe ($M = 11.73$) ($M = 20.65$)	$t(20) = -3.65$, $p = 0.002$, $d = 1.20$
Mild to moderate and moderate to severe ($M = 21.71$) ($M = 30.48$)	$t(20) = -2.81$, $p = 0.011$
Mild to moderate and severe to very severe ($M = 21.71$) ($M = 20.65$)	$t(20) = .37$, $p = 0.714$
Moderate to severe and severe to very severe ($M = 30.48$) ($M = 20.65$)	$t(20) = 3.18$, $p = 0.005$, $d = 1.05$

The results from this study support assertions that patients combined pain affect with other pain elements in their ratings on unidimensional pain scales (19, 22), and made comparisons with previous pain experiences and reported pain in idiosyncratic ways (17) in a complex decision process (16). Two themes in particular are similar to those described by Robinson-Papp et al. (18) with outpatients: the multiple influences on pain rating, and the individuality of referents for the anchor points. The distances between categories, derived from representing the verbal descriptors in spatial terms, corroborate previous findings that categories are not equidistant (4), as they would be were the VRS an interval scale.

Consistent with other research about low adherence to pain management protocols (28, 29), this study also found that participants reported multiple instances of improper use of the pain scale by staff, such as completing it without consulting the patient. This may reflect poor training, weak adherence or inadequate implementation of assessment policies (30), or other organisational or practical issues that influence use of the scale by staff (31, 32). Since this study did not sample staff experience, explanations can only be speculative. Nonetheless, the findings reported here extend the known difficulties with pain management protocols by describing some of the impact these behaviours have on patients. These included a reluctance to report pain due to a fear of being adversely judged as a person, with overall detriment to clinician understanding of the patient. It was encouraging to obtain accounts of positive experiences, of feeling "cared for," enabling participants to report their pain and, to some extent, to manage it themselves. Similarly, staff should be aware of the different ways that pain can be expressed, especially in chronic pain patients, and not believe that it can be determined simply by global impression of behaviour, mood or facial expression.

This research elaborated on the way that pain ratings from unidimensional pain scales such as the VRS, but also including numerical rating scales and visual analogue scales, combine

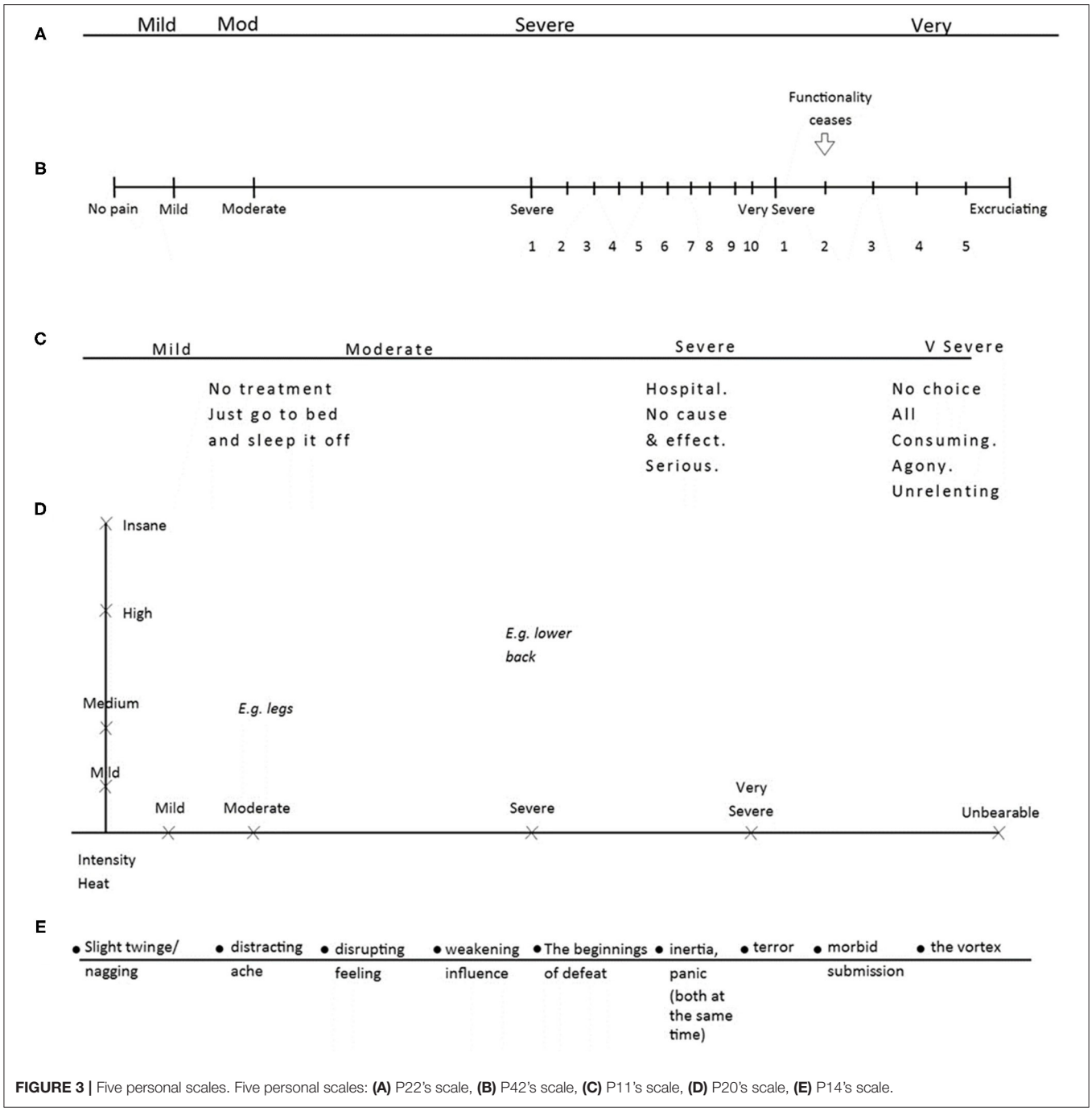


FIGURE 3 | Five personal scales. Five personal scales: **(A)** P22's scale, **(B)** P42's scale, **(C)** P11's scale, **(D)** P20's scale, **(E)** P14's scale.

multiple elements of the pain experience, including pain affect, disability, coping and magnitude, in an ordinal but non-linear and idiosyncratic fashion. To turn to analgesics for all of these, expressed in high pain ratings, is clearly ineffective. While there are more detailed pain measures, such as the McGill Pain Questionnaire (7) that attempt to segregate the various components of pain, they are not practical for routine hospital care. High ratings on a verbal or numerical scale should instead invite further questions to determine what intervention or support would be most helpful. Repeated and consistent use of the unidimensional scale with a follow-up exploration of support options would allow staff and patients to develop expertise in managing pain.

The finding of uneven distances between categories of the verbal rating scale means that interval-level scoring is inappropriate, and some categories, particularly "*Moderate*," may represent a wide span of intensities, overlapping with adjacent categories. In clinical settings this may mean that some changes are more meaningful than others. For example, a pain rating increase of moderate to severe could represent a greater increase

than mild to moderate. Although numerical and spatial scales avoid this problem, change on any unidimensional scale may represent increase or decrease in pain severity or improvement or deterioration in other functions such as mood, mobility, or sleep.

Participants incorporated their capacity to endure pain in the categories they chose, but that capacity was fluid, varying with context and emotional states. Addressing the emotional needs of patients is likely to be a more useful intervention than analgesics when emotional contexts make pain difficult to manage. In particular, feeling low and anxious were the most frequently reported emotional consequences of pain, and these may respond to support for coping, clarifying expectations of pain, providing information about pain, validating pain and providing reassurance. Likewise, consistent and responsive care by staff helped patients cope with the anxiety-provoking nature of pain and the hospital environment.

Similarly, some of the VRS use was goal-orientated. For example, people reported higher pain levels at night, when pain might interfere with sleep, in order to request analgesics. Staff should be aware that if pain is interfering with a valued activity, pain levels are likely to be rated higher. It may be useful to explore this with the patient, aiming for problem-solving. Equally, participants often described keeping occupied as a way to cope with pain, and providing the means to do so, such as liberal visiting hours, can help them to use this strategy.

A strength of this study was the examination of the VRS in an ecologically valid setting. It showed that when staff requested a pain rating, it was often perceived by patients to be a question about whether they required analgesia. This may be a feature of the ward environment and system; the measure is probably used rather differently in a research setting.

Limitations

The study has several limitations. Interviews took place at the bedside in open wards, without confidentiality, and this may have discouraged participants from disclosing sensitive issues, such as distress or loneliness. Second, there are limits to the accuracy with which people can describe their decision-making processes, being unaware of unconscious biases and subject to self-presentation to the researcher as an honest witness. Third, potential participants were identified by the nurse-in-charge as suitable, in order not to disturb those who were too ill or cognitively impaired to consent or participate, but this may have skewed selection toward more articulate or amenable patients, or those more likely to give a good account of their interactions with staff. The participant group was mainly white British and female, and so may underrepresent male viewpoints or those associated with particular ethnic groups. This may be particularly relevant in the approach to coping with pain, where culture and gender roles influence social expectations and norms, and affect preferences. However, the study has strengths in representing a range of patient diagnoses, time in pain, and ages.

Conclusion

Inpatients using the VRS combined multiple dimensions of pain in idiosyncratic ways, including sensory, affective, cognitive and functional dimensions. Each participant made sense of each VRS category, and the distances between categories, in unique ways. The VRS was widely used as a tool to express need for analgesics, and scores were adjusted according to the participant's wish for analgesia and expectations of staff. These results have implications for staff training in using the pain scale and interpreting scores, and in involving patients in this process. Pain scale ratings should not be assumed to represent simple pain intensity and need further investigation in setting such as this where they are widely used for monitoring care.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by National Health Service IRAS project ID: 16/YH/0417. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LB carried out the empirical work, wrote a long version for his doctoral thesis, and helped prepare this version for publication. AW was principal supervisor for the empirical work and wrote the first draft of the paper from LB's thesis. KH assisted the empirical work as external supervisor, and was involved in rewriting this version from the thesis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpain.2021.723520/full#supplementary-material

REFERENCES

1. IASP Subcommittee on taxonomy. Pain terms: a list with notes on usage. *Pain.* (1979) 6:249–52.
2. 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) 161:1976–82. doi: 10.1097/j.pain.0000000000001939
3. McCaffery M. *Nursing Practice Theories Related to Cognition, Bodily Pain, and Man-Environment Interactions*. Los Angeles: University of California at LA Students Store (1968).
4. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk CD, Melzack R, editors. *Handbook of Pain Assessment*. 3rd edn. London: The Guildford Press (2011). p. 19–44.
5. Huskisson EC. Measurement of pain. *Lancet.* (1974) 304:1127–31. doi: 10.1016/S0140-6736(74)90884-8

6. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain.* (1975) 1:277–99. doi: 10.1016/0304-3959(75)90044-5
7. Katz J, Melzack R. The McGill Pain Questionnaire: development, psychometric properties, and usefulness of the long form, short form, and short form-2. In: Turk DC, Melzack R, editors. *Handbook of Pain Assessment*. 3rd edn. New York, NY: Guilford Press (2011). p. 45–66.
8. Uher J. Quantitative data from ratings scales: an epistemological methodological enquiry. *Front Psychol.* (2018) 9:1–27. doi: 10.3389/fpsyg.2018.02599
9. Chiarotto A, Maxwell LJ, Ostelo RW, Boers M, Tugwell P, Terwee CB. Measurement properties of visual analogue scale, numeric rating scale, and pain severity subscale of the brief pain inventory in patients with low back pain: a systematic review. *J Pain.* (2019) 20:245–63. doi: 10.1016/j.jpain.2018.07.009
10. Karcioglu O, Topacoglu H, Dikme O, Dikme O. A systematic review of the pain scales in adults: Which to use? *Am J Emerg Med.* (2018) 36:707–14. doi: 10.1016/j.ajem.2018.01.008
11. Ho-Jin L, Yongjung C, Hyundeok J, Jae Yeong J, Young-Eun J, Jin-Tae K. Comparative study of verbal rating scale and numerical rating scale to assess postoperative pain intensity in the post anesthesia care unit. *Medicine.* (2021) 100:e24314. doi: 10.1097/MD.0000000000024314
12. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, et al. Studies comparing numerical rating scales, verbal rating scales and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage.* (2011) 41:1073–93. doi: 10.1016/j.jpainsymman.2010.08.016
13. Safikhani S, Gries KS, Trudeau JJ, Reasner D, Rüdell K, Coons SJ, et al. Response scale selection in adult pain measures: results from a literature review. *J Patient Rep Outcomes.* (2018) 2:40. doi: 10.1186/s41687-018-0053-6
14. Wood BM, Nicholas MK, Blyth F, Asghari A, Gibson S. Assessing pain in older people with persistent pain: the NRS is valid but only provides part of the picture. *J Pain.* (2010) 11:1259–66. doi: 10.1016/j.jpain.2010.02.025
15. Boonstra AM, Preuper HRS, Balk GA, Stewart RE. Cut-off points for mild, moderate, and severe pain on the visual analogue scale for pain in patients with chronic musculoskeletal pain. *Pain.* (2014) 155:2545–50. doi: 10.1016/j.pain.2014.09.014
16. Rollman GB. Signal detection theory measurement of pain: a review and critique. *Pain.* (1977) 3:187–211. doi: 10.1016/0304-3959(77)90002-1
17. Williams ACdeC, Davies HTO, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. *Pain.* (2000) 85:457–63. doi: 10.1016/S0304-3959(99)00299-7
18. Robinson-Papp J, George MC, Dorfman D, Simpson DM. Barriers to chronic pain measurement: a qualitative study of patient perspectives. *Pain Med.* (2015) 16:1256–64. doi: 10.1111/pme.12717
19. Morone NE, Weiner DK. Pain as the fifth vital sign: exposing the vital need for pain education. *Clin Ther.* (2013) 35:1728–32. doi: 10.1016/j.clinthera.2013.10.001
20. Cedraschi C, Delezay S, Marty M, Berenbaum F, Bouhassira D, Henrotin Y, et al. "Let's talk about OA pain": a qualitative analysis of the perceptions of people suffering from OA. Towards the development of a specific pain OA-related questionnaire, the Osteoarthritis Symptom Inventory Scale (OASIS). *PLoS ONE.* (2013) 8:1–10. doi: 10.1371/journal.pone.0079988
21. Main C, Keefe F, Jensen M, Vlaeyen J, Vowles K. *Fordyce's Behavioral Methods for Chronic Pain and Illness: Republished With Invited Commentaries*. Philadelphia: Wolters Kluwer Health (2015).
22. Sullivan MD, Ballantyne JC. Must we reduce pain intensity to treat chronic pain? *Int Assoc Study Pain.* (2016) 157:65–9. doi: 10.1097/j.pain.0000000000000336
23. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* (2006) 3:77–101. doi: 10.1191/1478088706qp063oa
24. Barker C, Pistrang N, Elliott R. *Research Methods in Clinical Psychology: An Introduction for Students and Practitioners*. Chichester, England: John Wiley & Sons (2002).
25. Patton MQ. *Qualitative Evaluation and Research Methods*. 2nd edition. London: Sage (1990).
26. Elliot R, Fischer CT, Rennie DL. Evolving guidelines for publication of qualitative research studies in psychology and related fields. *Br J Clin Psychol.* (1999) 38:215–29. doi: 10.1348/014466599162782
27. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member checking: a tool to enhance trustworthiness or merely a nod to validation? *Qual Health Res.* (2016) 26:1802–11. doi: 10.1177/1049732316654870
28. Vuille M, Foerster M, Foucault E, Hugli O. Pain assessment by emergency nurses at triage in the emergency department: a qualitative study. *J Clin Nurs.* (2018):669–76. doi: 10.1111/jocn.13992
29. Chow KM, Chan JC. Pain knowledge and attitudes of nursing students: a literature review. *Nurse Educ Today.* (2014) 35:366–72. doi: 10.1016/j.nedt.2014.10.019
30. Ista E, van Dijk M, van Achterberg T. Do implementation strategies increase adherence to pain assessment in hospitals? A systematic review. *Int J Nursing Stud.* (2013) 50:552–68. doi: 10.1016/j.ijnurstu.2012.11.003
31. Drake G, Williams ACdeC. Nursing education interventions for managing acute pain in hospital settings: a systematic review of clinical outcomes and teaching methods. *Pain Manage Nurs.* (2016) 18:3–15. doi: 10.1016/j.pmn.2016.11.001
32. Bell L, Duffy A. Pain assessment and management in surgical nursing: a literature review. *Br J Nurs.* (2009) 18:153–6.S doi: 10.12968/bjon.2009.18.3.39042

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pain Research

HYPOTHESIS AND THEORY
published: 02 November 2021
doi: 10.3389/fpain.2021.756680

What Is the Numerical Nature of Pain Relief?

Andrew D. Vigotsky[1,2], *Siddharth R. Tiwari*[2,3], *James W. Griffith*[2,4] *and A. Vania Apkarian*[2,5]*

[1] Departments of Biomedical Engineering and Statistics, Northwestern University, Evanston, IL, United States, [2] Center for Translational Pain Research, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States, [3] Illinois Mathematics and Science Academy, Aurora, IL, United States, [4] Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States, [5] Departments of Neuroscience, Anesthesia, and Physical Medicine & Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Pain relief, or a decrease in self-reported pain intensity, is frequently the primary outcome of pain clinical trials. Investigators commonly report pain relief in one of two ways: using raw units (additive) or using percentage units (multiplicative). However, additive and multiplicative scales have different assumptions and are incompatible with one another. In this work, we describe the assumptions and corollaries of additive and multiplicative models of pain relief to illuminate the issue from statistical and clinical perspectives. First, we explain the math underlying each model and illustrate these points using simulations, for which readers are assumed to have an understanding of linear regression. Next, we connect this math to clinical interpretations, stressing the importance of statistical models that accurately represent the underlying data; for example, how using percent pain relief can mislead clinicians if the data are actually additive. These theoretical discussions are supported by empirical data from four longitudinal studies of patients with subacute and chronic pain. Finally, we discuss self-reported pain intensity as a measurement construct, including its philosophical limitations and how clinical pain differs from acute pain measured during psychophysics experiments. This work has broad implications for clinical pain research, ranging from statistical modeling of trial data to the use of minimal clinically important differences and patient-clinician communication.

Keywords: pain, clinical trials, treatment effects, statistical models, ANCOVA

OPEN ACCESS

Edited by:
Loren Martin,
University of Toronto, Canada

Reviewed by:
Anaïs Lacasse,
Université du Québec en Abitibi Témiscamingue, Canada
Ali Khatibi,
University of Birmingham, United Kingdom

***Correspondence:**
A. Vania Apkarian
a-apkarian@northwestern.edu

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *10 August 2021*
Accepted: *21 September 2021*
Published: *02 November 2021*

Citation:
Vigotsky AD, Tiwari SR, Griffith JW and Apkarian AV (2021) What Is the Numerical Nature of Pain Relief? Front. Pain Res. 2:756680. doi: 10.3389/fpain.2021.756680

1. INTRODUCTION

Pain is highly prevalent, burdensome, and a common reason for doctor visits (1–4). In an attempt to understand the severity of patients' pain, doctors and researchers ask patients about the intensity of the their pain, requiring patients to condense and transmute their subjective experience to a single number. Despite its abstract and reductionist nature, self-reports of pain intensity are moderately-to-strongly correlated with several patient-reported outcome variables, including quality of life, disability, and more (5, 6). Moreover, self-reports of pain intensity are remarkably easy and inexpensive to collect. These pragmatic and measurement properties make a reduction in self-reported pain, which we define as pain relief, the gold standard for assessing pain improvement.

Clinical studies of pain commonly quantify pain relief as the primary outcome. However, how pain relief is quantified and reported roughly falls into one of two categories: absolute reductions in pain and relative (or percent) reductions in pain. For example, studies that report absolute reduction may state that a drug decreased pain by 2/10 numerical rating scale (NRS) units or 23/100

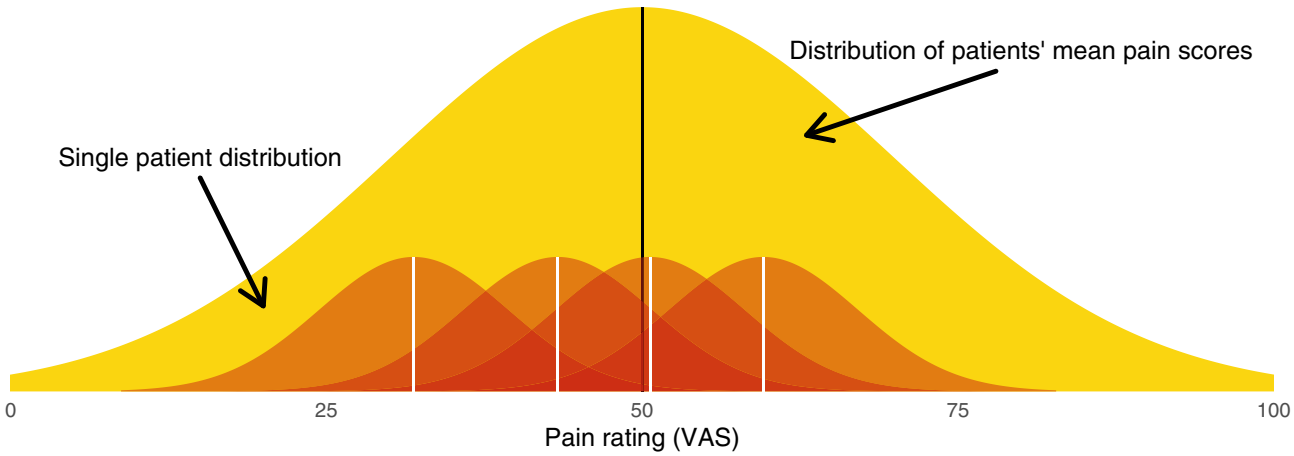


FIGURE 1 | Graphical illustration of the hierarchical model from which patients' pain scores are sampled. The broad yellow (light gray) distribution is the between-patient distribution (level 2), from which each patient's mean pain score is sampled. Each red (dark gray) distribution is a within-patient distribution (level 1), from which single measurements are sampled.

visual analog scale (VAS) units. Alternatively, studies that report relative reductions may state that pain decreased by 13% units more in the drug group relative to the placebo group. Although both approaches to reporting pain reductions are common, they are conceptually incompatible (unless baseline pain is perfectly homogeneous; see section 2). Their incompatibility begs the question as to whether one approach is more appropriate than the other.

In this paper, we aim to illuminate the issue of absolute vs. relative pain relief[1]. We rely on statistical theory to provide researchers and statistically-minded clinicians with the background necessary to understand these measurement models, for which readers are assumed to be familiar with linear regression. In addition, we empirically analyze four datasets to reinforce and make tangible our conceptual discussion.

2. STATISTICAL BACKGROUND

Whenever one uses data to make a calculation, they are building a model. Every model has assumptions, but still, models should accurately reflect the data they are intending to simplify and thus represent. With regards to modeling pain relief, when reporting absolute changes in pain, one is assuming the process is additive. Alternatively, when reporting percent changes in pain, one is assuming the process is multiplicative. These assumptions have corollaries that *prima facie* may be unclear. In this section, we aim to explain the processes that would generate each of these models and the theoretical implications of these measurement and modeling assumptions.

[1]For simplicity, herein, we will refer to self-reported pain intensity simply as pain.

2.1. Additive Model

The additive model and its implications are best understood by defining a *data-generating process*. This involves creating a mathematical model that reflects how one thinks the data are created. Because longitudinal pain relief is of interest, there is commonly at least one pain rating at the beginning of the study (x_i) and at least one or more follow-up ratings (y_i) for each subject *i*. The additive model of pain relief uses the simple difference between these pain ratings to calculate absolute pain relief ($\delta_i = y_i - x_i$), where negative δ_i's indicate relief and positive δ_i's indicate worsening of pain. Although straightforward, this is a gross oversimplification.

In reality, pain data are messy. For one, between-patient heterogeneity is appreciable—pain ratings at intake will often range from the minimum required for study entry (e.g., 4/10 NRS) to the scale's maximum (e.g., 10/10 NRS). In addition, patients' pain fluctuates from minute-to-minute, hour-to-hour, day-to-day, and so on. To complicate matters further, the process of converting a qualia to a number is undoubtedly fuzzy, meaning the pain ratings themselves will have noise associated with them. Thus, there are two sources of variance to consider: between patients and within patients. These sources of variance can be thought of hierarchically (**Figure 1**).

Between-patient heterogeneity is a natural place to start. The entire sample of patients will have a mean pain score μ. Each patient's mean at baseline, α_i, will be dispersed around this group mean according to the between-subject variance τ^2. We can say that patient means are distributed

$$\alpha_i \sim \mathcal{N}\left(\mu, \tau^2\right).$$

This distribution of patient means is illustrated in yellow in **Figure 1**.

The notion of within-patient heterogeneity implies there will be variance around each patient's mean pain. When we "sample" a patient's pain rating, we do not observe α_i; rather, we obtain a value $\alpha_i \pm \sigma$. These within-patient distributions are illustrated in red in **Figure 1**. Together, the within- and between-patient models form a hierarchical model (**Appendix A1**).

Because the patient's pre- and post-intervention pain ratings have variability associated with them, the observed difference scores are subject to regression toward the mean (RTM). RTM is a statistical phenomenon whereby higher initial scores are likely to be followed by lower measurements, and similarly, lower initial scores are likely to followed by higher measurements. For example, suppose someone's diastolic blood pressure is normally around 70 mmHg. If a doctor measures that individual's blood pressure and finds it to be 90 mmHg, it is highly probable that the next time it is measured, it will be lower than 90 mmHg. Individuals whose measurements deviate more from their mean will thus appear to undergo greater changes. In the case of a pain study, those who start off with greater pain levels will regress toward the mean, in turn creating larger change scores. This is depicted graphically in **Figure 2B**, which shows that those who have greater pre-intervention pain scores (x-axis) have smaller change scores (y-axis). Importantly, this phenomenon is purely statistical and can be explained by the reliability of the measurement.

Measurement reliability is commonly quantified using the intraclass correlation coefficient (ICC). The simplest version of the ICC is the ratio of the between-patient variance to the total variance,

$$\frac{\tau^2}{\tau^2 + \sigma^2},$$

where τ^2 is the between-patient variance and σ^2 is the within-patient variance. Since σ^2 defines the variance between individual measurements from a single patient, the ICC can be improved by using the mean of several measurements from a single patient rather than a single measurement. Doing so allows us to substitute σ^2 with the variance of the sample mean, $\frac{\sigma^2}{n}$, giving us an ICC that is a function of the number of data points sampled from each patient,

$$\frac{\tau^2}{\tau^2 + \frac{\sigma^2}{n}}.$$

Note, this quantity approaches 1 (perfect reliability) as $n \to \infty$.

Importantly, the above concepts generalize to post-intervention scores as well. If we assume τ^2 and σ^2 do not change, and instead, there is a simple shift in mean scores without ceiling and floor effects, then the ICC also defines the Pearson correlation between pre- and post-intervention scores. The Pearson correlation is useful because it gives us direct insight into RTM—the slope between the pre-intervention scores and change scores approaches zero as the correlation between pre- and post-intervention scores approaches 1 (**Figure 3**).

All of these properties come together and should be considered when statistically modeling pain relief and the effect of an intervention.

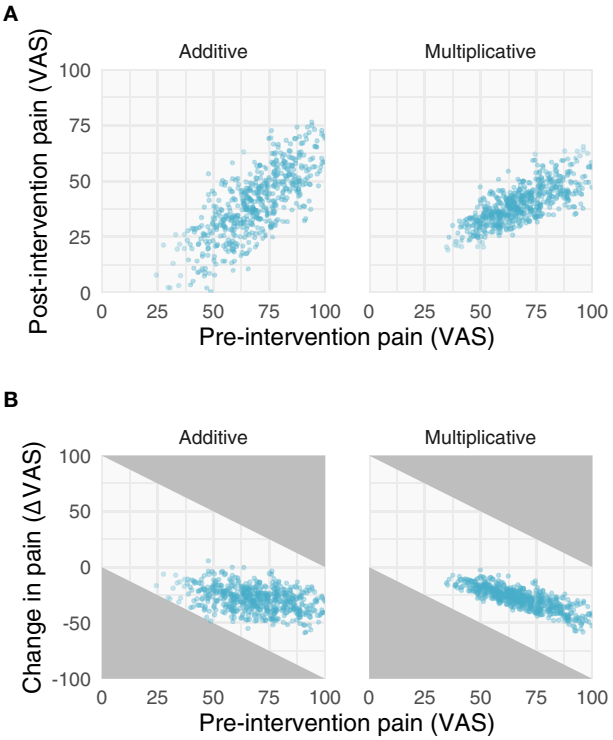


FIGURE 2 | Properties of additive and multiplicative data. We simulated data with additive (left) and multiplicative (right) assumptions. **(A)** Relationships between pre- and post-intervention pain scores when improvements are additive (left) and multiplicative (right). Note the additive post-intervention scores are relatively homoscedastic, while the variance of multiplicative post-intervention scores increases with increasing pre-intervention scores. **(B)** Negative relationships between change scores and pre-intervention scores. Gray areas in **(B)** represent regions where points are not possible due to measurement constraints; that is, because a change score cannot be > |100|.

2.2. Multiplicative Model

The multiplicative model is still mathematically simple but its implications are more complex. If pain relief is multiplicative, then it can be modeled as a relative reduction; i.e., $\phi = \frac{\delta_i}{x_i}$. This would imply that each person's post-intervention pain (y_i) is a fraction of their starting pain (x_i); i.e., $y_i = (\phi + 1)x_i$. However, ratios and relative reductions have unfavorable statistical properties. Instead, it is preferable to work on the log scale (7–9). In particular, recall $\log \frac{y_i}{x_i} = \log y_i - \log x_i$, enabling us to linearize the multiplicative process. Similarly, from this, one may realize that it is natural to model multiplicative effects as being generated from *log*-normal distributions rather than normal distributions (**Appendix A2**).

The implications of the log-normal distribution and its multiplicative properties are shown and described in **Figures 2**, **3**. Note that the multiplicative pain reductions follow a different distribution than additive effects owing to their errors compounding rather than adding. This results in a "fanning" (or heteroscedasticity) of post-intervention scores as a function of greater pre-intervention scores (**Figure 2A**). This is a hallmark

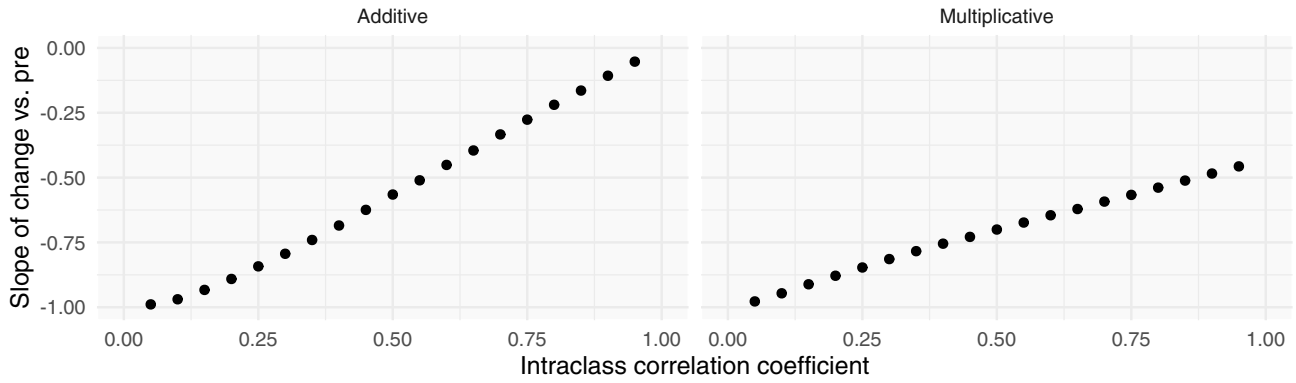


FIGURE 3 | Simulations of additive and multiplicative changes reveal the effect of different intraclass correlation coefficients on the slope between change scores and pre-intervention scores. Additive effects have slopes that trend toward zero with increasing ICC's, while multiplicative effects always have a negative slope no matter their ICC.

of multiplicative processes that can be evaluated empirically. In addition to this fanning, it is quickly apparent that even with zero measurement error (**Figure 3**), multiplicative effects can look like RTM since greater pre-intervention scores will result in greater decreases in pain (**Figure 2B**). However, as opposed to additive processes in which greater pre-intervention scores are attributable to RTM (i.e., measurement error), this relationship is indeed "real" for multiplicative processes.

The multiplicative nature does not only apply to the relationship between pre- and post-intervention pain, but also the effect of a treatment. This is described in further detail in the next subsection.

2.3. Statistical Models of Pain Relief

Randomized controlled clinical trials aim to compare pain between two groups. To do so, investigators commonly compare the absolute or percent pain relief itself (e.g., a *t*-test on the change scores). However, such analyses are ill-conceived. Instead, especially for studies that record one or few follow-up measures (as opposed to time-series), it is recommended that the data-generating process be modeled using an analysis of covariance (ANCOVA) with pre-intervention scores as a covariate (8, 10). The reasons for this are manifold:

1. The response variable in a statistical model should be the result of an experiment. Because patients enter studies with their baseline score, it is not the result of the experiment so it should not be treated as a dependent variable (e.g., like in a group×time analysis of variance).
2. Accounting for RTM. Instead of a group×time analysis of variance, one could perform a simple *t*-test on the change scores. However, such an analysis ignores RTM, and, especially in the case of baseline imbalances, can produce biased estimates. ANCOVA can adjust for such effects.
3. Improving statistical efficiency. ANCOVA has greater statistical efficiency, resulting in greater power and more precise intervals.
4. Post-intervention scores are arguably more interesting than change scores. Patients must live with the pain following the intervention, not the change in pain. However, regressing post-intervention pain *or* change in pain produces the same group effect (8).

These statistical and philosophical advantages are well-established in the biostatistics literature (8, 10–14). Note, the benefits of ANCOVA primarily apply to randomized studies, as ANCOVA may produce biased estimates in non-randomized studies depending on the allocation mechanism (15).

For the additive case, the ANCOVA model takes the form

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 g_i + \epsilon_i,$$

where $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ and g_i is dummy-coded for group (e.g., 0 = placebo and 1 = drug). β_2 is the effect of interest: the average difference in post-intervention pain scores between groups after adjusting for pre-intervention scores. β_1 will typically be < 1, indicative of RTM, and the intercept may be nonsensical unless x_i is mean-centered. Of course, like any regression, one can add more covariates, especially those with prognostic value, which will further increase statistical efficiency.

The ANCOVA can also be generalized to the multiplicative case. Since multiplicative effects can be linearized by taking the log-transform, we can write the model as

$$y_i = B_0 \cdot x_i^{\beta_1} \cdot B_2^{g_i} \cdot E_i \quad (1)$$

$$= \exp\left\{\beta_0 + \beta_1 \log x_i + \beta_2 g_i + \epsilon_i\right\} \quad (2)$$

$$\implies \log y_i = \beta_0 + \beta_1 \log x_i + \beta_2 g_i + \epsilon_i. \quad (3)$$

This model reveals a few things. First, in (1), residuals will compound with increasing values of the predicted y_i (i.e., $\hat{y}_i$). Indeed, this is consistent with what we observed in the simulations above, so this functional form can capture the compounding error. Second, in (3), both y_i and x_i are logged, so when $\beta_1 = 1$, it is equivalent to modeling the percent

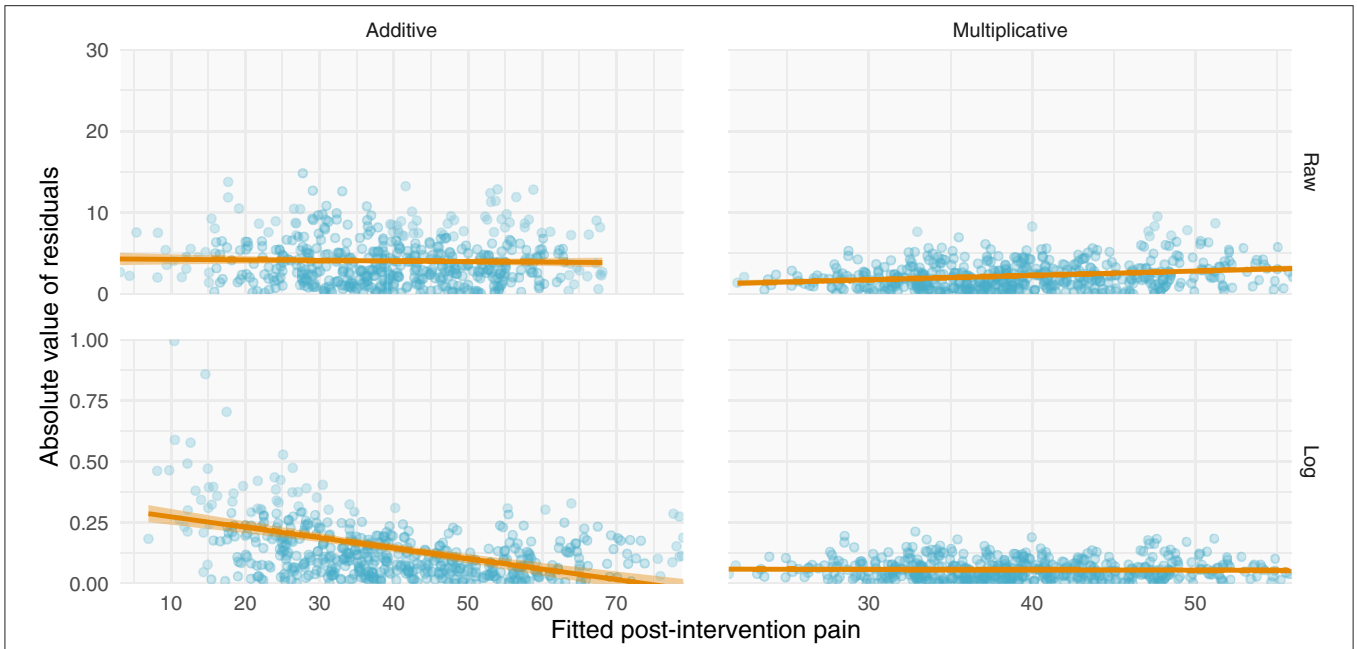


FIGURE 4 | Simulations of additive and multiplicative changes reveal differential residual behavior for raw and log-transformed ANCOVA models. **(Left)** data generated with have an additive structure have homoscedastic residuals when fit with a standard ANCOVA (top) but heteroscedastic residuals when fit with a log-transformed ANCOVA (bottom). **(Right)** data generated with a multiplicative structure have heteroscedastic residuals when fit on their raw scale (top) but homoscedastic residuals when log-transformed (bottom).

change; however, when $\beta_1 \neq 1$, there is a scaling to account for nonlinearities and RTM. Finally, B_2 is a multiplicative effect: when $B_2 = 1$, both groups are expected to have the same post-intervention score for a given pre-intervention score; when $B_2 > 1$, the experimental group is expected to have a greater post-intervention score for a given pre-intervention score; and so on. Since we are fitting β_2 rather than B_2, the fit coefficient will be on the log scale, so exponentiating the coefficient will make it more interpretable despite the log scale having nicer mathematical properties. Note, even this multiplicative ANCOVA is more efficient than analyzing percent changes (12).

3. EMPIRICAL DATA

As a proof of principle, we assessed the properties of four separate datasets. Two of the datasets were collected in patients with subacute back pain and the other two consist of patients with chronic back pain. Ideally, data are analyzed using intention-to-treat. However, here, we included individuals for whom we had enough ratings to complete our analyses as the data are being used for illustrative purposes and we are not looking to draw inferences.

3.1. Datasets

3.1.1. Placebo I (Chronic Back Pain)

3.1.1.1. Overview

The purpose of this study was to investigate factors associated with placebo analgesia in chronic pain patients (16). This was the first trial designed to study chronic pain patients receiving placebo vs. no treatment. The total duration of the study lasted ∼ 15 months. Protocol and informed consent forms were approved by Northwestern University IRB and the study was conducted at Northwestern University (Chicago, IL, USA).

3.1.1.2. Participants

To meet inclusion criteria, individuals had to be 18 years or older with a history of lower back pain for at least 6 months. This pain should have been neuropathic (radiculopathy confirmed by physical examination was required), with no evidence of additional comorbid chronic pain, neurological, or psychiatric conditions. Individuals had to agree to stop any concomitant pain medications and had to be able to use a smartphone or computer to monitor pain twice a day. Additionally, the enrolled patients had to report a pain level of at least 5/10 during the screening interview, and their averaged pain level from the smartphone app needed to be higher than 4/10 during the baseline rating period before they were randomized into a treatment group. A total of 82 patients were randomized. Here, we include 18 participants from the no treatment group and 42 participants from the placebo group for whom we had complete rating data [cf. Supplementary Figure 1 in (16)].

3.1.1.3. Pain Data

Data were collected using a custom pain rating phone app through which patients could rate their pain (0–10 NRS). Patients were asked to enter their pain 2 times/day over the course of the entire study. For the purposes of demonstration, here we averaged pain ratings within a single day.

3.1.2. Placebo II (Chronic Back Pain)

3.1.2.1. Overview

The purpose of this study was to validate a prognostic model for classifying chronic pain patients based on their predicted improvement with placebo (17). Protocol and informed consent forms were approved by Northwestern University IRB and the study was conducted at Northwestern University (Chicago IL, USA).

3.1.2.2. Participants

Individuals with chronic low back pain were recruited for this study. Patients must have had low back pain for at least 6 months, with or without symptoms of radiculopathy, a minimum VAS score of 5/10 at the screening visit and a minimum average pain of 4/10 over a 2-week period prior to their first visit. A total of 94 patients were randomized to no treatment, placebo, or naproxen. Here, we include 12 participants from the no treatment group, 33 participants from the placebo group, and 35 participants from the naproxen group for whom we had complete rating data [cf. Figure 1 in (17)].

3.1.2.3. Pain Data

Data were collected using a custom pain rating phone app through which patients could rate their pain (0–10 NRS), as in Placebo I. Patients were asked to enter their pain 2 times/day over the course of the entire study. For the purposes of demonstration, here we averaged pain ratings within a single day.

3.1.3. Levodopa Trial (Subacute Back Pain)

3.1.3.1. Overview

The purpose of this trial was to investigate whether levodopa (l-DOPA) can block patients' transition from subacute to chronic back pain (18). This 24-week double-blind parallel group randomized controlled trial was conducted at Northwestern University (Chicago, IL, USA). Protocol and informed consent form were approved by Northwestern University IRB as well as NIDCR/NIH. All enrolled participants provided written informed consent. The trial was registered on ClinicalTrials.gov, under registry NCT01951105.

3.1.3.2. Participants

Individuals with a recent onset of low back pain were recruited. Criteria for enrollment included history of low back pain with a duration between 4 and 20 weeks with signs and symptoms of radiculopathy and average reported pain intensity $>$ 4 (on an NRS scale from 0 to 10) on the week before baseline assessments and the week preceding treatment start. Participants were randomized to one of three groups: no treatment (completed $n = 10$), naproxen + placebo ($n = 28$), naproxen + l-DOPA/c-DOPA ($n = 21$). Here, we will use data from 47 patients who had complete rating data (naproxen + placebo = 27; naproxen + l-DOPA/c-DOPA = 20) [cf. Figure 1B in (18)].

3.1.3.3. Pain Data

Data were collected using a custom pain rating phone app through which patients could rate their pain (0–10 NRS). Patients were asked to enter their pain 3 times/day over the course of the entire study (28 weeks). For the purposes of demonstration, here we averaged pain ratings within a single day.

3.1.4. Prospective Cohort (Subacute Back Pain)

3.1.4.1. Overview

The purpose of this study was to identify predictive biomarkers to identify individuals who will vs. will not recover from subacute back pain (19). Protocol and informed consent forms were approved by Northwestern University IRB as well as NIDCR/NIH, and the study was conducted at Northwestern University (Chicago, IL, USA). All enrolled participants provided written informed consent. All participants were right-handed and were diagnosed by a clinician for back pain. An additional list of criteria was imposed including: pain intensity $>$ 40/100 on the visual analog scale (VAS) and duration $<$ 16 weeks.

3.1.4.2. Participants

Eighty individuals with a recent onset (within 16 weeks) of lower back pain and an average reported pain intensity $>$ 40/100 (on the VAS) who completed at least three follow-up visits (i.e., 30 weeks following the initial visit).

3.1.4.3. Pain Data

Data were collected at five separate visits using the short form of the McGill Pain Questionnaire (MPQ). The computed sensory and affective scores from the MPQ for each visit are used as individual pain scores for each subject.

3.2. Data Properties

To evaluate whether each dataset was more compatible with an additive or multiplicative process, we conducted the same analyses from the Statistical Background section (**Figures 2**–**4**) on these data. In particular, we investigated properties of the raw and log-transformed data, in addition to the properties of ANCOVAs fit to the data. To do so, all data were converted to a 0–100 scale. Before log-transforming, we added 1 to the raw scores to avoid log(0)=`NaN`. In doing so, we demonstrate how the aforementioned principles apply to real data.

All datasets have positive relationships between pre- and post-intervention scores (**Figure 5**). Interestingly and in contrast to the other studies, the variance of the post-intervention scores in the levodopa trial appears to increase with greater pre-intervention scores, consistent with a multiplicative effect. Finally, with the exception of the prospective cohort study, there are negative relationships between changes in pain and pre-intervention scores. These negative relationships may be explained by multiplicative effects or RTM. Further examination is needed to ascertain the nature of these data.

Including more points in the calculation of pre-intervention and post-intervention scores increases the ICC, thereby increasing the reliability and decreasing the effect of RTM (**Figure 3**). Since three of the four datasets contained ecological momentary assessments of pain, we were able to sample and average more than one point from the beginning and end of each study. We averaged an increasing number of a pre- and post-intervention points and recalculated the slope between

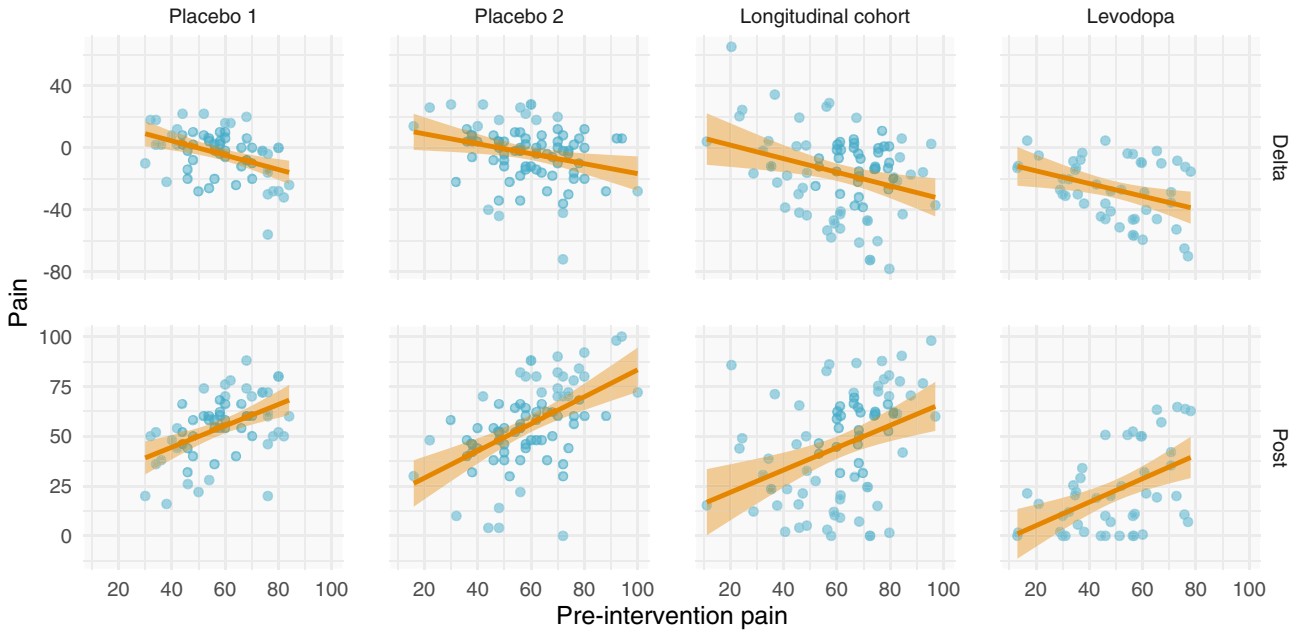


FIGURE 5 | Relationships between pre-intervention scores and change scores (top) and post-intervention scores (bottom). **(Top)** Relationship between pre-intervention scores and change scores. Note that most of the studies have a negative relationship. This could be explained by regression toward the mean or multiplicative effects, in addition to ceiling/floor effects. **(Bottom)** Relationship between pre-intervention and post-intervention pain scores across all studies. Each study shows a positive relationship between pre- and post-intervention scores; however, the Levodopa study appears to have greater variance in post-intervention scores with greater pre-intervention scores.

change score and pre-intervention score (i.e., plot from **Figure 5**, top). If the slopes strongly trend toward zero by increasing the number of points, this indicates that the data have additive properties. Slopes that stay negative regardless of increasing reliability (number of points) indicate that the data may be multiplicative. For the studies included in this analysis (Placebo I, Placebo II, Levodopa Trial), Placebo I and Placebo II's slopes have slight upward trends: as the number of points in the calculation of pre-intervention and post-intervention scores increases, the negative slope due to RTM increases. In contrast, the Levodopa trial's negative slopes remain stable (**Figure 6**). This again hints at the notion that the levodopa trial's data may be multiplicative, while Placebo I and Placebo II may be additive.

Perhaps the most direct assessment of additive vs. multiplicative properties is to model the data and assess the model fits. When assessing and utilizing a model, one should ensure that the model's assumptions are met and that the model captures salient features of the data. Because multiplicative data-generating processes lead to compounding residuals, we can observe these effects when fitting ANCOVAs. In **Figure 7**, we focus specifically on the variance observed in **Figure 5**, illustrating the relationship between fitted values (using the ANCOVA models from **Figure 5**) and the absolute value of the residuals. As shown in **Figure 2**, multiplicative relationships possess higher variance as pre-intervention scores increase, compared to additive relationships which are homoscedastic. For this reason, we should observe a null correlation between fitted values and absolute residual error for data that have exhibited additive properties (Placebo I, Placebo II, Prospective Cohort) thus far, and observe a positive correlation between fitted values and absolute residual error for data that have exhibited multiplicative properties (Levodopa Trial). As predicted, the Placebo I, Placebo II, and Prospective Cohort data all display this additive quality, as their residual error does not increase as fitted values increase. In contrast, the Levodopa Trial data display multiplicative properties, as its residual error increases as fitted values increase. The description and analyses of these data can be seen below (**Figure 7**).

From these plots, it is clear that the Placebo I, Placebo II, Prospective Cohort demonstrate additive properties while the Levodopa Trial demonstrates multiplicative properties. An understanding of these concepts and model assumptions have real implications. In **Table 1**, we include the average absolute (additive) and log-transformed (multiplicative) change in pain scores for each dataset. As an example, the effect of naproxen relative to no treatment in Placebo II is −15 (−27, −3) for the additive model but 0.7 (0.4, 1.1) for the multiplicative model. The 95% CI is much wider for the multiplicative model since it is misspecified, which in turn may lead an investigator or clinician to be less certain conclusions about the treatment effect.

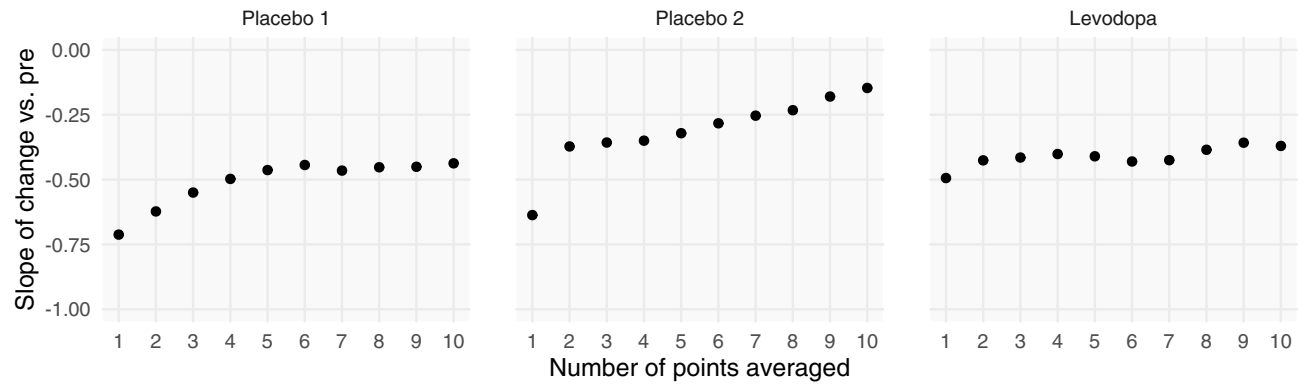


FIGURE 6 | Increasing the number of points used for each patient's pre- and post-intervention scores increases the slope between change scores and pre-intervention scores. Each patient's pre- and post-intervention scores were calculated using the mean of *x* points. By averaging over more points, we should increase the intraclass correlation coefficient. Negative slopes between change scores and pre-intervention scores are indicative of one of two things: (1) regression toward the mean or (2) multiplicative effects. In the datasets that show evidence of being additive, we see marked increases in slopes, indicating that we are decreasing regression toward the mean by including more points. However, because the Levodopa Trial displays multiplicative properties, it is only minimally affected by adding more points.

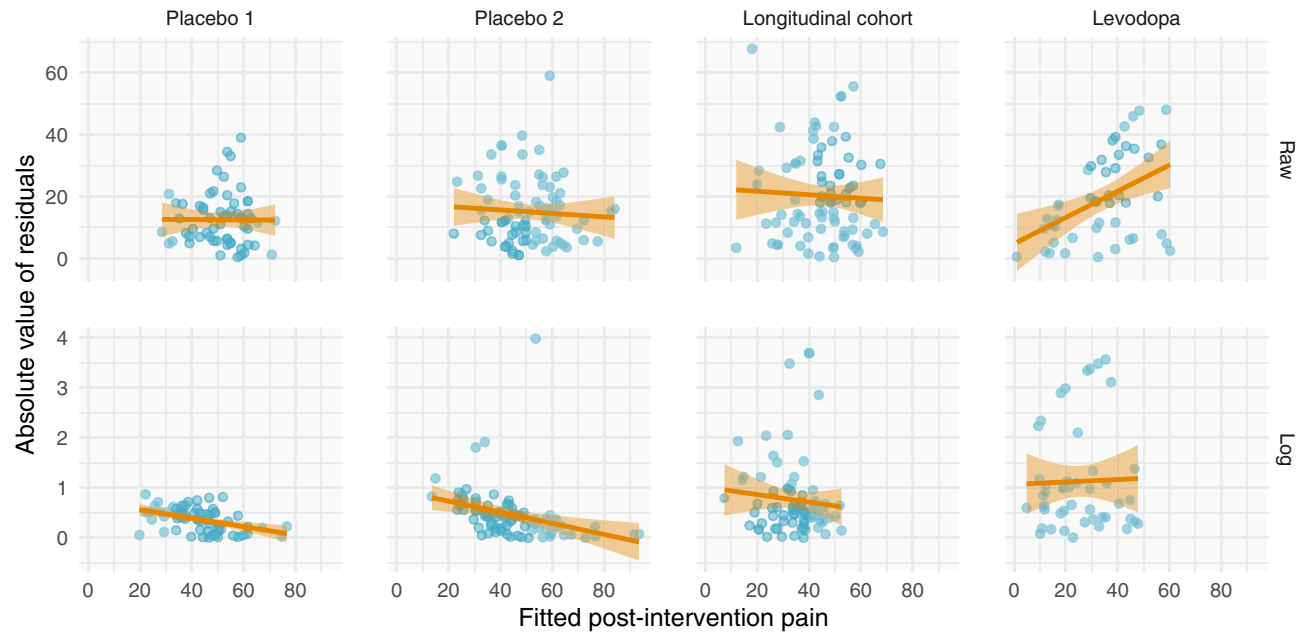


FIGURE 7 | Absolute values of residuals from additive ANCOVA models. We fit an ANCOVA to each dataset using pre-intervention score and group membership as covariates. From these models, we plotted the absolute values of the residuals as a function of the fitted value. Additive models should be homoscedastic, meaning the magnitudes of the residuals do not change as a function of the response variable. However, multiplicative models have compounding error, such that if you fit them using an additive model, greater predicted values will be associated with larger magnitudes of residual error. Placebo I, Placebo II, and the Prospective Cohort study all exhibit features of additive data. However, the Levodopa Trial exhibits multiplicative properties, as evidenced by the increasing error residual magnitude with increasing fitted values.

4. DISCUSSION

Pain relief is a ubiquitous clinical trial outcome with direct treatment implications. Treatments that yield appreciable pain relief will be employed in the clinic, and findings from these trials may be communicated to patients. However, if data from trials are not properly modeled, then the resulting treatment effects may be both biased and highly variable, which in turn may mislead researchers, clinicians, and patients. In this theory-based paper, we have emphasized the difference between additive and multiplicative treatment effects from mathematical, statistical, and empirical perspectives.

TABLE 1 | Additive and multiplicative effects by dataset.

Dataset	Additive model (NRS), $\hat{\beta}$ ($CI_{95\%}$)	Multiplicative model (AU), $\hat{\beta}$ ($CI_{95\%}$)
Placebo I	−3 (−12, 5)	0.9 (0.8, 1.1)
Placebo II	Placebo: −9 (−21, 4) Naproxen: −15 (−27, −3)	Placebo: 0.8 (0.5, 1.3) Naproxen: 0.7 (0.4, 1.1)
Levodopa trial	4 (−7, 15)	1.5 (0.7, 3.3)

All effects were modeled using ANCOVA with pre-intervention scores as a covariate. Multiplicative effects use the log-transformed scores and represent the exponentiated coefficients which can be interpreted as the relative effect of treatment group vs. the control group [e.g., post-intervention pain in the placebo group (Placebo I) will be 90% of the post-intervention pain in the no treatment group].

It is clear that the assumptions behind these effects are not interchangeable and thus should be more thoughtfully considered when planning and analyzing clinical trial data. Moreover, how pain relief is conceptualized will propagate into the interpretation of effects, which we briefly discuss herein.

4.1. Minimal Clinically Important Differences

Pain intensity ratings can be difficult to interpret—they are a reductionist, unidimensional measurement intended to capture a single aspect of a private, complex, incommunicable experience (20, 21). To help make sense of improvements, researchers and clinicians commonly rely on minimal clinically important differences (MCID). In clinical pain research, MCIDs are commonly derived by mapping changes in pain ratings onto a different scale, such as global impression of change (22). For example, what absolute change in NRS and relative change in NRS correspond to "much improved"? This mapping is then commonly used as a guidepost for interpreting other studies, and in some cases, individual patient changes (23).

Although commonly derived and used without justification, absolute and relative MCIDs are not interchangeable since they are mathematically incompatible with one another. Suppose patient A starts with an 8/10 pain and patient B starts with a 4/10 pain. If the treatment has an additive effect, both patients may improve by 2/10, but this would result in markedly different percent reductions: 25 and 50% for patients A and B, respectively. Farrar et al. (22) suggest that an MCID for pain relief is 2/10 NRS or 30%; here, these would yield two different conclusions since both patients achieved a 2/10 decrease but only one patient achieved a 30% decrease. Much attention has been and continues to be given to both additive and multiplicative MCIDs without considering the conceptual difference between the two. This conceptual incompatibility needs to be reconciled if MCIDs are to be used in a meaningful way. However, there are also larger issues that warrant addressing.

Across studies and ignoring the numerical nature of treatment effects, MCIDs have a linear relationship with baseline pain ratings, with an x-intercept corresponding to roughly 30/100 and a slope of 1 (i.e., MCID $\approx$ baseline $-$ 30) (24). This relationship calls into question both absolute and relative MCIDs. If absolute MCIDs were valid, then we would expect the MCID to be constant across all baseline pain scores. If relative MCIDs were valid, then we would expect a y-intercept of 0 and a slope equal to the MCID. Rather, this relationship suggests MCIDs are more compatible with a post-intervention pain rather than a change score, and this post-intervention pain is equal to 30/100. In other words, the MCID is the change in pain needed to obtain a 30/100. If true, this would be consistent with the idea that it is a patient's pain, not change in pain, that is important.

More generally, MCIDs arguably represent a conflation of constructs. MCIDs typically involve dichotomizing a measurement by mapping it onto some other measurement using some loss function—a form of "dichotomania" (25). For example, researchers may threshold and dichotomize changes in VAS into improvement vs. non-improvement using the global impression of change scale (22). This dichotomization of pain scores is then applied to other studies. Yet, such an approach is curious—it implies we are actually interested in global impression of change but use pain scores as a noisy proxy. If a researcher is interested in global impression of change, they should measure global impression of change as an outcome in their sample. Further, the ontological basis for dichotomous change scores is arguably ill-conceived. The insipid use of MCIDs in pain research and practice deserves greater scrutiny. From this perspective, it has been argued that greater context is needed in deriving metrics of clinical importance (26, 27) for which decision theory may provide a rigorous foundation.

In addition to using MCIDs for interpreting findings, researchers have used MCIDs for "responder analysis." For example, a researcher may split patients into groups of "responders" and "non-responders" based on whether their change in pain exceeded the MCID [see section 4.5 in (23)]. However, such analyses have undesirable properties on both the individual and group levels. On the individual level, inferences cannot be made regarding response magnitude for several reasons. First, individual counterfactuals are not observed in parallel group trials; for example, we do not know what an individual's pain would have been had they been randomized to the placebo group instead of the drug group. An individual's observed improvement or worsening may have been due to the intervention or alternatively, RTM, natural history, or some other unmeasured, stochastic process. Second, the individual may not reliably attain the same improvement each time the trial is performed; for example, 60% of individuals may respond 100% of the time *or* 100% of individuals may respond 60% of the time (or some mixture of the two). Third, this dichotomization assumes an improvement of, say, 30 and 100% are equivalent, and similarly, that an improvement of 29 and 0% are equivalent (assuming MCID = 30%) by treating improvements as a binary step function rather than continuous—such an assumption strains credulity. These issues have been previously discussed in great detail (28–31). On the group level, dichotomizing individual responses turns each patient's pain improvement into a 0 ("non-responder") or 1 ("responder"), which discards information and, in turn, markedly decreases statistical efficiency and power (32, 33). Thus, the dichotomization of improvements

is arguably unethical since it discards information, effectively decreasing the sample size (32) and, in turn, the ability to quantify (or rule out) meaningful intervention effects. Rather than being treated as an analytical tool, MCIDs are perhaps better viewed from an interpretive and decision-making perspective.

Notwithstanding MCID's limitations, it is perhaps most useful at the planning stage of clinical research. A clinically important difference is just one approach to justifying an effect size of interest for a study (34), which may be used for sample size calculations or stopping rules in adaptive trials. However, beyond planning, dichotomizing trial and especially individual patient outcomes using an MCID is a questionable practice that commonly ignores context and variability (9).

4.2. Scale Assumptions

Psychological measurement scales have a rich history across the fields of psychometrics and psychophysics (35). Anchors determine the extremes within which a participant must rate their experience, ultimately constraining the measurement construct and how accurately participants understand what they are rating (36). Bounded by these anchors, the measurements themselves can be on one of a number of scales: nominal, ordinal, interval, ratio, and absolute. Nominal scales assume a one-to-one mapping between the desired quantity x' and the measured quantity x; ordinal scales assume a monotonic mapping; interval scales assume an affine mapping ($x' = ax + b$); ratio scales assume a linear mapping with an absolute zero ($x' = ax$); and absolute scales assume a perfect mapping ($x' = x$) (37). Several renowned psychophysicists have argued—not without criticism (38, 39)—that perceptual ratings are or can easily be converted to ratio scale (35, 37). Importantly, the additive and multiplicative models rely on interval and ratio assumptions, respectively. Thus, the validity of these assumptions for clinical pain must be considered.

The numerical nature of clinical pain is an open, controversial, and perhaps unanswerable question. Early psychophysics work argues that VAS and NRS pain scales are ratio for both experimental and clinical pain. Price et al. (40) used cross-modality matching to argue that clinical pain, like heat pain, is a ratio scale. However, by mapping clinical pain onto heat pain, this finding is arguably tautological—they assessed whether clinical pain-matched heat pain follows the same power law as heat pain. Others have used item-response theory to argue that pain ratings are ordinal scale (nonlinear) rather than ratio or interval scale (41). Since the authors used unidimensional measures and a Rasch model, this conclusion is based on stationarity assumptions and ratings' reliability, which are not necessary conditions for interval or ratio scales. Although the perceptual ratings from psychophysics are undoubtedly related to clinical pain, assessing the measurement properties of clinical pain is much more complex since we cannot precisely control the sensory input. Thus, clinical pain measurement scale assumptions arguably cannot be rigorously evaluated, reinforcing that they are indeed assumptions. However, the strength of assumption varies, with interval scales (additive) having weaker assumptions than ratio scales (multiplicative). The assumptions a researcher makes directly affects the model they should choose.

TABLE 2 | Hallmarks of additive and multiplicative effects.

Plot	Additive	Multiplicative
Slope of change score vs. pre-intervention score (y) vs. number of points (x)	Slopes approach zero as the number of points utilized in calculating pre- and post-intervention pain scores increases by increasing ICC (**Figure 3**, left).	Slopes increase minimally with increasing number of points (**Figure 3**, right).
Absolute value of residuals (y) vs. fitted values (x)	No relationship between absolute residual error and fitted (post-intervention) values.	Positive, heteroscedastic relationship between absolute residual error and fitted (post-intervention) values.

4.3. Statistical Modeling and Applications

The choice of a statistical model can greatly affect the inferences drawn from the same dataset. Here, we observed that applying a multiplicative model to a dataset that exhibits additive properties can create wide CIs, making it difficult to interpret the results of an experiment (**Table 1**). This is consistent with the idea that a properly specified model will be more statistically efficient (12), and perhaps most importantly, it will better represent the underlying data.

We presented two ways of modeling data: additively and multiplicatively. Both rely on ANCOVA, with the former using raw pain scores and the latter using log-transformed pain scores. These models have different assumptions about the underlying data and, as a result, have different interpretations. If authors feel the linearity and ratio assumptions are too strict, there are other models that can be used; e.g., ordinal regression and semiparametric (or nonparametric) ANCOVA (42), in addition to intensive longitudinal and time-series analysis (43). Indeed, there are good examples in the pain literature of ANCOVA-type models being implemented with more complicated data structures [e.g., multiple study endpoints, see (44)]. In any case, researchers should be aware of the assumptions of their statistical models of the properties of their data, and of course, researchers are encouraged to collaborate with statisticians (45).

4.4. Recommendations

We have clearly demonstrated the mathematical, conceptual, and interpretive differences between additive and multiplicative effects. From this explication, there are tangible takeaways and recommendations for clinical researchers. Specifically, we suggest that researchers include and consider the following:

1. When deciding which metric to use—absolute pain decreases or percent pain decreases—use the data as a guide unless there is a principled reason to choose one or the other. Since it is unclear what influences the presence of additive or multiplicative characteristics in

pain data, it is safer to use the metric that accurately represents the properties of the data. **Table 2** summarizes the differences between additive and multiplicative properties. In time, we may develop a better understanding of pain conditions and improvements such that more general recommendations can be provided. We view this data-driven approach as being no different than checking statistical model assumptions.

2. When reporting descriptive statistics, use the arithmetic mean to calculate between-subject (average) intervention for additive data; conversely, use geometric mean for multiplicative data.
3. Ensure that patients' pre-intervention scores are heterogeneous for drawing conclusions about the nature of the data. By including a wide range of pre-intervention scores, it makes the additive or multiplicative properties more apparent. If the data are not heterogeneous, false conclusions may be made about the data's additive or multiplicative properties.

5. CONCLUSION

The properties of changes in self-reported pain are commonly implicitly assumed to be additive, multiplicative, or are conflated. Ignoring the properties of pain relief can result in model mis-specification, in turn leading to bias and statistical inefficiency. These errors further propagate into metrics such as minimal clinically important differences. We contend that more attention should be paid to the statistical properties of pain relief to ensure model assumptions are met. By paying closer attention to these properties, we can gain more insight from and make better use of data from pain clinical trials.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: OpenPain.org.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Northwestern University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AV, ST, and AVA conceptualized the paper. AV drafted the paper. AV and ST produced the figures. AV, ST, JG, and AVA read, provided feedback on, approved the final version of, and agree to be accountable for the contents of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the National Institutes of Health (1P50DA044121-01A1). This material was based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1324585.

REFERENCES

1. Gaskin DJ, Richard P. The economic costs of pain in the United States. *J Pain*. (2012) 13:715–24. doi: 10.1016/j.jpain.2012.03.009
2. St Sauver JL, Warner DO, Yawn BP, Jacobson DJ, McGree ME, Pankratz JJ, et al. Why patients visit their doctors: assessing the most prevalent conditions in a defined American population. *Mayo Clin Proc*. (2013) 88:56–67. doi: 10.1016/j.mayocp.2012.08.020
3. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. (2020) 396:1204–22. doi: 10.1016/S0140-6736(20)30925-9
4. Mäntyselkä P, Kumpusalo E, Ahonen R, Kumpusalo A, Kauhanen J, Viinamäki H, et al. Pain as a reason to visit the doctor: a study in Finnish primary health care. *Pain*. (2001) 89:175–80. doi: 10.1016/S0304-3959(00)00361-4
5. Yazdi-Ravandi S, Taslimi Z, Jamshidian N, Saberi H, Shams J, Haghparast A. Prediction of quality of life by self-efficacy, pain intensity and pain duration in patient with pain disorders. *Basic Clin Neurosci*. (2013) 4:117–24.
6. Nasution IK, Lubis NDA, Amelia S, Hocin K. The correlation of pain intensity and quality of life in chronic LBP patients in Adam Malik general hospital. *IOP Conf Ser*. (2018) 125:012183. doi: 10.1088/1755-1315/125/1/012183
7. Keene ON. The log transformation is special. *Stat Med*. (1995) 14:811–9. doi: 10.1002/sim.4780140810
8. Senn S. Change from baseline and analysis of covariance revisited. *Stat Med*. (2006) 25:4334–44. doi: 10.1002/sim.2682
9. Senn S, Julious S. Measurement in clinical trials: a neglected issue for statisticians? *Stat Med*. (2009) 28:3189–209. doi: 10.1002/sim.3603
10. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. *BMJ*. (2001) 323:1123–4. doi: 10.1136/bmj.323.7321.1123
11. Vickers AJ. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. *BMC Med Res Methodol*. (2001) 1:1–4. doi: 10.1186/1471-2288-1-6
12. Vickers AJ. *Change/Percent Change From Baseline*. Wiley StatsRef: Statistics Reference Online. (2014). doi: 10.1002/9781118445112.stat06963
13. Frison L, Pocock SJ. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Stat Med*. (1992) 11:1685–704. doi: 10.1002/sim.4780111304
14. Borm GF, Fransen J, Lemmens WA. A simple sample size formula for analysis of covariance in randomized clinical trials. *J Clin Epidemiol*. (2007) 60:1234–8. doi: 10.1016/j.jclinepi.2007.02.006
15. Van Breukelen GJ. ANCOVA versus change from baseline: more power in randomized studies, more bias in nonrandomized studies. *J Clin Epidemiol*. (2006) 59:920–5. doi: 10.1016/j.jclinepi.2006.02.007
16. Vachon-Presseau E, Berger SE, Abdullah TB, Huang L, Cecchi GA, Griffith JW, et al. Brain and psychological determinants of placebo pill response in chronic pain patients. *Nat Commun*. (2018) 9:1–15. doi: 10.1038/s41467-018-05859-1
17. Vachon-Presseau E, Abdullah TB, Berger SE, Huang L, Griffith JW, Schnitzer TJ, et al. Validating a biosignature predicting placebo pill response in

chronic pain in the settings of a randomized controlled trial. *Pain*. (2021). doi: 10.1097/j.pain.0000000000002450. [Epub ahead of print].
18. Reckziegel D, Ttreault P, Ghantous M, Wakaizumi K, Petre B, Huang L, et al. Sex-specific pharmacotherapy for back pain: a proof-of-concept randomized trial. *Pain Ther*. (2021). doi: 10.1007/s40122-021-00297-2
19. Baliki MN, Petre B, Torbey S, Herrmann KM, Huang L, Schnitzer TJ, et al. Corticostriatal functional connectivity predicts transition to chronic back pain. *Nat Neurosci*. (2012) 15:1117–9. doi: 10.1038/nn.3153
20. Chapman C, Loeser J. *Issues in Pain Measurement*. New York, NY: Raven Press (1989).
21. Turk D, Melzack R. *Handbook of Pain Assessment*. New York, NY: Guilford Press (2011).
22. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. (2001) 94:149–58. doi: 10.1016/S0304-3959(01)00349-9
23. Smith SM, Dworkin RH, Turk DC, McDermott MP, Eccleston C, Farrar JT, et al. Interpretation of chronic pain clinical trial outcomes: IMMPACT recommended considerations. *Pain*. (2020) 161:2446–61. doi: 10.1097/j.pain.0000000000001952
24. Olsen MF, Bjerre E, Hansen MD, Tendal B, Hilden J, Hrbjartsson A. Minimum clinically important differences in chronic pain vary considerably by baseline pain and methodological factors: systematic review of empirical studies. *J Clin Epidemiol*. (2018) 101:87–106. doi: 10.1016/j.jclinepi.2018.05.007
25. Senn SJ. Dichotomania: an obsessive compulsive disorder that is badly affecting the quality of analysis of pharmaceutical trials. In: *Proceedings of the International Statistical Institute, 55th Session*. Sydney, NSW (2005).
26. Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Nascimento DP, et al. A critical review of methods used to determine the smallest worthwhile effect of interventions for low back pain. *J Clin Epidemiol*. (2012) 65:253–61. doi: 10.1016/j.jclinepi.2011.06.018
27. Ferreira M. Research Note: The smallest worthwhile effect of a health intervention. *J Physiother*. (2018) 64:272–4. doi: 10.1016/j.jphys.2018.07.008
28. Senn S. Individual therapy: new dawn or false dawn? *Drug Inform J*. (2001) 35:1479–94. doi: 10.1177/009286150103500443
29. Senn S. Individual response to treatment: is it a valid assumption? *BMJ*. (2004) 329:966–8. doi: 10.1136/bmj.329.7472.966
30. Senn S. Mastering variation: variance components and personalised medicine. *Stat Med*. (2016) 35:966–77. doi: 10.1002/sim.6739
31. Senn S. Statistical pitfalls of personalized medicine. *Nature*. (2018) 563:619–21. doi: 10.1038/d41586-018-07535-2
32. Cohen J. The cost of dichotomization. *Appl Psychol Meas*. (1983) 7:249–53. doi: 10.1177/014662168300700301
33. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ*. (2006) 332:1080. doi: 10.1136/bmj.332.7549.1080
34. Cook JA, Julious SA, Sones W, Hampson LV, Hewitt C, Berlin JA, et al. DELTA2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. *BMJ*. (2018) 363:k3750. doi: 10.1136/bmj.k3750
35. Stevens SS. *Psychophysics: Introduction to Its Perceptual, Neural and Social Prospects*. New York, NY: Routledge (2017). doi: 10.4324/9781315127675
36. Yokobe J, Kitahara M, Matsushima M, Uezono S. Preference for different anchor descriptors on visual analogue scales among Japanese patients with chronic pain. *PLoS ONE*. (2014) 9:e99891. doi: 10.1371/journal.pone.0099891
37. Bolanowski SJ Jr, Gescheider GA, editors. Introduction to conference on ratio scaling of psychological magnitudes. In: *Ratio Scaling of Psychological Magnitude*. Hillsdale, NJ: Psychology Press (2013). doi: 10.4324/9780203772201
38. Ellermeier W, Faulhammer G. Empirical evaluation of axioms fundamental to Stevens's ratio-scaling approach: I. Loudness production. *Percept Psychophys*. (2000) 62:1505–11. doi: 10.3758/BF03212151
39. Zimmer K. Examining the validity of numerical ratios in loudness fractionation. *Percept Psychophys*. (2005) 67:569–79. doi: 10.3758/BF03193515
40. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain*. (1983) 17:45–56. doi: 10.1016/0304-3959(83)90126-4
41. Kersten P, White PJ, Tennant A. Is the pain visual analogue scale linear and responsive to change? An exploration using Rasch analysis. *PLoS ONE*. (2014) 9:e99485. doi: 10.1371/journal.pone.0099485
42. Harrell FE. *General Aspects of Fitting Regression Models*. Cham: Springer International Publishing (2015).
43. Fitzmaurice G. *Applied Longitudinal Analysis*. Hoboken, NJ: Wiley (2011). doi: 10.1002/9781119513469
44. Mathieson S, Maher CG, McLachlan AJ, Latimer J, Koes BW, Hancock MJ, et al. Trial of pregabalin for acute and chronic sciatica. *N Engl J Med*. (2017) 376:1111–20. doi: 10.1056/NEJMoa1614292
45. Sainani KL, Borg DN, Caldwell AR, Butson ML, Tenan MS, Vickers AJ, et al. Call to increase statistical collaboration in sports science, sport and exercise medicine and sports physiotherapy. *Br J Sports Med*. (2021) 55:118–22. doi: 10.1136/bjsports-2020-102607

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

A. DATA GENERATING PROCESSES

A.1. Additive Model

The additive model can be conceptualized hierarchically. First, we will assume each individual's average pre-intervention pain, α_i for patient i, is sampled from a larger population,

$$\alpha_i \sim \mathcal{N}\left(\mu, \tau^2\right).$$

Since α_i represents an individuals *average* pre-intervention pain, it is a latent construct and ignores measurement error and natural pain variability; for example, minute-to-minute, hour-to-hour, and day-to-day fluctuations in pain intensity. In actuality, an experiment will sample an individual's pain ratings and will be affected by measurement error. Thus, a given measurement of a patient's pre-intervention pain will be

$$x_{ij} = \alpha_i + \epsilon_{ij},$$

where $\epsilon_{ij} \sim \mathcal{N}\left(0, \sigma^2\right)$ for measurement j from patient i, assuming all patients have the same within-patient variability (**Figure 1**). If we sample and average n measurements from patient i, we obtain

$$x_{i\cdot} \sim \mathcal{N}\left(\alpha_i, \frac{\sigma^2}{n}\right).$$

Similarly, assuming homogeneous improvement and treatment effects, the average post-intervention pain rating for patient i is

$$y_{i\cdot} \sim \mathcal{N}\left(\alpha_i + \delta + \theta g_i, \frac{\sigma^2}{n}\right),$$

where δ is the improvement in the control group, θ the treatment effect of interest, and g_i is a dummy variable for group (0 = control; 1 = intervention). Without loss of generality via the additive assumption of treatment effects, we will ignore treatment groups (θ) to simplify the problem and describe the properties of these distributions, giving us the simplified post-intervention pain distribution

$$y_{i\cdot} \sim \mathcal{N}\left(\alpha_i + \delta, \frac{\sigma^2}{n}\right).$$

For both the pre and post model, the intraclass correlation coefficient (ICC) is

$$ICC = \frac{\tau^2}{\tau^2 + \frac{\sigma^2}{n}},$$

which is also the correlation between pre- and post-intervention scores. Luckily, ICC is sensitive to the number of data points from which each patient's pre- and post-intervention mean pain scores are calculated,

$$\lim_{n\to\infty} \frac{\sigma^2}{n} = 0 \implies \lim_{n\to\infty} \frac{\tau^2}{\tau^2 + \frac{\sigma^2}{n}} = 1.$$

With more data points, the slope attributable to RTM disappears. Since the ICC is equivalent to a Pearson's r in this case, we can write the joint pre-post distribution of averaged pain scores can be written as a multivariate normal,

$$\begin{pmatrix} x_{i\cdot} \\ y_{i\cdot} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu + \delta \end{pmatrix}, \begin{bmatrix} \tau^2 + \frac{\sigma^2}{n} & \tau^2 \\ \tau^2 & \tau^2 + \frac{\sigma^2}{n} \end{bmatrix}\right).$$

A.2. Multiplicative Model

The log-normal distribution is an exponentiated normal distribution, meaning the log of the log-normal distribution is a normal distribution. Therefore, we have

$$\log \alpha_i \sim \mathcal{N}\left(\log\left(\frac{\mu^2}{\sqrt{\mu^2 + \tau^2}}\right), \log\left(1 + \frac{\tau^2}{\mu^2}\right)\right).$$

And like the additive case, a single pre-intervention score j for patient i can be described as being centered around their individual mean,

$$\log x_{ij} \sim \mathcal{N}\left(\log \alpha_i, \frac{\sigma}{\mu}\right).$$

Similarly, a patient's post-intervention pain is scaled rather than shifted by the change in pain, δ,

$$\log y_{ij} \sim \mathcal{N}\left(\log \alpha_i + \log\left(1 + \frac{\delta}{\mu}\right), \frac{\sigma}{\mu}\right).$$

frontiers in Pain Research

ORIGINAL RESEARCH
published: 24 November 2021
doi: 10.3389/fpain.2021.725893

Stress and Pain Before, During and After the First Wave of the COVID-19 Pandemic: An Exploratory Longitudinal Mixed Methods Study

M. Gabrielle Pagé[1,2,3]*, *Lise Dassieu*[3], *Élise Develay*[3], *Mathieu Roy*[4,5], *Étienne Vachon-Presseau*[5,6,7], *Sonia Lupien*[8,9] and *Pierre Rainville*[10,11]

[1] Department of Anesthesiology and Pain Medicine, Faculty of Medicine, Université de Montréal, Montreal, QC, Canada, [2] Department of Psychology, Faculty of Arts and Science, Université de Montréal, Montreal, QC, Canada, [3] Research Center of the Centre hospitalier de l'Université de Montréal (CRCHUM), Montreal, QC, Canada, [4] Department of Psychology, Faculty of Science, McGill University, Montreal, QC, Canada, [5] Alan Edwards Centre for Research on Pain, McGill University, Montreal, QC, Canada, [6] Faculty of Dentistry, McGill University, Montreal, QC, Canada, [7] Department of Anesthesia, Faculty of Medicine, McGill University, Montreal, QC, Canada, [8] Centre for Studies on Human Stress, Montreal Mental Health University Institute, Montreal, QC, Canada, [9] Department of Psychiatry, Faculty of Medicine, University of Montreal, Montreal, QC, Canada, [10] Centre de recherche de l'Institut universitaire de gériatrie de Montréal, CIUSSS Centre-sud-de l'île de Montréal, Montreal, QC, Canada, [11] Department of Stomatology, Université de Montréal, Montreal, QC, Canada

OPEN ACCESS

Edited by:
Kenneth Craig,
University of British Columbia, Canada

Reviewed by:
Yang Wang,
University of Maryland, Baltimore, United States
C. Meghan McMurtry,
University of Guelph, Canada

Delane Linkiewich contributed to the review of C. Meghan McMurtry

***Correspondence:**
M. Gabrielle Pagé
gabrielle.page@umontreal.ca

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *16 June 2021*
Accepted: *22 October 2021*
Published: *24 November 2021*

Citation:
Pagé MG, Dassieu L, Develay É, Roy M, Vachon-Presseau É, Lupien S and Rainville P (2021) Stress and Pain Before, During and After the First Wave of the COVID-19 Pandemic: An Exploratory Longitudinal Mixed Methods Study. Front. Pain Res. 2:725893. doi: 10.3389/fpain.2021.725893

Aims: This study explores the association between subjective feeling of stress and pain experience in the context of the COVID-19 pandemic with a focus on characteristics known to trigger a physiological stress response [sense of low control, threat to ego, unpredictability and novelty (STUN)].

Methods: This exploratory longitudinal convergent mixed methods design consisted of online questionnaires over three time points (before, during and after the 1st wave of the COVID-19 pandemic) ($N = 49$) and qualitative interviews ($N = 27$) during the 1st wave of the pandemic on distinct samples of individuals living with chronic pain (CP). Both types of data sources were mixed upon integration using joint display.

Results: Mean pain intensity scores remained stable across time points, while pain unpleasantness and pain interference scores significantly improved. Global impression of change scores measured during the first wave of the pandemic do not entirely concord with pain scores evolution. Two thirds of participants reported a global deterioration of their pain condition at the beginning of the pandemic. Stress and pain catastrophizing before the pandemic were associated with pain scores throughout the pandemic; while most specific measures of stress due to the novel, uncontrollable, unpredictable and threatening nature of the pandemic were not. Qualitative data demonstrated that the deterioration reported in pain status reflected additional dimensions, including spatial expansion of the painful area, reduced access to treatments and challenges in adapting pain management strategies.

Conclusions: Helping individuals to negotiate stressful aspects of the pandemic might help offset the negative impacts of stress on pain status in this context or other important life events.

Keywords: chronic pain (MeSH), stress, COVID-19, pandemic, mixed methods, control, unpredictability

INTRODUCTION

The SARS-CoV-2 was identified in January 2020 as the cause of the coronavirus disease 2019 (COVID-19). Since then, this pandemic has been associated with more than 3 million deaths and 235 million confirmed cases as of October 7th 2021, more than 20 months after the first case was detected (1). In the province of Quebec, Canada, almost 1,000 cases and 150 deaths due to COVID-19 were reported daily during the first wave, for a population of 8.1 million inhabitants. The province of Quebec enforced lockdown of schools, office buildings, sports installations, restaurants, shopping malls in addition to postponing most non-urgent medical appointments. Notably, reopening was announced and postponed several times, until the end of May. The Quebec context during that specific time offered a unique opportunity to study the interaction between stress and chronic pain.

These effects constitute potential sources of stress that might have a particularly devastating impact on individuals living with chronic pain (2). Pain might deteriorate during the COVID-19 lockdown because of the direct impact of stress on pain (3, 4), or through indirect effects such as unpredictable access to pain care and management facilities, increased social isolation, and poor sleep (2, 5–10).

A multitude of studies have documented the complex associations between stress and pain, varying from stress-induced analgesia to stress-induced hyperalgesia (11–14). Furthermore, stress has also been identified as an important factor that could increase risks of comorbid psychological distress such as depression in this population (15, 16). Not every individual react the same way to sources of stress however, and understanding how individual appraisal of the threat and challenges posed by a new stressor such as the pandemic, as well as identifying vulnerability and resilience factors can help better understand the experience of individuals and its impact on pain evolution and its management (17, 18).

To better understand stress reactions, it is necessary to understand what stress is and how it triggers a physiological response (body's response to the detection of a threat—i.e., secretion of cortisol, noradrenaline). Decades of research have shown that when individuals perceive being in a situation over which they have a sense of low control (S), that poses a social-evaluative threat (T), is unpredictable (U) and/or is novel (N)—[thereby referred to the STUN characteristics], this will activate the hypothalamic-pituitary-adrenal axis and produce a stress response (19–21). The STUN framework appears to be an interesting and comprehensive approach to understanding the associations between stress and chronic pain, especially considering that they are not traditionally explored together within a comprehensive framework (22). Being able to characterize individuals' pain, stress and psychological characteristics and understand how these factors change once they are simultaneously exposed to a world-wide outbreak presents a unique opportunity to further our understanding of how and for whom stress has a significant impact on pain and psychological distress.

OBJECTIVES

The overall study goal was to explore the evolution of pain experiences among individuals living with chronic pain before and during the first wave of the COVID-19 pandemic in the province of Quebec, Canada, and understand how individual appraisal of the threats and challenges posed by the pandemic, influence this evolution. The specific quantitative (Study 1), qualitative (Study 2) and mixed methods (MM) goals were as follow:

Study 1—Quantitative examination:

(1) Examine the evolution of pain intensity, unpleasantness and interference scores at baseline, during and after the first wave of the pandemic.
(2) Document individuals' perceived global impression of change in pain status, and psychological distress during and after the first wave of the COVID-19 pandemic.
(3) Identify pre-pandemic stress-related indices (STUN characteristics, global perceived stress and pain catastrophizing) associated with the evolution of pain and psychological distress (anxiety and depressive symptoms) across the first wave of the COVID-19 pandemic.

Study 2 – Qualitative inquiry:

(4) Explore the dynamic impact of stress on the pain experience during the first wave of the COVID-19 pandemic from the perspective of people with chronic pain; and

Mixed Methods Integration:

(5) Obtain a more comprehensive understanding of the relationship between stress and pain experience during the COVID-19 pandemic by exploring convergence and divergence of the quantitative and qualitative findings.

The purpose of mixed methods in this study was thus to provide complementary and more comprehensive views of the phenomena under study and to take into account the diversity of perspectives on the experience (23).

OVERALL STUDY DESIGN

This study adopted a longitudinal convergent design with triangulation in which quantitative (Study 1) and qualitative (Study 2) data were collected in parallel using different samples and integrated using previously described methods (24, 25). **Figure 1** shows the timeline of the QUAN and QUAL studies, overlapping with the progression of the COVID-19 pandemic.

The study was approved by the research ethics board of the *Center hospitalier de l' Université de Montréal* (18.368-YP) and written consent was obtained from study participants. Here we first present the methodology, results and brief discussion of Study 1 and Study 2 separately, and finally the methods

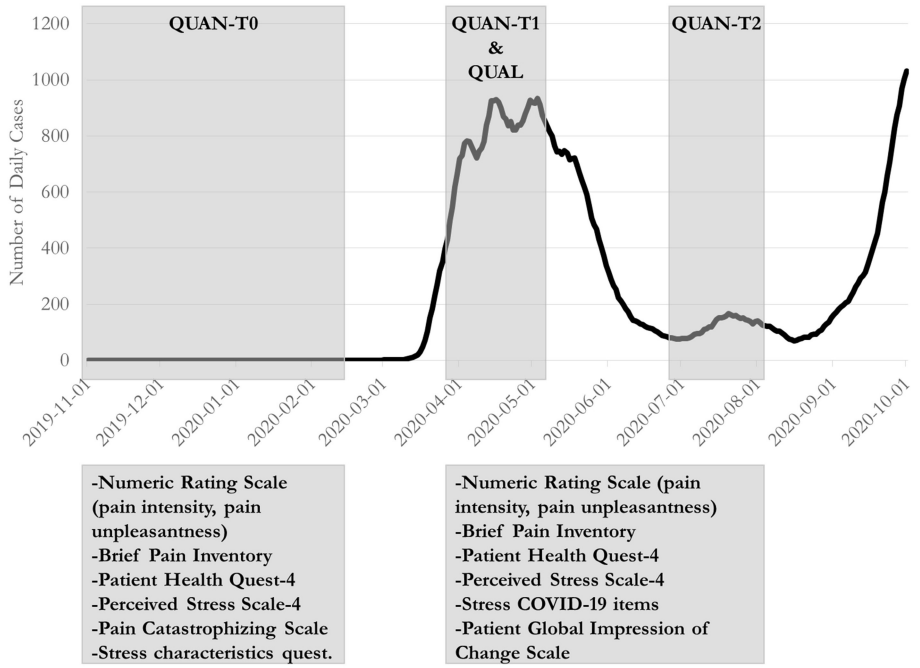


FIGURE 1 | Timing of baseline and follow-up measures in relation to the evolution of the COVID-19 pandemic in Quebec, Canada. The shaded gray zones refer to the periods during which participants completed the time point measures or interviews.

and results of the mixed methods integration. Samples were independent for Study 1 and Study 2.

STUDY 1—QUANTITATIVE STRESS AND PAIN INVESTIGATION

Materials and Methods

Study Design

The quantitative study adopted a longitudinal, prospective study design with three distinct time points: T0 (before the pandemic), T1 (during the first wave of the pandemic), and T2 (after the first wave of the pandemic).

Participants

Participants were initially recruited through an ad sent electronically in November 2019 to all members (approximately 9,000 individuals) of a community-based organization for individuals living with chronic pain for one of three studies exploring the associations between stress characteristics and chronic pain. Out of more than 600 individuals who manifested interest in the study, 54 were enrolled in this particular project until February 2020. At that time enrollment stopped because of the potentially confounding impact of the pandemic. When COVID-19 pandemic began and after obtaining ethics approval, participants who had already completed the study were solicited to participate in additional follow-up measurements to capture the impact of the pandemic on stress and pain. Eligibility criteria were assessed by phone and included having non-cancer pain of more than 3 months duration and of moderate to severe intensity (>3 on a 0–10 point scale), living in the province of Quebec, being fluent in written and spoken French, being aged 18 years or older, and having access to the Internet. Participants were excluded if they had a cognitive or physical impairment that made it impossible to complete self-reported questionnaires.

Procedures

Baseline (T0): After providing written consent electronically, participants completed an online battery of questionnaires documenting their overall stress, pain-related stress, pain characteristics and quality of life. They also completed a 1-week electronic diary that aimed to explore optimal methodological approaches to collect daily information on stress and pain, but the results are not presented as part of this study.

T1 and T2: During the first wave of the COVID-19 pandemic, all participants were re-contacted and invited to participate in a follow-up study to re-examine the associations between stress and pain during the pandemic. Forty-nine out of the 54 participants agreed to participate in these additional time points [during (T1) and after (T2) the first wave of the pandemic]. They followed the same procedure established for the baseline assessment to document stress related to the pandemic, overall stress, and pain characteristics using online questionnaires and electronic diary. Participants were compensated a total of $60 for the study.

Measures

Measures were selected to assess general and pain-specific stress and psychological responses to pain that might influence pain and psychological distress during the pandemic.

The following measures were administered across all three time points:

The *Brief Pain Inventory* [BPI (26)] is a measure of the impact of pain on daily function, pain location, pain medication, and amount of pain relief over 24-h period. Seven items, each rated on a 0–10 scale, document the extent to which pain impacts on daily function. This composite score had good reliability and validity in various chronic pain populations (27). In this study, $\alpha = 0.78–0.85$ at T0–T2.

The *Patient Health Questionnaire-4* [PHQ-4 (28)] is a brief measure of psychological distress with the following classification: normal (0–2), mild (3–5), moderate (6–8) and severe (9–12). Two items are drawn from the Patient Health Questionnaire-9 and evaluates depressive symptoms and two items are drawn from the Generalized Anxiety Disorder-7 scale and evaluates anxious symptoms. The PHQ-4 had good validity and adequate reliability (28). In this study, $\alpha = 0.70–0.78$ at T0–T2.

The *Perceived Stress Scale-4* [PSS4 (29)] is a 4-item self-reported measure that assesses the extent to which individuals perceive their life as being unpredictable, uncontrollable and overloaded over the previous month. The scale had excellent validity and internal consistency (29). In this study, $\alpha = 0.74–0.83$ at T0–T2.

The following measures were administered at baseline only:

The *Pain Catastrophizing Scale* [PCS (30)] measures the extent to which individuals ruminate, feel helpless, and magnify their pain experience. Each item is rated on a scale from 0 to 4, and items are summed to create a total score that ranges from 0 to 52. The PCS has been shown to have adequate internal consistency, reliability and sensitivity to change over time (31). In this study, $\alpha = 0.93$ at T0.

The *Stress Characteristics Questionnaire* [SCQ; (32)] measures one's sensitivity to each of the four characteristics associated with a physiological response to stress (19), namely Sense of low control, Threat to ego (one's personality), Unpredictability, and Novelty. Each dimension is measured by summing 5 Likert-type items that ask participants to rate on a scale from 0 (not stressful at all) to 10 (extremely stressful) the extent to which they would find each situation described as stressful. Higher scores indicate higher stress responsivity. In this study, $\alpha = 0.64$ for control subscale, $\alpha = 0.72$ for the ego subscale, $\alpha = 0.76$ for unpredictability subscale, and $\alpha = 0.80$ for the novelty subscale at T0. The psychometric properties of the original questionnaire have not yet been published. As such, the validity of the questionnaire is unknown.

Additional *pain characteristics* were measured, including pain duration. Pain intensity (mean, and worst pain intensity) and pain unpleasantness over the past 7 days were assessed using a Numeric Rating Scale (33, 34) (NRS, duration).

The following measures were administered at T1 and T2:

A series of questions on *Stress related to the COVID-19 pandemic* were administered on a 0–10 scale to assess the extent to which individuals found the pandemic to be stressful. Two questions aimed to measure overall stress related to the pandemic: "To what extent do you find the COVID-19 pandemic stressful," and "To what extent do you find the lockdown measures associated with the COVID-19 pandemic stressful". Four questions aimed to measure the four dimensions of the STUN model (Ego: "My behaviors and emotions about the COVID-19 pandemic have a negative impact on the opinion I have of myself;" Control: "The feeling of having no control over the evolution of the pandemic causes me stress," Novelty: The novelty of the current pandemic causes me stress;" Unpredictability: "The unpredictable evolution of the pandemic causes me stress"). These questions were developed by expert consensus (i.e., authors and other researchers with expertise in pain and/or stress research) at the beginning of the pandemic.

The *Patient Global Impression of Change scale* [PGIC (35)] was administered to document whether participants perceived a change in their pain status since the beginning of the pandemic on a scale ranging from 1 (completely deteriorated) to 7 (completely improved); a score between 1 and 3 indicates some deterioration; a score of 4 indicates no change and a score above 4 indicates some improvement. An open-ended question was also included that asked participants to describe how and why their pain status had changed. The PGIC has been recommended by the Initiative on Methods, Measurement and Pain Assessment in Clinical Trials (IMMPACT) group (33) and has been shown to mediate individual differences in a number of chronic pain outcomes associated with expectations (36).

Data Analysis

Descriptive statistics (frequencies, percentages, means and standard errors) were used to characterize the study sample and follow the evolution of pain and stress over time.

Objective 1 examined the evolution of pain and psychological distress across the first wave of the pandemic using linear mixed effect analysis. Models included a linear and a quadratic time trend, and a random effect of participant with a restricted maximum likelihood estimation used to determine whether scores on pain intensity (NRS pain intensity), pain unpleasantness (NRS pain unpleasantness), pain interference (BPI) or psychological distress (PHQ-4 total score) significantly changed across time. If quadratic term was significant it was retained in the model for Obj. 2, otherwise it was not included. Box plots were also used to compare changes in NRS pain intensity scores between T0 and T1, and between T1 and T2 according to participants' global impression of change in their pain status at T1 and T2, respectively.

Objective 2 examined characteristics associated with evolution of pain and psychological distress using linear mixed effect analysis. Models were used to identify baseline stress-related characteristics (pain catastrophizing, perceived stress (PSS-4) and scores on each of the four dimensions of the SCQ) associated with (a) pain intensity (NRS pain intensity), (b) pain interference (BPI), and (c) psychological distress (PHQ-4) across the first wave of the pandemic. Intercept was included as a random effect.

Alpha was set at 0.05. No correction was applied for multiple comparisons given that it further contributes to reducing statistical power, increases risks of Type II errors, and contributes to negative publication bias (37). Information regarding the clinical meaningfulness of statistically significant results is provided when relevant. Sensibility analyses were also conducted

TABLE 1 | Socio-demographic, pain, stress and psychological characteristics of individuals living with chronic pain before, during and after the first wave of the COVID-19 pandemic.

	Study time points relative to the 1st wave of the COVID-19 pandemic		
	Before ($n = 49$)	**During ($n = 48$)**	**After ($n = 46$)**
Socio-demographic characteristics			
Gender (*N* [%])		–	–
Woman	36 [75.0]		
Man	12 [25.0]		
Missing	1		
Age mean ± sd	51.13 ± 10.6	–	–
Min	30		
Max	78		
Living Environment (*N* [%])		–	–
Rural	9 [19.6]		
Urban	37 [80.4]		
Missing	3		
Race (*N* [%])			
White	44 [89.8%]		
Prefer not to answer/missing data	5 [10.2%]		
Education level (*N* [%])		–	–
High school	7 [14.3]		
Technical degree	26 [53.0]		
University	16 [32.7]		
Living condition (*N* [%])		–	–
Alone	12 [25.0]		
Family members	36 [75.0]		
Missing	1		
Work status (*N* [%])		–	–
Working	16 [28.6]		
Invalidity	26 [57.1]		
Retired	7 [14.3]		
Work status change (*N* [%])	–		
Same as pre-pandemic		5 [10.4]	10 [21.7]
Temporarily laid-off		5 [10.4]	3 [6.5]
Remote working		6 [12.5]	3 [6.5]
Not applicable		28 [58.4]	28 [60.9]
Missing		4 [8.3]	2 [4.4]
Psychological and stress characteristics			
Pain catastrophizing (PCS)	20.35 ± 11.8		
Stress characteristics based on the STUN framework (SCQ) (0-50)		–	–
Sense of low control	24.24 ± 10.4		
Threat to ego	27.08 ± 10.1		
Unpredictability	25.67 ± 10.6		
Novelty	31.10 ± 8.3		
Perceived stress scale (PSS-4)	7.41 ± 3.1	7.10 ± 2.4	7.48 ± 3.1
Psychological distress (PHQ-4)			
None-mild (0–5)	31 [63.3]	28 [58.3]	31 [67.4]
Moderate-severe (6–12)	18 [36.7]	20 [41.7]	15 [32.6]
Stress associated with COVID-19 pandemic (0–10)	–	7.16 ± 2.4	6.72 ± 2.4
Stress associated with lockdown measures (0–10)	–	5.86 ± 28.1	5.03 ± 2.6
Stress associated with (0–10):	–		
Sense of low control related to pandemic		5.67 ± 3.0	4.76 ± 2.9
Threat to the ego related to pandemic		2.48 ± 2.6	2.41 ± 2.8

(Continued)

TABLE 1 | Continued

	Study time points relative to the 1st wave of the COVID-19 pandemic		
	Before ($n = 49$)	**During ($n = 48$)**	**After ($n = 46$)**
Unpredictability of pandemic		6.48 ± 2.7	5.89 ± 2.5
Novelty of pandemic		5.69 ± 2.9	5.26 ± 2.6
Pain and health-related characteristics			
Pain duration (years)	15.11 ± 11.3	–	–
Mean Pain Intensity (NRS-11)	5.86 ± 1.4	6.08 ± 2.0	5.63 ± 1.8
Worst Pain Intensity (NRS-11)	8.16 ± 1.3	8.06 ± 1.5	7.70 ± 1.8
Pain Unpleasantness (NRS-11)	7.33 ± 1.8	6.42 ± 2.4	6.35 ± 2.2
Pain Interference (BPI)	5.90 ± 1.8	5.11 ± 2.1	5.09 ± 2.1
Global impression of change—pain status (*N* [%])	–		
Considerably deteriorated		2 [4.2]	3 [6.5]
Moderately deteriorated		9 [18.8]	4 [8.7]
Slightly deteriorated		21 [43.8]	15[32.6]
Unchanged		13 [27.1]	19[41.3]
Slightly improved		3 [6.3]	3[6.5]
Moderately improved		0 [0.0]	0[0.0]
Greatly improved		0 [0.0]	2[4.3]
Reason for pain deterioration (*N* [%])	–		
Increased stress		20 [62.5]	9 [40.9]
Delayed pain treatments		5 [15.6]	5 [22.7]
Other		6 [18.8]	6 [27.3]
Missing		1 [3.1]	2 [9.1]

The statistics are represented as mean ± sd unless otherwise specified.

PCS, Pain Catastrophizing Scale; SAM-S-P, Stress Appraisal Measure-Stressfulness subscale applied to pain; PSS-4, Perceived Stress Scale-short version; SCQ, Stress Characteristics Questionnaire; NRS-11, 0-10 Numeric Rating Scale; BPI, Brief Pain Inventory; QofL, Quality of Life.

to examine the unique contribution of the SCQ variables alone. The linear mixed effect models were thus re-run to exclude the PCS and PSS.

Results

Out of 54 individuals initially recruited, 49 completed at least one of the follow-ups and thus were included in the quantitative analyses. Participants' characteristics are shown in **Table 1**. Three quarters of participants identified as female ($n = 36$; 75.0%). More than half of participants ($n = 26$; 57.1%) were not working due to disability and the average pain duration was 15.11 years (sd = 11.3; min = 2, max = 43). Average pain intensity scores varied by <10% across the three time points, which is considered to be below what is considered as clinically meaningful (38).

Obj 1. Evolution of Pain and Psychological Distress. There were no significant linear or quadratic effect of time for pain intensity or psychological distress across the first wave of the pandemic ($p > 0.05$). There were significant linear ($\beta = -1.93$, $p = 0.010$) and quadratic ($\beta = 0.37$, $p = 0.042$) effects of time on levels of pain interference (BPI). Finally, there was a significant linear effect of time for pain unpleasantness ($\beta = -2.17$, $p = 0.036$). Results of the linear mixed effects models are shown in **Table 2**.

As shown in **Table 1**, 32 participants reported deteriorated pain during the first wave of the COVID-19 pandemic using the Patient Global Impression of Change Scale. Reasons reported by participants for this deterioration included stress ($n = 19/32$; 59.8%), postponed pain treatments ($n = 5/32$; 15.6%), ergonomic issues associated with working from home ($n = 2/32$; 6.3%), and sleep difficulties ($n = 2/32$; 6.3%). Thirteen reported unchanged pain. Only three participants reported improved pain that they attributed to a slower pace during the pandemic (e.g., less scheduled activities, not having to commute to work). Twenty-two participants (out of 46; 47.8%) reported that their pain deteriorated after the first wave (T2) compared to during the first wave (T1) of the pandemic.

In **Figure 2**, boxplots are displayed that show the differences in participants' report of pain intensity scores at the different time points and their corresponding reports of pain status change based on participants' global impression of change in their pain status.

Obj 2. Baseline stress characteristics associated with evolution of pain and psychological distress. Results of the linear mixed effects models are shown in **Table 3**. Higher levels of pain catastrophizing ($\beta = 0.04$, $p = 0.028$) at baseline were associated with higher pain intensity levels throughout the pandemic. Higher levels of pain catastrophizing ($\beta = 0.05$, $p = 0.031$) and perceived stress ($\beta = 0.07$, $p = 0.048$), and lower degree of vulnerability to perceived social-evaluative threat ($\beta = -0.08$, $p = 0.032$) were associated with higher levels of pain unpleasantness throughout the pandemic. Higher levels of perceived stress ($\beta = 0.10$, $p = 0.012$) and lower degree of vulnerability to unpredictability ($\beta = -0.09$, $p = 0.037$) were

TABLE 2 | Linear mixed effects models examining the within-person evolution of pain and psychological distress ($N = 49$).

Fixed effects	β	SE	t	p
Pain intensity				
Intercept	5.02	0.63	7.99	<0.001
Time	1.16	0.68	1.70	0.091
Time^2	−0.33	0.17	−1.04	0.054
Pain unpleasantness				
Intercept	9.08	0.91	9.93	<0.001
Time	**−2.17**	**1.01**	**−2.15**	**0.036**
Time^2	0.41	0.25	1.65	0.105
Pain interference (BPI)				
Intercept	7.46	0.68	10.98	<0.001
Time	**−1.93**	**0.72**	**−2.67**	**0.010**
Time^2	**0.37**	**0.18**	**2.09**	**0.042**
Psychological distress (PHQ)				
Intercept	4.53	1.44	3.15	0.003
Time	0.62	1.61	0.39	0.700
Time^2	−0.18	0.40	−0.44	0.659

B, unstandardized regression coefficients; SE, standard error; BPI, Brief Pain Inventory; PHQ, Patient Health Questionnaire-4. Bold values indicate $p < 0.05$.

associated with higher levels of pain interference throughout the pandemic. Finally, higher levels of baseline perceived stress ($\beta = 0.14$, $p = 0.002$) and pain catastrophizing ($\beta = 0.09$, $p < 0.001$) were associated with higher levels of psychological distress throughout the pandemic.

Sensibility analyses did not show any significant effects of the individual STUN components (when examined in a model without the PCS and PSS or in models where SCQ subscales were examined individually with the PCS and PSS), all $p > 0.05$.

Discussion Study 1

This study has investigated the experience of pain and stress among individuals living with chronic pain during the COVID-19 pandemic. Pain intensity scores on the NRS for the overall sample varied by <10% throughout the pandemic and pain unpleasantness and pain interference scores have improved. However, two-thirds of individuals reported that their pain status deteriorated during its first wave using the PGIC scale. Many studies report a high degree of concordance in individuals' pain intensity ratings and global impressions of change (39). However, a study of patients recruited from multidisciplinary pain clinics showed an overall subjective deterioration in pain but failed to show a significant difference in pain intensity ratings before and during the pandemic (40). Such discrepancy between pain scores and global impression of change in pain experience likely reflects the multidimensional and complex nature of the pain experience that goes beyond its intensity.

Stress was identified by more than half of participants with deteriorated pain as an important contributor to changes in pain status during the pandemic. The present study showed that individuals' tendency to ruminate, feel helpless, and magnify their pain experience, and those with higher levels of perceived stress are more likely to report higher levels of pain and psychological distress throughout the pandemic compared to those reporting lower levels at baseline.

Study results also show a small protective effect of social-evaluative threat on pain unpleasantness. This might be because the pandemic had sheltered us from social interactions and indirectly decreased the likelihood of encountering events that pose a social-evaluative threat. As such, those individuals most vulnerable to this type of stress experienced the largest benefits on pain unpleasantness. In addition, study results showed a small protective effect of sensitivity to unpredictable events on levels of pain interference. It is possible that those vulnerable to unpredictable situation react to this vulnerability by being more proactive in their environment in an attempt to reduce as much as possible sources of uncertainties. This attitude might in turn lead to increased levels of engagement in daily activities and thus reducing pain interference.

Global and multifactorial measures of stress (PCS and PSS) seem to have a stronger impact on pain outcomes however, compared to individual components of the STUN model. Perhaps given the magnitude of the pandemic, a global measure that captures many dimensions of stress would capture more variance in pain outcomes compared to individual components of the STUN framework. Many scales are now available to measure stress specifically in the context of the pandemic, such as the COVID Stress Scales (41) and the COVID-19 Phobia Scale (42).

STUDY 2—QUALITATIVE EXPLORATION OF STRESS AND PAIN DURING THE PANDEMIC

Material and Methods

Design of the QUAL Study

Semi- structured one-on-one interviews were carried out between March and May 2020 to explore the associations between stress and chronic pain in a pandemic context among individuals living with chronic pain. These individuals were recruited among a sample of 41 individuals who had participated in a focus group about stress and pain in 2019 (22). Given the different objectives of these two phases and content of the interview guides, these data are not analyzed jointly and here we focus only on the semi-structured interview data.

Participants

Out of 41 eligible individuals, 32 participants (16 women and 16 men) were randomly contacted by phone to inform them of the project until optimal sample size was achieved. Twenty-seven participants agreed to take part in an online interview and provided written consent electronically. Those participants were 18 years of age or older and living in the province of Quebec, fluent in spoken French, and living with chronic pain (>3 months) of moderate to severe intensity (>3/10). Final sample size was determined based on a number of factors, including timeline (interviews had to be conducted over the shortest time period possible in order to have the most homogeneous public health restrictions in place when participants were interviewed) (43) and methodological considerations for thematic analysis, including data saturation

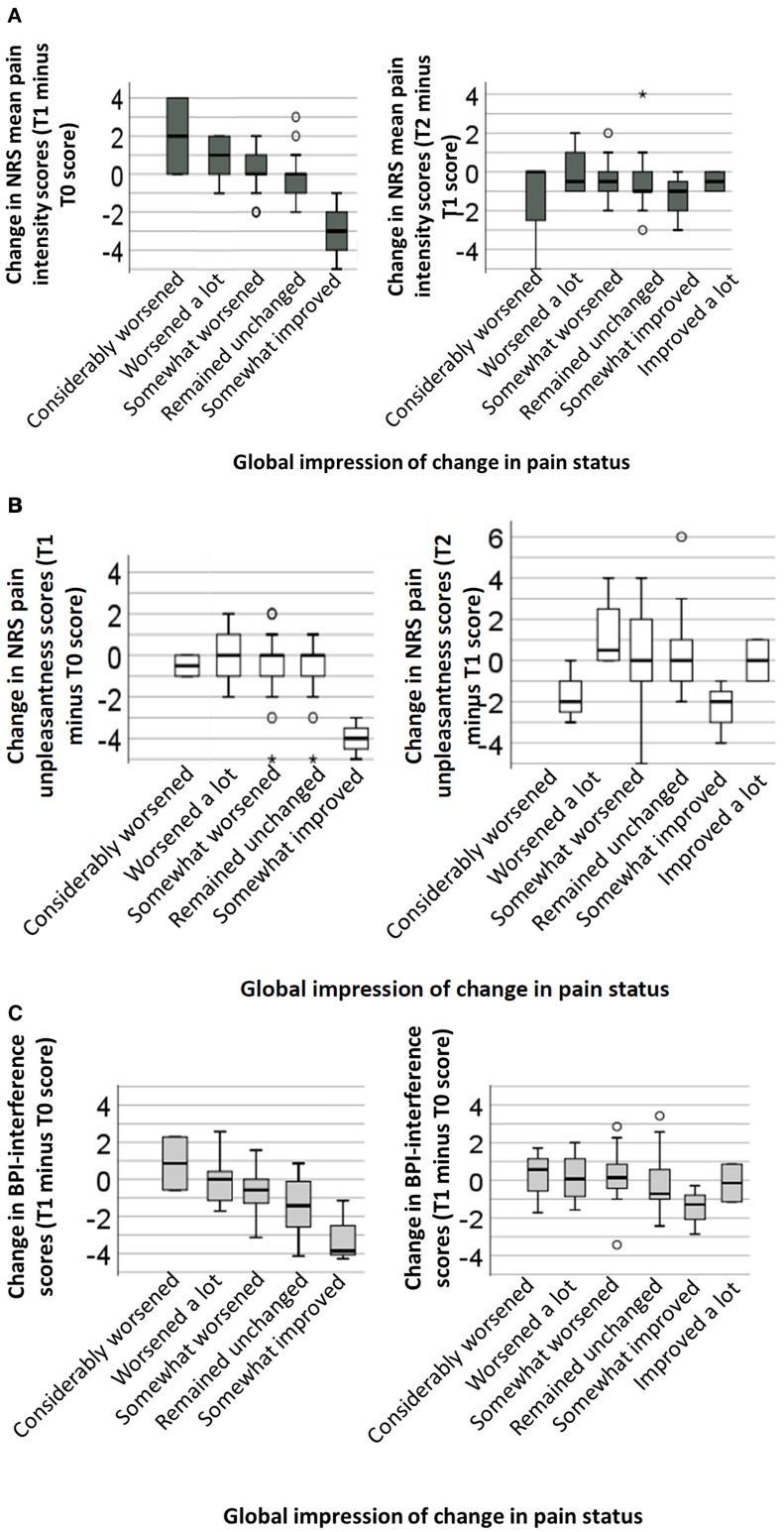


FIGURE 2 | Box plots of pain changes between the first wave of the pandemic (T1) and pre-pandemic pain levels (left graph) and between the end of the first wave (T2) and during the first wave (T1) of the pandemic (right graph). Row **(A)** represents changes in pain intensity scores, row **(B)** represents changes in pain unpleasantness scores, and row **(C)** represents changes in pain interference scores. Each box represents the first (Q1) and third (Q3) quartile and the middle line represent the median. The whiskers represent the minimum and maximum (Q1 or Q3–1.5*interquartile range) of the score distribution, with circles representing outliers. A score above zero on the y-axis indicates an increase in pain/interference scores (i.e., pain deterioration) from baseline to T1 (left graph) or from T1 to T2 (right graph), while a score below zero on the y-axis represents a decrease in pain/interference (i.e., pain relief). The x-axis represents individuals' global impression of change in pain status.

TABLE 3 | Linear mixed effects models examining the within-person evolution of pain and psychological distress taking into account baseline stress characteristics ($N = 49$).

Fixed effects	β	SE	t	p
Pain intensity				
Intercept	4.50	0.88	5.07	<0.001
Time	−0.10	0.12	−0.90	0.370
SCQ.control	−0.004	0.04	−0.11	0.913
SCQ.unpred	−0.02	0.03	−0.62	0.544
SCQ.ego	−0.44	0.03	−1.53	0.135
SCQ.new	0.59	0.03	1.98	0.056
PSS	0.58	0.03	2.00	0.053
PCS	**0.04**	**0.02**	**2.30**	**0.028**
Pain unpleasantness				
Intercept	6.48	1.13	5.74	<0.001
Time	**−0.56**	**0.15**	**−3.86**	**<0.001**
SCQ.control	0.003	0.05	0.07	0.942
SCQ.unpred	0.003	0.04	0.08	0.941
SCQ.ego	**−0.08**	**0.04**	**−2.21**	**0.032**
SCQ.new	0.04	0.04	0.96	0.342
PSS	**0.07**	**0.04**	**2.04**	**0.048**
PCS	**0.05**	**0.02**	**2.23**	**0.031**
Pain interference (BPI)				
Intercept	4.10	1.30	3.16	0.002
Time	−1.91	0.76	−2.54	0.015
Time^2	0.37	0.19	1.92	0.060
SCQ.control	0.08	0.05	1.70	0.097
SCQ.unpred	**−0.09**	**0.04**	**−2.15**	**0.037**
SCQ.ego	0.02	0.04	0.45	0.655
SCQ.new	<0.001	0.04	−0.002	0.998
PSS	**0.10**	**0.04**	**2.62**	**0.012**
PCS	0.04	0.02	1.76	0.086
Psychological distress (PHQ)				
Intercept	0.11	1.30	0.09	0.933
Time	−0.08	0.16	−0.52	0.603
SCQ.control	−0.01	0.05	−0.15	0.882
SCQ.unpred	−0.03	0.05	−0.51	0.611
SCQ.ego	0.003	0.04	0.08	0.938
SCQ.new	0.04	0.04	0.95	0.348
PSS	**0.14**	**0.04**	**3.39**	**0.002**
PCS	**0.09**	**0.02**	**3.82**	**<0.001**

B, unstandardized regression coefficients; SE, standard error; BPI, Brief Pain Inventory; PHQ, Patient Health Questionnaire-4; SCQ.control, control subscale of the Stress Characteristics Questionnaire; SCQ.unpred, unpredictability subscale of the Stress Characteristics Questionnaire; SCQ.ego, threat to the ego subscale of the Stress Characteristics Questionnaire; SCQ.new, novelty subscale of the Stress Characteristics Questionnaire; PSS, Perceived Stress Scale-4; PCS, Pain Catastrophizing Scale. Bold values indicate $p < 0.05$.

and thematic prevalence (44), narrow study aim, moderate sample specificity, and case analysis strategy (45).

Characteristics of participants involved in this qualitative part of the study are shown in **Table 4**. Information on participants' ethno-racial background was not collected in this study. None of the participants had been diagnosed with COVID-19 but 3 reported symptoms at the time of the interview.

TABLE 4 | Participant characteristics of the qualitative study ($N = 27$).

Variables	*N* (%)
Sex	
Males	12 (44.4%)
Females	15 (55.6%)
Age range	
<40 years	4 (14.8%)
40–69 years	18 (66.7%)
>70 years	5 (18.5)
Education level	
High school or less	0 (0%)
College or technical degree	12 (44.4%)
University	15 (55.6%)
Exposure to the COVID-19	
Diagnosed with the COVID-19	0 (0%)
Currently presenting symptoms of the COVID-19	3 (11.1%)
Been in contact with someone diagnosed with the COVID-19	1 (3.7%)
Pain duration	
0–2 years	1 (3.7%)
3–5 years	3 (11.1%)
6–10 years	4 (14.8%)
11–20 years	7 (25.9%)
21–30 years	8 (29.6%)
>30 years	4 (14.8%)
0–10 pain intensity (original study—2019)	
4-6	19 (70.4%)
≥7	8 (29.6%)
0–10 pain intensity (phase 2—2020)	
0–3	5 (18.5%)
4–6	13 (48.1%)
≥7	9 (33.4%)
Public health safety measures that directly impacted participants	
Dependent children at home	5 (18.5%)
Remote work	3 (11.1%)
Temporary loss of employment	3 (11.1%)
Canceled medical appointments	18 (66.7%)
Decreased medical assistance	9 (33.3%)
Reduction in assistance received from relatives	9 (33.3%)
Restrictions on leaving home (e.g., >70 years old, immunocompromised)	12 (44.4%)
Voluntary 14-day confinement	13 (48.1%)

Procedure

Participants completed a sociodemographic questionnaire online prior to engaging in an individual interview online via the platform Zoom that lasted between 30 and 80 min. Interviews were conducted using a semi-structured guide. Interview topics included overall stress experience in the context of the pandemic, the impact of stress related to the pandemic on their pain condition and its treatment and management and coping with stress and pain during the pandemic. Conversations were audio-recorded and transcribed verbatim.

Interviews were conducted by one of two interviewers (MP or ÉD). MP is a female clinical psychologist and pain researcher

trained in qualitative and mixed methods. ÉD is a female sociologist trained in qualitative research. Participants were informed about the study goals, i.e., to revisit the relationship between stress and pain but this time in the context of the COVID-19 pandemic. All interviews were conducted in French and data analysis was also conducted in that language in line with recommendations for qualitative analysis and result dissemination in a different language than the one of data collection (46). Final themes and selected quotes were translated into English by a professional translator.

Data Analysis

Reflexive thematic analysis was used as the primary data analysis method, using patterns of shared meaning (47, 48). An inductive approach was mainly used to explore specifically characteristics of stress and pain present in the data. Contextualization of these characteristics within the broader lived experience of participants in the context of the COVID-19 pandemic was then explored. While the STUN framework helped to interpret results from the analysis, it was not used to identify theme or classify types of stress experienced by participants. Attempts were made however, to evaluate whether the STUN characteristics are relevant to the experience of stress during the pandemic.

The lead analyst and two other team members established a preliminary and evolving codebook (49); frequent meetings were held to arrive at a common codebook. Process and open codes through a line-by-line analysis were used to move toward an interpretive level of analysis and the generation of themes (50). An iterative approach moving several times between raw data and ongoing interpretation and reflections on participants' experiences was used. Several team meetings took place to construct themes. Member checking and audit trails were used to enhance trustworthiness of the data (51). NVivo-12 (52) was used to code data into domain summaries.

Qualitative Results (QUAL)

Data analysis aimed to explore the dynamic impact of stress on the pain experience during the first wave of the COVID-19 pandemic. The experiences of participants were heterogeneous, with some reporting little to no impact of the pandemic on their stress, pain or daily routine, while others described feeling heavily the bidirectional effects of the pandemic and pain conditions. Five themes were identified: (1) status quo: between philosophy and stability of life and health stages; (2) pain management in socially exposed and disrupted environments; (3) further complicating access to pain care: adding insult to injury; (4) avoidance as a stress response to an invisible threat; and (5) silver lining: regaining control of pain during an uncontrollable pandemic.

Status Quo: Between Philosophy and Stability of Life and Health Stages

A few participants reported minimal disruptions to their daily life during the ongoing pandemic. These individuals described having well-established pain care plans that were not disrupted by the lockdown measures, or their work and social statuses were less likely to be disrupted because they were retired for example.

> "How is my stress… well yesterday we baked bread! (laughs) And so we don't have food problems anymore. And me and my spouse we have been married for 46 years, so we get along very well.… We talk, we do things together.… Things are going well. And on top of that, there are less people coming over for dinner! And we don't go to other people's places for diner! So there is less going back and forth and that suits me very well!" (P.13, M, 64 years old)

Beyond the stability of one's life and health stage, one's philosophy also helped minimize the impact of the pandemic on their pain and stress levels. Those individuals tended to focus on aspects that they could control, and on the present moment. By doing so they were able to reduce the perceived lack of control and unpredictability of the pandemic.

> "I often say that nowadays: I don't accept my pain, but I'm learning to live with it. It'll be the same thing with the pandemic. I don't accept the virus, but I've got no choice but to learn to live with it! And learn to live differently! It's the same as with my pain!" (P.11, M, 78 years old)

For others however, the absence of added stress from the pandemic came rather from the perspective that one's situation was already so poor that it could not further deteriorate.

> "The pandemic hasn't had any particular impact, because I was already all destroyed, or almost. This is normal as health problems like this one play out. It can go as far as social isolation! You can no longer have a social life with people in the same way!" (P.20, M, 58 years old)

Pain Management in Socially Exposed and Disrupted Environments

Managing pain during a pandemic was difficult for many participants who struggled with the increased cognitive and physical workload brought on by the pandemic in their personal and/or professional lives. This translated for some into increased difficulties to apply pain management strategies.

> "[My pain] has completely increased… It is hard to manage right now. We have so many other things to manage, other things to think about. As stress increases, pain management becomes more difficult." (P.6, F, 34 years old)

Furthermore, the altered social environment, such as all family members suddenly staying home, exposed people's pain to broad daylight, making it much more challenging to hide it. This confinement decreased their ability to manage pain and hide it from others, thus threatening their ego.

> "Then, on top of it, they see the pain I've got. Normally they don't see it so much. They'd see it in the evening, but now when you're with somebody all day long and then the person… you see that they're in pain all the time. [...] It's like showing your family another part of the pain, so it's harder." (P.6, F, 34 years old)

The constant tension between needing to engage in self-care or pain management and caring for the needs of others (e.g., having to care for children who are at home) was highly stressful and led

to a vicious cycle of increased pain that then fed their stress. This was particularly discussed by younger participants. At times, this vicious cycle led to an under-utilization of non-pharmacological means to manage pain and increased pain medication intake.

> "Before, since I was all alone during the day, I'd take a nap and often in the afternoon I'd feel better… I was able to do things to reduce the pain. And now, since the kids are here all the time and they're asking for things, they keep asking, asking!... So, I just can't manage my pain like I used to. Now I'm managing my kids. So, my pain, I manage it more with meds now." (P.12, F, 39 years old)

Further Complicating Access to Pain Care: Adding Insult to Injury

Many participants feared that unstable access to pain management because of postponed or canceled appointments would lead (or had already led) to significant deterioration of their pain condition. This was also the case for physical and psychological pain management strategies. In this case, the threat was not only a worsening of their pain condition, but also of their social and psychological well-being.

> "When they closed the gyms that was a big deal for me. Because it's the only physical exercise I can do. It's good for the... it allows circulation in what I've got left. I'm afraid I'm going to lose this arm [the right arm] at some point. This arm [the left arm] is starting to get cold. I'm afraid that algodystrophy will get into it. It's a bit of a drag. Going to the gym has psychological benefits... It's like you took away my pub, by doing that. They've taken away my social club. By closing the gym, they took a lot away from me. It's a big deal." (P.18, M 57 years old)

During the pandemic, there has been new solutions, such as online care and activities, to provide social interactions and meet the self-care needs of the general population. However, for those who needed health care services, these solutions seemed to increase one's frustration and stress.

> "We're told: be creative! Do some meditation at home. Do some painting, and all the rest. It's all very well to do some painting... Instead of just making life livable for that person, to say: to forget your misery, you can… make paper-maché sculptures! Well I don't want paper-maché, I want massage therapy. They make me do paper-maché to help me forget that I don't have any services, that my life is just poop. But (laughs) at one point, it's NO! That's enough!" (P.27, F, 48 years old)

Avoidance as a Stress Response to an Invisible Threat

The virus posed an invisible threat for many individuals who perceived themselves as being at higher risks of dying should they get infected. Many individuals perceived their health as fragile in part because of chronic pain.

> "I already have so much difficulties trying to be the woman I used to be. I will never be that woman again. But if I catch [the virus] on top of it, I don't think I'll be able to get through this. Just sometimes I cough let's say because I have a dry throat. It pulls, it really hurts, I am writhing in pain. If I should catch something like this virus, I won't survive." (P.6, F, 34 years old)

The novelty of the virus and lack of knowledge about modes of transmissions, the unpredictability of one's chance of surviving if they get infected, and the lack of control over the situation were important sources of stress. This perceived threat had a significant impact on their behaviors, including increased hesitation at seeking medical care for pain.

> "I would not want to end up in the hospital… We don't want pain to increase, but we don't want to end up in the health care system for COVID-19." (P.21, M, 59 years old)

Silver Lining: Regaining Control of Pain During an Uncontrollable Pandemic

The stability of one's pain condition and ongoing treatment prior to the COVID-19 pandemic had a large impact on an individuals' stress appraisal of the pandemic. For some, the pandemic had positive impacts on their pain management opportunities by providing them with more time to devote to their pain care.

> "And I take advantage of it because I'm slower in my personal activities. So, I take at least two breaks each day, for my treatments [TENS]. And I can only have this treatment when the pain isn't too intense, because when it is, I can't take these electric shocks." (P3, M, 78 years old)

This was also the case for those who perceived the pandemic as a break from having to push the limits of their physical capacities in order to meet their basic needs and as a temporary protection against the threat of pain on one's sense of identity.

> "It's less confrontational not to have to do something than it is to have to do it and tell you, well, right now I'm dragging a 40-pound load, the stairs in the subway are out of order, I have two huge landings, and no one's stopping to help me… So now it's more like "don't take public transit!" So now what you're doing to me, is that I don't have to suffer, and I don't have to be humiliated? Cool!" (P27, F, 48 years old)

Discussion Study 2

This study explored the dynamic impact of stress on the pain experience during the first wave of the COVID-19 pandemic. Stress and pain responses to the pandemic were heterogeneous and seemed influenced by many factors, such as one's life stage and social situation, pain condition and stability of ongoing treatments, degree of precariousness, and level of adaptability. While many reported negative impacts of the pandemic on their pain and overall well-being, others perceived opportunities to further adapt their pain management strategies or focus on elements that were within their control to minimize stress.

As mentioned previously, the STUN framework identifies four specific characteristics of situations that will trigger a physiological stress response: novelty, unpredictability, threat to the ego and sense of low control (19–21). Many of these characteristics could be observed in participants' narratives.

Those perceiving having little control over an unpredictable pandemic, pain condition, and access to pain care often felt overwhelmed, stressed and with increased pain levels. Few participants described that focusing on aspects over which they had control, such as strictly following public health safety recommendations, served as a buffer against potentially stressful situations. Threat to the ego was also discussed in the context of increased in-home social interactions, making it more difficult to hide their pain. For others however, this social threat of pain was decreased because they either lived alone or had significantly less interactions with the outer world. Among the four characteristics of the STUN framework, the novelty was the least discussed. This might be in part because chronic pain requires one to constantly navigate new challenges and as such the pandemic wasn't such a departure from their constant need for adaptation. This could also be in part because this pandemic was a novel experience for most of the world population and as such wasn't discussed specifically in the interviews focused on living with chronic pain.

MIXED METHODS INTEGRATION

Considering the quantitative findings, the mixed methods objective was to better understand how individuals' pain condition evolved throughout the first wave of the pandemic, and the extent to which stress played a role in this evolution. The integration process was particularly focused on obtaining a deeper understanding of the incongruent pain status and pain intensity reports in the quantitative finding, and how stress and other dimensions of pain could help better understand individual experiences. This was done by merging to two databases for analysis.

Methodology for Mixed Methods Integration

The quantitative data was given more weight in the integration, and qualitative data was mainly used to elucidate puzzling quantitative findings regarding discrepancies between scores on the global impression of change in pain status and pain intensity. The databases were examined contiguously, and then together to compare and contrast findings using primarily joint displays (53–55). Codes and themes obtained from the qualitative analysis were examined in the context of the quantitative results. More specifically, three frequently occurring types of quantitative profiles characterizing changes associated with the pandemic were identified in the data using a cross-tabulation: global impression of change worsened and pain increased (NRS score); global impression of change worsened and pain decreased; and global impression of change unchanged/improved and pain decreased. Then, participants of the qualitative study were also categorized based on change in their NRS pain score obtained before each interview as increased, unchanged or decreased pain intensity score before and during the first wave of the pandemic. Their qualitative data was then coded to capture their subjective impression on the evolution of their pain during the pandemic (e.g., how participants described their pain evolution since the beginning of the pandemic). This allowed to explore whether the quantitative profiles found in Study 1 were also present in Study 2 and whether other profiles could also be identified. This structure allows for exploration of conceptual similarities between the variables and themes and how they interact (55). This data integration was presented in the form of joint displays, which are visual integration of quantitative and qualitative findings that aim to generate new insights.

Integration of QUAN and QUAL Findings

Given that participants in the qualitative study were recruited from an earlier focus group study, their individual pain scores pre-pandemic were available and examined (see **Table 4**). Specific profiles of narratives offering a deeper understanding of the quantitative pain intensity and global pain status ratings are shown in the joint display in **Figure 3**. This display highlights the presence of three distinct profiles of individuals who participated in the quantitative study. The profiles were derived using a cross-tabulation using the change in NRS pain score from the first wave of the pandemic compared to their pre-pandemic score, and their global impression of change score related to their pain status since the beginning of the pandemic. The first profile represents individuals who report a worsened pain status on the PGIC and report an increased pain on the NRS-11 from pre- to during the pandemic. The second profile represents individuals who also report a worsened pain status on the PGIC but report a decreased pain on the NRS-11 from pre- to during the pandemic. Last, the third profile represents individuals who report an unchanged or improved pain status on the PGIC while reporting a decreased pain on the NRS-11 from pre- to during the pandemic. No individuals reported an unchanged/improved pain one the PIGC while reporting a deteriorated pain on the NRS-11 from pre- to during the pandemic. Participants in the qualitative study were also categorized based on their NRS-11 scores as having deteriorated or improved/unchanged. Their narrative were analyzed to understand their impression of pain evolution, in order to further our understanding of the quantitative profiles.

Increased Pain Intensity Ratings and Global Impression of Pain Deterioration

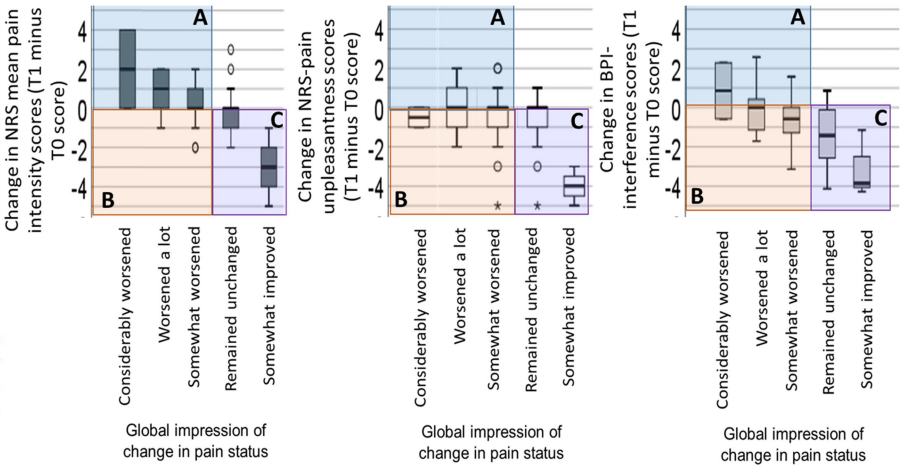
As shown in quadrant A in **Figure 3**, many individuals reported coherent ratings of pain deterioration in the measure of global impression of change combined with increased pain scores, reflecting struggles to adapt to the pandemic in terms of both stress and pain. The intensity of their pain itself increased, and this is to be understood in the context of a global deterioration of their physical condition, well-being, and often social environments.

Decreased Pain Intensity Ratings and Global Impression of Pain Deterioration

As shown in quadrant B in **Figure 3**, a few participants, despite reporting a decreased pain intensity score from baseline to the first wave of the pandemic, reported that their pain condition had deteriorated. Pain being a biopsychosocial experience, degradation of the psychological and social components of pain negatively impacted individuals' perception of their overall pain

"I'm not proud to say it, but I feel like I deal with pain more poorly, I lose a lot. So I feel that this situation makes me weaker psychologically, so when I have to manage my pain… Before I had tons of strategies and it was going well, I had my routine. When I was hurting I would do this and that. But now I'd say that I'm so stuck in all of [the pandemic] that I forget my strategies and I cry." Pain intensity: pre-covid: 5/10; during covid: 8/10 (P.17, F, 54. y.o)

"Of course there's an added layer to the pain. Not being able to have control over the [pandemic], it triggers pains that are not related to the disease, but that are associated with stress. For example, my shoulders are tight, I have a bit more pain. I already have normal neck pain, but it seems like there is more pilling up in there. And I'm often tired because I struggle more to manage this aspect of the pain. It's like an added layer of pain that is not there normally but that I can control to some extent. I take more medications." Pain intensity: pre-covid: 7; during covid: 4.5 (P.22, M, 61 y. o.)



"I try to stay calm and it reassures people around me. And it makes me feel useful… I take advantage of this because during that time, the chronic pain that I have, it's almost like if I forget that it`s there. I have my own slogan that I repeat all of the time, I have to keep occupied so I don't get preoccupied… I'm not saying the pain is gone! I've been living with pain for 21 years. I've been on Fentanyl patches for 18 years to be able to tolerate the pain. Of course the virus has turned our lives upside down, but there are people worse off than me, that are stuck in hospital beds, that are dying. So thinking about it that way helps me to keep going." Pain intensity: pre-covid: 8/10; during covid: 5/10 (P.11, M, 78 y. o.)

FIGURE 3 | Quantitative reports of pain intensity (left graph), pain unpleasantness (middle graph) and pain interference (right graph) scores and global impression in pain status from the quantitative study and citations of qualitative study participants. For each graph, the x-axis represents individuals' global impression of change in pain status. The y-axis represents the difference in pain intensity, unpleasantness or interference scores from T1 minus T0. **(A)** Pain worsened consistent with a deterioration in the global impression of change in pain status. This was the case for 26 individuals for pain intensity, 18 individuals for pain unpleasantness, and 12 individuals for pain interference. **(B)** Pain improved but the global impression of change in pain status suggests a general deterioration reflecting other pain/stress-related factors. This was the case for 6 individuals for pain intensity, 14 individuals for pain unpleasantness, and 20 individuals for pain interference. **(C)** Pain improved and global impression of change in pain status suggests stability or some improvement. This was the case for 7 individuals for pain intensity, 9 individuals for pain unpleasantness, and 13 individuals for pain interference. The qualitative study allowed to identify similar profiles of participants that provide context to those ratings.

condition. For example, the fact that others were less emotionally available, having less access to a social network, increased level of suffering, and uncertainty about the resolution of the turmoil produced by the pandemic were discussed as being embedded in their overall pain experience.

Decreased Pain Intensity Ratings and Unchanged/Improved Global Impression of Change

As shown in quadrant C in **Figure 3**, approximately one-third of individuals in the quantitative study reported no change or in few instances an improvement of their pain condition and reported decreased pain intensity ratings during the first wave of the pandemic compared to baseline. These individuals tended to have relatively stable life circumstances (e.g., not being in the workforce), living with a partner, having a stable source of income, and having a well-established pain care plan unaffected by the lockdown measures. Adopting an empathic stance toward those affected more directly by the pandemic helped decrease the social threat posed by such novel event and turned the focus away from pain.

DISCUSSION

This research has investigated the experience of pain and stress among individuals living with chronic pain during the COVID-19 pandemic using a mixed methods approach. Several key findings emerged from this research.

Levels of Stress and Pain During the Pandemic

A significant proportion of individuals in this study reported a deterioration of their pain condition during the first wave of the pandemic, in agreement with another Canadian study (10). Results of linear models exploring pain scores over time showed however no change, or slight improvement in the case of pain unpleasantness and pain interference over time. This would be consistent with studies from the United States and Europe which found that most participants reported unchanged pain severity at the beginning of the pandemic (56, 57). One has to be careful when comparing data across countries, given differences in the local state of the pandemic during data collection, strictness of lockdown measures in place, and extent of disruption of the health care system. Nonetheless all studies identified subgroups of individuals who faired relatively well during the pandemic.

As documented in another qualitative Canadian study, some individuals perceived an improved quality of life during the pandemic, either because the world had slowed down to a pace that is more compatible with their level of functioning (e.g., decreased requests for social outings) or because they had more time to focus on pain management (58). It is also possible that some individuals with chronic pain have developed a resilience to overcome challenges and obstacles and a flexibility to engage in new or alternative pain management strategies. Psychological flexibility has been identify in another study as an important contributor to individuals with chronic pain' psychological well-being and pain interference during the pandemic (59). The concept of resilience as facilitating adaptation to chronic pain has also been documented in other contexts (60).

Many studies have documented that stress is common (61–65). In the pain literature, while levels of stress are also generally more elevated during the pandemic, there is controversy regarding whether this increased stress leads to worsened pain (66, 67). Not all were equal in the face of stress, however. In the present study, those whose daily routine were disrupted or were in more precarious socioeconomic situations were more likely to face multiple stressors due to the pandemic. This is consistent with a European study that found that levels of economic vulnerability increased one's risk of experiencing anxiety, depression, and stress during the lockdown to control the spread of COVID-19 (68).

Multidimensional Impact of Stress on Pain

Sources of stress were numerous and diverse during the pandemic. For individuals living with chronic pain, this included environmental stressors of the pandemic itself (exposure to the virus, lockdown measures), but also pain-specific stressors (e.g., postponement of medical appointments, decreased help from others) (59, 69). Given individuals' vulnerabilities to stress, social context and pain condition, the impact of the pandemic on their pain journey was heterogeneous (58, 70, 71). Pain appeared to be affected by stress in multiple ways, including overwhelming cognitive load that made it more difficult to engage in pain management, decreased social contexts conducive to pain management, anxiety, fatigue and apathy that decreases one's ability to cope with pain (57). Given the observed heterogeneity in participants' contextual factors (e.g., stability of pain treatments, socioeconomic status, social support), the association between stress during the pandemic and pain outcomes remains complex and multifactorial.

Global Perceived Stress vs. Individual Components of the STUN Framework

Quantitative and qualitative studies identified individual components of the STUN framework associated with participants' experience of the pandemic and its impact on pain. For example, lack of control over and unpredictability of the pandemic and pain dynamics led some participants overwhelmed and feeling vulnerable to the escalation of both stress and pain. For others, focusing on controllable aspects of their day-to-day life seemed to decrease their levels of stress. Perceived control over time was identified in one study as an important factor associated with anxiety and fear of COVID-19 pandemic (72). The unpredictable evolution of the pandemic and its overall stress load were also identified as important determinants of burnout syndromes in different populations, such as healthcare workers (73). Quantitative results, however, suggest that general measures of stress and pain catastrophizing before the pandemic are associated with pain dimensions during the first wave of the pandemic, more so that one's vulnerability to individual components of the STUN model. This might be because the pandemic at is onset disrupts so many aspects of individuals' lives, including work, social relations, health care behaviors, and survival that not one single component will capture all these facets. As individuals learn to live in a pandemic and as specific pandemic-related issues emerge (e.g., polarization of opinions on confinement measures or vaccines), specific characteristics of the STUN model (e.g., social-evaluative threat) would have a larger influence on individuals' experiences.

STRENGTHS AND LIMITATIONS

One important strength of this study is the capture of stress and pain data at baseline, before the beginning of the pandemic. This provided a rare opportunity to explore how stress, pain and their associations evolved during and after a natural world-wide stress exposure. The use of mixed methods also added value to both quantitative and qualitative findings and provided new insights that would not have been possible without data integration. Nonetheless this study also has some limitations. The sample size of the quantitative component is relatively small, but it was not possible to increase sample size once the state of emergency had been declared in Canada due to its influence on baseline stress data. As a result, the number of independent variables examined was limited. The small sample size might have also introduced a selection bias, and limits generalizability of study findings to different chronic pain populations. Participants were recruited from a single province, namely the one reporting the highest number of COVID-19 cases during the first wave of the pandemic. As such, results might not be generalizable to individuals from other provinces or other countries. In addition, the level of education, particularly in the qualitative sample was high and might not reflect the situation of many individuals living with chronic pain. Also, some study questions, such as the Stress Characteristics Questionnaire, do not have published data on their psychometric properties and as a result their validity and reliability have not yet been demonstrated. Finally, both samples had socioeconomic diversity but lacked in ethnic diversity with participants being predominantly White in study 1 and this information was not captured in study 2.

CONCLUSIONS

Multiple sources of stress associated with the COVID-19 pandemic were identified among individuals with chronic pain. While some participants reported little impact of the pandemic on their stress and pain status, most identified significant difficulties in managing pain and stress in this context. For future COVID-19 waves and pandemics, it will be crucial to develop

interventions (e.g., individual and/or family programs aimed at optimizing well-being, stress and pain management in the context of shifted routines and roles) and community support (e.g., programs adapted to the specific challenges faced during the pandemic) that are tailored to the needs and physical capacities of individuals living with chronic pain.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author, A. Lacasse, upon reasonable request and conditionally to a proper ethical approval for a secondary data analysis. The data are not publicly available since participants did not initially provide consent to open data.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité d'éthique à la recherche du Center hospitalier de l'Université de Montréal. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MP, MR, PR, ÉV-P, and ÉD were involved in study design. MP and ÉD collected data. MP, ÉD, and LD were the primary authors involved in the qualitative analysis and mixed methods integration. Results were discussed with all study authors and subsequently refined. All authors contributed to the writing and/or revisions of the manuscripts and all approved its final version.

FUNDING

This study was funded by a stimulus grant awarded to MP from the Quebec Pain Research Network, a network funded by the Fonds de recherche du Québec-santé. MP and ÉV-P are Junior 1 research scholars from the *Fonds de recherche du Québec-santé*. MR holds a Canada Research Chair in Brain Imaging of Experimental and Chronic pain. SL holds a Canada Research Chair on Human Stress.

ACKNOWLEDGMENTS

We are thankful to all participants who took the time and share their stories.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpain.2021.725893/full#supplementary-material

REFERENCES

1. World Health Organization. Coronavirus Disease (COVID-19) pandemic 2020 [Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed October 7, 2021).
2. Nieto R, Pardo R, Sora B, Feliu-Soler A, Luciano JV. Impact of COVID-19 lockdown measures on spanish people with chronic pain: an online study survey. *J Clin Med*. (2020) 9:558. doi: 10.3390/jcm9113558
3. Vachon-Presseau E, Martel MO, Roy M, Caron E, Albouy G, Marin MF, et al. Acute stress contributes to individual differences in pain and pain-related brain activity in healthy and chronic pain patients. *J Neurosci.* (2013) 33:6826–33. doi: 10.1523/JNEUROSCI.4584-12.2013
4. Burke NN, Finn DP, McGuire BE, Roche M. Psychological stress in early life as a predisposing factor for the development of chronic pain: clinical and preclinical evidence and neurobiological mechanisms. *J Neurosci Res.* (2017) 95:1257–70. doi: 10.1002/jnr.23802
5. Aremka LM, Andridge RR, Fagundes CP, PAlfano CM, Povoski SP, Lipari AM, et al. Pain, depression and fatigue: loneliness as a longitudinal risk factor. *Health Psychol.* (2014) 33:948–57. doi: 10.1037/a0034012
6. Voitsidis P, Gliatas I, Bairachtari V, Papadopoulou K, Papageorgiou G, Parlapani E, et al. Insomnia during the COVID-19 pandemic in a Greek population. *Psychiatry Res.* (2020) 289:113076. doi: 10.1016/j.psychres.2020.113076
7. Smith TO, Dainty JR, Williamson E, Martin KR. Association between musculoskeletal pain with social isolation and loneliness: analysis of the English Longitudinal Study of Ageing. *Br J Pain.* (2018) 13:82–90. doi: 10.1177/2049463718802868
8. Dassieu L, Heino A, Develay E, Kabore JL, Page MG, Hudspith M, et al. Conversations about opioids: impact of the opioid overdose epidemic on social interactions for people who live with chronic pain. *Qual Health Res.* (2021) 31:1657–69. doi: 10.1177/10497323211003063
9. Lacasse A, Page MG, Dassieu L, Sourial N, Janelle-Montcalm A, Dorais M, et al. Impact of the COVID-19 pandemic on the pharmacological, physical, and psychological treatments of pain: findings from the Chronic Pain & COVID-19 Pan-Canadian Study. *Pain Rep.* (2021) 6:e891. doi: 10.1097/PR9.0000000000000891
10. Page MG, Lacasse A, Dassieu L, Hudspith M, Moor G, Sutton K, et al. A cross-sectional study of pain status and psychological distress among individuals living with chronic pain: the Chronic Pain & COVID-19 Pan-Canadian Study. *Health Promot Chronic Dis Prev Can*. (2021) 41:141–52. doi: 10.24095/hpcdp.41.5.01
11. Butler RK, Finn DP. Stress-induced analgesia. *Prog Neurobiol.* (2009) 88:184–202. doi: 10.1016/j.pneurobio.2009.04.003
12. Jennings EM, Okine BN, Roche M, Finn DP. Stress-induced hyperalgesia. *Prog neurobiology.* (2014) 121:1–18. doi: 10.1016/j.pneurobio.2014.06.003
13. Olango WM, Finn DP. Neurobiology of stress-induced hyperalgesia. *Curr Top Behav Neurosci.* (2014) 20:251–80. doi: 10.1007/7854_2014_302
14. Vachon-Presseau E. Effects of stress on the corticolimbic system: implications for chronic pain. *Prog Neuropsychopharmacol Biol Psychiatry.* (2017). doi: 10.1016/j.pnpbp.2017.10.014
15. Martin AL, Halket E, Asmundson GJ, Flora DB, Katz J. Posttraumatic stress symptoms and the Diathesis-Stress Model of Chronic Pain and Disability. *Clin J Pain.* (2009) 26:518–27. doi: 10.1097/AJP.0b013e3181e15b98
16. Turk DC. A diathesis-stress model of chronic pain and disability following traumatic injury. *Pain Res Manag.* (2002) 7:9–19. doi: 10.1155/2002/252904
17. Lazarus RS. *Psychological Stress and the Coping Process*. New York, NY: McGraw-Hill. (1966).
18. Lazarus RS, Folkman S. *Stress, Appraisal and Coping*. New York, NY: Springer Press. (1984).
19. Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull.* (2004) 130:355–91. doi: 10.1037/0033-2909.130.3.355

20. Mason JW. A review of psychoendocrine research on the sympathetic-adrenal medullary system. *Psychosom Med.* (1968) 30:631–53. doi: 10.1097/00006842-196809000-00022
21. Lupien SJ, Ouellet-Morin I, Trepanier L, Juster RP, Marin MF, Francois N, et al. The DeStress for Success Program: effects of a stress education program on cortisol levels and depressive symptomatology in adolescents making the transition to high school. *Neuroscience.* (2013) 249:74–87. doi: 10.1016/j.neuroscience.2013.01.057
22. Pagé MG, Dassieu L, Develay É, Roy M, Vachon-Presseau É, Lupien SJ, et al. The stressful characteristics of pain that drive you NUTS: a qualitative exploration of a stress model to understand the chronic pain experience. *Pain Med.* (2021) 22:1095–108. doi: 10.1093/pm/pnaa370
23. Tashakkori A, Teddie C. Quality of inferences in mixed methods research: calling for an integrative framework. In: Bergman MM, editor. *Advances in Mixed Methods Research: Theories*. London, UK: Sage Publications Ltd. (2008). p. 101–19.
24. Creswell JW. *A Concise Introduction to Mixed Methods Research*. Washington DC: Sage Publications (2015). 114 p.
25. Creswell JW, Plano Clark VL. *Designing and Conducting Mixed Methods Research*. 6th ed. Thousand Oaks, CA: SAGE (2018).
26. Cleeland CS. *The Brief Pain Inventory User Guide*. The University of Texas MD Anderson Cancer Center (2009).
27. Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS. Validity of the brief pain inventory for use in documenting the outcomes of patients with noncancer pain. *Clin J Pain.* (2004) 20:309–18. doi: 10.1097/00002508-200409000-00005
28. Kroenke K, Spitzer RL, Williams JB, Lowe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics.* (2009) 50:613–21. doi: 10.1176/appi.psy.50.6.613
29. Dupret E, Bocéréan C. La mesure du stress en milieu professionnel avec l'échelle de stress perçu (Perceived Stress Scale): pertinence des versions en dix et quatre items. *Psychologie du Travail et des Organisations.* (2013) 19:362–84. doi: 10.1016/S1420-2530(16)30049-8
30. Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess.* (1995) 7:524–32. doi: 10.1037/1040-3590.7.4.524
31. McWilliams LA, Kowal J, Wilson KG. Development and evaluation of short forms of the pain catastrophizing scale and the pain self-efficacy questionnaire. *Eur J Pain.* (2015) 19:1342–9. doi: 10.1002/ejp.665
32. Lupien S. *Stress Characteristics Questionnaire*. Montreal, Canada: Centre for Studies on Human Stress (2019). Available from: https://humanstress.ca/stress/nuts-questionnaire-2020/
33. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain.* (2005) 113:9–19. doi: 10.1016/j.pain.2004.09.012
34. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, editors. *Handbook of Pain Assessment*. 2nd ed. New York, NY: Guilford Press (2001). p. 15–34.
35. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain.* (2004) 8:283–91. doi: 10.1016/j.ejpain.2003.09.004
36. Cormier S, Lavigne GL, Choinière M, Rainville P. Expectations predict chronic pain treatment outcomes. *Pain.* (2016) 157:329–38. doi: 10.1097/j.pain.0000000000000379
37. Nakagawa S. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behav Ecol.* (2004) 15:1044–5. doi: 10.1093/beheco/arh107
38. Rowbotham MC. What is a 'clinically meaningful' reduction in pain? *Pain.* (2001) 94:131–32. doi: 10.1016/S0304-3959(01)00371-2
39. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain.* (2001) 94:149–58. doi: 10.1016/S0304-3959(01)00349-9
40. Lassen CL, Siam L, Degenhart A, Klier TW, Bundscherer A, Lindenberg N. Short-term impact of the COVID-19 pandemic on patients with a chronic pain disorder. *Medicine.* (2021) 100:e25153. doi: 10.1097/MD.0000000000025153
41. Taylor S, Landry CA, Paluszek MM, Fergus TA, McKay D, Asmundson GJG. Development and initial validation of the COVID Stress Scales. *J Anxiety Disord.* (2020) 72:102232. doi: 10.1016/j.janxdis.2020.102232
42. Arpaci I, Karatas K, Baloglu M. The development and initial tests for the psychometric properties of the COVID-19 Phobia Scale (C19P-S). *Pers Individ Dif.* (2020) 164:110108. doi: 10.1016/j.paid.2020.110108
43. Varpio L, Ajjawi R, Monrouxe LV, O'Brien BC, Rees CE. Shedding the cobra effect: problematising thematic emergence, triangulation, saturation and member checking. *Med Educ.* (2017) 51:40–50. doi: 10.1111/medu.13124
44. Guest G, Bunce A, Johnson LK. How many interviews are enough? An experiment with data saturation and variability. *Fields Methods.* (2006) 18:59–82. doi: 10.1177/1525822X05279903
45. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. *Qual Health Res.* (2016) 26:1753–60. doi: 10.1177/1049732315617444
46. van Nes F, Abma T, Jonsson H, Deeg D. Language differences in qualitative research: is meaning lost in translation? *Eur J Ageing.* (2010) 7:313–6. doi: 10.1007/s10433-010-0168-y
47. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* (2006) 3:77–101. doi: 10.1191/1478088706qp063oa
48. Braun V, Clarke V, Hayfield N, Terry G. Thematic analysis. In: Liamputtong P, editor. *Handbook of Research Methods in Health Social Sciences*. Singapore: Springer (2019). p. 843–60. doi: 10.1007/978-981-10-5251-4_103
49. Berends L, Johnston J. Using multiple coders to enhance qualitative analysis: the case of interviews with consumers of drug treatment. *Addict Res Theory.* (2005) 13:373–81. doi: 10.1080/16066350500102237
50. Corbin JM, Strauss A. *Basics of Qualitative Research*. 4th ed. Los Angeles, CA: SAGE (2015). 456 p.
51. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med.* (2014) 89:1245–51. doi: 10.1097/ACM.0000000000000388
52. *NVivo qualitative data analysis software*. 12 ed. Doncaster, Australia: QSR International Pty Ltd. (2018).
53. Fetters MD, Freshwater D. The 1 + 1 = 3 Integration Challenge. *J Mix Methods Res.* (2015) 9:115–7. doi: 10.1177/1558689815581222
54. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. *Health Serv Res.* (2013) 48:2134–56. doi: 10.1111/1475-6773.12117
55. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann Fam Med.* (2015) 13:554–61. doi: 10.1370/afm.1865
56. Kersebaum D, Fabig SC, Sendel M, Sachau J, Lassen J, Rehm S, et al. The early influence of COVID-19 pandemic-associated restrictions on pain, mood, and everyday life of patients with painful polyneuropathy. *Pain Rep.* (2020) 5:e858. doi: 10.1097/PR9.0000000000000858
57. Mun CJ, Campbell CM, McGill LS, Aaron RV. The early impact of COVID-19 on chronic pain: a cross-sectional investigation of a large online sample of individuals with chronic pain in the United States, April to May, 2020. *Pain Med.* (2021) 22:470–80. doi: 10.1093/pm/pnaa446
58. Dassieu L, Page MG, Lacasse A, Laflamme M, Perron V, Janelle-Montcalm A, et al. Chronic pain experience and health inequities during the COVID-19 pandemic in Canada: qualitative findings from the chronic pain & COVID-19 pan-Canadian study. *Int J Equity Health.* (2021) 20:147. doi: 10.1186/s12939-021-01496-1
59. Yu L, Kioskli K, McCracken LM. The psychological functioning in the COVID-19 pandemic and its association with psychological flexibility and broader functioning in people with chronic pain. *J Pain.* (2021). doi: 10.1016/j.jpain.2021.02.011
60. Sturgeon JA, Zautra AJ. Resilience: a new paradigm for adaptation to chronic pain. *Curr Pain Headache Rep.* (2010) 14:105–12. doi: 10.1007/s11916-010-0095-9
61. Rajkumar RP. COVID-19 and mental health: a review of the existing literature. *Asian J Psychiatr.* (2020) 52:102066. doi: 10.1016/j.ajp.2020.102066
62. Bakioglu F, Korkmaz O, Ercan H. Fear of COVID-19 and positivity: mediating role of intolerance of uncertainty, depression, anxiety, and stress. *Int J Ment Health Addict.* (2020) 31:1–4. doi: 10.1007/s11469-020-00331-y

63. Stein MB. COVID-19 and anxiety and depression in 2020. *Depress Anxiety*. (2020) 37:302. doi: 10.1002/da.23014
64. Stein MB. Mental health participation in the fight against the COVID-19 pandemic. *Depress Anxiety*. (2020) 37:404. doi: 10.1002/da.23027
65. Pedrosa AL, Bitencourt L, Froes ACF, Cazumba MLB, Campos RGB, de Brito S, et al. Emotional, behavioral, and psychological impact of the COVID-19 pandemic. *Front Psychol*. (2020) 11:566212. doi: 10.3389/fpsyg.2020.566212
66. Sagat P, Bartik P, Prieto Gonzalez P, Tohanean DI, Knjaz D. Impact of COVID-19 quarantine on low back pain intensity, prevalence, and associated risk factors among adult citizens residing in Riyadh (Saudi Arabia): a cross-sectional study. *Int J Environ Res Public Health*. (2020) 17:7302. doi: 10.3390/ijerph17197302
67. Serrano-Ibanez ER, Esteve R, Ramirez-Maestre C, Ruiz-Parraga GT, Lopez-Martinez AE. Chronic pain in the time of COVID-19: Stress aftermath and central sensitization. *Br J Health Psychol*. (2021) 26:544–52. doi: 10.1111/bjhp.12483
68. Codagnone C, Bogliacino F, Gomez C, Charris R, Montealegre F, Liva G, et al. Assessing concerns for the economic consequence of the COVID-19 response and mental health problems associated with economic vulnerability and negative economic shock in Italy, Spain, and the United Kingdom. *PLoS ONE*. (2020) 15:e0240876. doi: 10.1371/journal.pone.0240876
69. Zambelli Z, Fidalgo AR, Halstead EJ, Dimitriou D. Acute impact of a national lockdown during the COVID-19 pandemic on wellbeing outcomes among individuals with chronic pain. *J Health Psychol*. (2021) 18:1359105321995962. doi: 10.1177/1359105321995962
70. Hovey R, Linkiewich D, Brachaniec M. Chronic pain, vulnerability and human spirit while living under the umbrella of COVID-19. *Patient Exp J*. (2020) 7:Article 8. doi: 10.35680/2372-0247.1504
71. Colloca L, Thomas S, Yin M, Haycock NR, Wang Y. Pain experience and mood disorders during the lockdown of the COVID-19 pandemic in the United States: an opportunistic study. *Pain Rep*. (2021) 6:e958. doi: 10.1097/PR9.0000000000000958
72. Caci B, Miceli S, Scrima F, Cardaci M. Neuroticism and fear of COVID-19. The interplay between boredom, fantasy engagement, and perceived control over time. *Front Psychol*. (2020) 11:2578. doi: 10.3389/fpsyg.2020.574393
73. Rodriguez BO, Sanchez TL. The Psychosocial Impact of COVID-19 on health care workers. *Int Braz J Urol*. (2020) 46:195–200. doi: 10.1590/s1677-5538.ibju.2020.s124

Conflict of Interest: MP received honoraria as a speaker for Canopy Growth for content unrelated to the work presented here.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Pain Research

REVIEW
published: 13 December 2021
doi: 10.3389/fpain.2021.788606

Computer Mediated Automatic Detection of Pain-Related Behavior: Prospect, Progress, Perils

Kenneth M. Prkachin[1]* *and Zakia Hammal*[2]

[1] *Department of Psychology, University of Northern British Columbia, Prince George, BC, Canada,* [2] *The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, United States*

Pain is often characterized as a fundamentally subjective phenomenon; however, all pain assessment reduces the experience to observables, with strengths and limitations. Most evidence about pain derives from observations of pain-related behavior. There has been considerable progress in articulating the properties of behavioral indices of pain; especially, but not exclusively those based on facial expression. An abundant literature shows that a limited subset of facial actions, with homologs in several non-human species, encode pain intensity across the lifespan. Unfortunately, acquiring such measures remains prohibitively impractical in many settings because it requires trained human observers and is laborious. The advent of the field of affective computing, which applies computer vision and machine learning (CVML) techniques to the recognition of behavior, raised the prospect that advanced technology might overcome some of the constraints limiting behavioral pain assessment in clinical and research settings. Studies have shown that it is indeed possible, through CVML, to develop systems that track facial expressions of pain. There has since been an explosion of research testing models for automated pain assessment. More recently, researchers have explored the feasibility of multimodal measurement of pain-related behaviors. Commercial products that purport to enable automatic, real-time measurement of pain expression have also appeared. Though progress has been made, this field remains in its infancy and there is risk of overpromising on what can be delivered. Insufficient adherence to conventional principles for developing valid measures and drawing appropriate generalizations to identifiable populations could lead to scientifically dubious and clinically risky claims. There is a particular need for the development of databases containing samples from various settings in which pain may or may not occur, meticulously annotated according to standards that would permit sharing, subject to international privacy standards. Researchers and users need to be sensitive to the limitations of the technology (for e.g., the potential reification of biases that are irrelevant to the assessment of pain) and its potentially problematic social implications.

Keywords: pain, measurement, facial expression, automation, assessment

OPEN ACCESS

Edited by:
Kai Karos,
Open University of the Netherlands, Netherlands

Reviewed by:
Joseph Walsh,
Bath Spa University, United Kingdom
Ali Khatibi,
University of Birmingham, United Kingdom

***Correspondence:**
Kenneth M. Prkachin
kmprk@unbc.ca

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *02 October 2021*
Accepted: *15 November 2021*
Published: *13 December 2021*

Citation:
Prkachin KM and Hammal Z (2021) Computer Mediated Automatic Detection of Pain-Related Behavior: Prospect, Progress, Perils. Front. Pain Res. 2:788606. doi: 10.3389/fpain.2021.788606

INTRODUCTION

The International Association for the Study of Pain's recent revision to the definition of pain ["an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage;" (1)] added several contextualizing notes. First, pain is "always a personal experience, influenced…by personal, psychological, and social factors." Second, "a person's report of an experience as pain should be respected." Lastly, verbal description is only one of several behaviors to express pain." The first and second recognize that the experience of pain is subjective and falls into the category of phenomena we call "feelings." The second addresses the common temptation, when a phenomenon is subjective, to be skeptical about its reality or its potential to be interrogated scientifically. The third recognizes that evidence about pain exists in various types of behavior. While we can acknowledge that there is much in the experience of pain that is unique and individual, if we are interested in advancing understanding of pain, either from a purely scientific point of view or for utilitarian purposes of management and control, then we must achieve some consensus on the evidence we use to infer its presence and properties.

The experience of pain cannot be directly measured. Instead, there are two general categories of pain indicators. One consists of changes in the body, especially but not limited to the central nervous system, that are believed to mark and quantify pain and that can be measured more-or-less directly by some form of instrumentation. The other consists of behavior. The vast majority of pain indicators, including verbal descriptions, fall into this category. Other behavioral pain indicators include instrumental acts, such as withdrawal or avoidance and expressive acts, such as vocalizations or grimacing.

In recent years advances in technology, accompanied by expanding analytic tools in the area of computer vision and machine learning (CVML), have been applied to some behavioral pain indicators in efforts to improve on them for both scientific and practical reasons. Until recently, most progress has been made toward automatic assessment of facial expression of pain (2, 3). Although in everyday pain experience we encounter associations between body movement and pain, the communicative functions of body movements in relation to pain have been fairly unexplored in automatic pain assessment. Notable exceptions are to be found in the work of Aung et al. [(4), see also Egede et al. (5)] who found association between pain and certain bodily protective behaviors, such as guarding/stiffness and bracing/support.

In this article, we describe the advent of such approaches, as they relate to facial expressions of pain, beginning with the behavioral roots that gave rise to them. We articulate the prospects foreseen for such approaches, then describe early progress in the form of "demonstrations of concept." We then go on to summarize key developments and address emergent applications of the work, including the development of commercial products. In the course of this narrative, we highlight emergent problems that, we believe, should qualify enthusiasm about the field.

VERBAL ASSESSMENT OF PAIN

While it is possible to gain insight about a subjective process, that insight often comes indirectly—by operationalizing it in the form of a measure. In the field of pain, operationalized verbal reports have become a standard—indeed it is common to see verbal report referred to as the "gold standard." Verbal reports of pain can be obtained about different dimensions but pain intensity is overwhelmingly the most frequently assessed. The widely used visual analog scale (VAS), in which the respondent marks a spot on a line of finite length to characterize their pain, is a variation on verbal report. In clinical and population-based studies, verbal descriptor or VAS scales are commonly used to characterize certain pain states or as outcome measures in studies of interventions. The 0–10 numeric rating scale was advocated for and implemented widely in health-care settings as a fifth vital sign.

Limitations of Verbal Assessment of Pain

The fact that verbal report techniques are used ubiquitously is a testament to their utility. However, concerns about their potential shortcomings are common. One concern is epistemological, reflecting an underlying belief that scientific inquiry should be based in measurements of things that are objectively observable. But there are others. For one, verbal reports bear an uncertain relation to the underlying experience. They can be shown to behave in a way that should coarsely correspond to an underlying pain state, such as when people use lower numbers or words reflecting lesser pain to describe their pain after being administered a known analgesic. However, when a patient with low back pain who initially gave a rating of 8 to their pain now gives a rating of 4 after a rehabilitation program does that mean they are in half as much pain? In the historical debates about pain measurement, this issue was at the center of several attempts to develop psychophysical techniques with ratio-scale properties (6–8).

Even if it can be shown that verbal ratings vary according to expectations in experimental and clinical studies, it is not possible to be certain that all individuals use the scales in the same way. Some people are more sensitive to variations in the experience and more precise reporters than others. Williams et al. (9), for example, reported a lack of concordance between patients and consistency within patients in their use of visual analog and numeric rating scales as they actively interpreted the meaning of their experiences.

Often, variations in the operationalization reveal inconsistencies in the characterization. When different techniques are used to assess the painfulness of the same level of nociceptive stimulation in experimental studies, or the same patient at the same time in clinical studies, the evaluations are often incommensurate. For example, in one of our recent studies, participants were asked first to rate cold pressor pain using a VAS. Then, at the end of the study, they were asked to rate the maximum pain using the pain intensity rating (PIR) of the McGill Pain Questionnaire. Participants who gave the maximum pain rating according to the VAS—a rating corresponding to

"worst pain imaginable" frequently gave a PIR rating implying pain of considerably lower intensity.

One of the most well-known features of verbal reports is their extraordinary malleability. This property has been known for a long time, featuring in Beecher's (10) classic *Measurement of Subjective Responses* in the form, among other things, of the placebo effect. Craig's early studies of the social modeling effect [e.g., (11)], showed that exposure to tolerant or intolerant social models could make participants rate electric shocks less or more painful, respectively. Such malleability may, of course, simply exemplify that pain is an extremely plastic phenomenon. On the other hand, recognizing that verbal report is under exquisite control of the perceiver raises concern whether what is being measured is instead the response to personal or social expectations embedded in the conditions of observation such as expectancy effects or demand characteristics, independent of any true effect on the pain experience itself. One example of the concern arose in studies of hypnosis that made use of the "hidden observer" technique (12). Participants under hypnosis were given suggestions that they would experience an analgesic state. They then rated the painfulness of cold in the cold-pressor test. Participants were also told that under hypnosis they would have access to the experience of their hidden observer—a part of them that would experience the pain as it was—and that they were to give the ratings of the hidden observer after they rated their own pain under hypnotic analgesia. The studies showed a dissociation between the ratings of the hypnotized subject and the same subject's hidden observer Spanos and Hewitt (13), however found that the hidden observer's ratings could be easily diverted by manipulations of what the participant expected that the researcher expected.

Similarly, self-presentation biases are likely to come into play and distort controlled verbal reports in a species as socially responsive as humans. A common self-presentation bias in the pain context is stoicism. When self-report is the criterion, studies (both clinical and experimental) routinely find, for example, that men report lower pain than women (14). It is, of course, possible that this reflects a true difference in pain sensitivity between the sexes, but there is an obvious socialization difference in which masculinity is equated with enduring pain that can also account for the difference.

A final shortcoming of verbal report in studies of pain is that there are important instances in which verbal reports cannot be obtained because the respondent is incapable of using words to describe their pain (for example, preverbal infants, people with profound verbal communication impairments and non-human animals), or people who, though capable of communicating verbally, are impaired in the ability to communicate reliably about pain (such as in types of dementia).

PHYSIOLOGICAL ASSESSMENT OF PAIN

There is a substantial history of search for alternatives to self-report. A diversity of physiological measures has been promoted over the years, including measures of autonomic responses such as electrodermal activity (15), oxygen saturation (16), heart-rate variability (17), and evoked potentials (18). With the advent of neuroimaging procedures measures of regional cerebral bloodflow have become ubiquitous in pain studies. Some have been promoted as true "central registers" of the pain experience, but none are widely recognized as such (19).

Limitations of Physiological Assessment of Pain

A physiological measure of pain has been a kind of "holy grail" among some researchers and clinicians. Physiological variables such as those noted are routinely deployed in both basic and clinical studies but have not achieved consensual status as measures of pain outcomes. Some, such as electrodermal activity or heart-rate variability, serve as indices of processes that are affected by pain, such as autonomic arousal. As measures of pain, they are sometimes overly responsive and therefore poorly discriminating of variations in pain states, sometimes insufficiently responsive and therefore also poorly discriminating, and sometimes covary with other affect states with which pain is correlated, such as fear. Neuroimaging procedures have identified various brain regions in which activation varies in accordance with other evidence of pain; however, they are distributed across networks in a manner that does not lend itself to simple interpretation as pain indicators. Most physiological assessment techniques are at least modestly invasive, involving special instrumentation and sometimes highly specialized laboratory environments and therefore do not lend themselves to study in ecologically normative conditions.

PAIN ASSESSMENT BASED ON FINE-GRAINED FACIAL OBSERVATION

The insight that behavior is fundamental to the understanding of pain gained currency with the development of behavioral approaches to pain management. As Fordyce (20) observed, a person has to do something for it to be known that they are in pain. The early behavioral approach was based in the learning theory of the day but did not make nuanced distinctions about the properties of pain-related behaviors that varied by topography.

The model brought an emphasis on observation and precise definition and assessment of behaviors that, curiously, dovetailed with the concerns of students of emotion.

The study of emotion had venerable roots in the work of Charles Darwin. In *The expression of the emotions in man and animals*, Darwin (21) argued that emotions are phylogenetically shared with other species. He described how various affective states, including pain, are represented in specific behavioral topographies, especially but not exclusively facial expression.

Interest in the role of the face in communication of affect revived in the late 1960's, reflecting in part the influence of studies supporting the idea that facial expressions of certain emotions are universal across human cultures (22). Subsequent refinements in methods for studying facial expressions laid the foundation for examining their role in communicating information about pain.

In 1978, Ekman and Friesen published the Facial Action Coding System (FACS). This is a system for deconstructing any facial movement into its constituent actions based on the changes that appear when an individual muscle or combination of muscles are activated. Observers trained to FACS proficiency then view facial expressions and describe their constituent actions in terms of 44 action units (AUs) or action descriptors (ADs). Most AUs can be described in terms of their intensity. Intensity coding for most AUs is on a 6-point A—E scale, where a code of A is assigned to a trace of an action, B to an action that meets minimum requirements for the action, E to an action that is as strong as it could be, and codes in between refer to gradations between meeting the minimum requirements and maximum intensity (note that, in quantitative analyses the alpha codes are transformed to numbers between 1 and 5; if the action has not occurred a code of 0 is assigned as default).

The system is thus anatomically based, atheoretical, and relatively objective ("relatively" because inferences are still involved; for example, when rating intensity). It is manualized such that, with intensive study, an observer can learn the system within about 100 h. Data quality when performed by observers who have established proficiency in the system by passing a proficiency test, is generally sufficient to meet conventional reliability standards and the system is generally considered to be the "gold standard" for assessing facial action.

The FACS has been applied extensively in studies to characterize the appearance of the face when a person is in pain. A systematic review of 37 studies (23) reported that, for both experimental and clinical pain, a subset of facial actions reliably discriminates between pain and no-pain conditions. These are: brow lowering (FACS AU 4), orbit tightening (AUs 6 or 7), levator tightening (AUs 9 or 10), and mouth opening (AUs 25, 26, or 27). Eyelid closing (AU 43) also consistently discriminates between pain and no pain in studies of clinical pain. The same actions discriminated pain from no pain independent of the participants' cognitive status (impaired vs. unimpaired).

Systems resembling FACS have been developed for studies of pain in children. The two systems that have been applied most widely are the Neonatal Facial Coding System [NFCS; (24)] and the Child Facial Coding System [CFCS; (25)]. Rather than being defined by the underlying facial musculature of the constituent actions, NFCS and CFCS codes are based on appearance changes. In both neonates and young children, the codes that have been found most consistently to discriminate pain from no pain conditions are homologous to the codes that distinguish pain from no pain conditions in adults, including seniors; namely, brow bulge (NFCS)/brow lower (CFCS), eye squeeze (both systems), nasolabial furrow /nose wrinkle (NFCS), nasolabial furrow, upper lip raiser (CFCS) (26). Various other facial actions have been associated with pain in neonates and children. Nevertheless, the smaller "core" subset appears with remarkable consistency across types of pain and the human life-span, including among the aged. There is also a noteworthy similarity with the facial actions reported to be associated with pain in non-human animals that have been studied to date [e.g., (27)].

Limitations of Fine-Grained Facial Observation of Pain

Somewhat remarkably, despite the substantial scientific literature documenting the properties of facial expressions of pain, the work has had little application in basic science or clinical studies of pain. The simple reason for this is that objective description of facial action by FACS or similar systems is burdensome. FACS is implemented by human observers who require training to render assessments that are sufficiently reliable for scientific purposes. Implementing FACS in scientific or clinical studies cannot be done practically in real-time because coding requires multiple observations of behavioral samples to identify the separate actions of separate muscle groups. Ordinarily it requires slow-motion and stop-action to settle on a final set of codes. This makes the coding process lengthy–a final code from a sample of behavior is typically estimated to require a coding time: real time ratio of around 100:1. Conducting studies with requisite numbers of participants and observations quickly becomes arduous and, for human observers, oppressive. Realistic application in clinical settings is impractical. Although some work has aimed at reducing training and coding time by focusing on only facial actions that have been empirically associated with pain (28), even modified procedures are problematically time-consuming. Further, the measurement rendered by human observers is insufficiently granular and continuous to render certain kinds of information that could provide the insights into pain processing that the face may be capable of; for example, temporal information about the onset and decay of certain facial actions that may be informative about such issues as the relative reflexivity or conscious modulation of the sufferer.

For these reasons, since the inception of fine-grained systems for measuring facial action, there has been an underlying question whether advances in information technology could render a technique as reliable and valid as facial coding by trained observers that would reduce the burden of observation, that would not be subject to human observers' susceptibility to fatigue and error, that might be more sensitive and better able to represent dynamic changes. Development of the field of affective computing appeared to address this prospect.

TOWARD AUTOMATED ASSESSMENT OF PAIN FROM NON-VERBAL BEHAVIOR

Affective computing has been defined as "computing that relates to, arises from, and deliberately influences emotion" (29, 30). It subsumes a wide range of topics and applications, one of which is the measurement and modeling of affective processes. Affective processes like pain have behavioral markers, including but not limited to facial expressions, that can be captured and stored by technology. Decoding their messages is a kind of pattern recognition. Advances in computer and data science enabled by the development of neural nets and machine learning, which had proved to be successful modeling pattern recognition, appeared to offer a technological solution to the burden associated with decoding facial expressions. Further potential benefits, such as rapid processing and the ability to render more precise

information about movement dynamics than can be effectively obtained from human observers, appeared possible.

Some of the earliest demonstrations of the feasibility of such automated analysis of facial expression appeared in work by Bartlett et al. (31) and Cohn et al. (32). Bartlett et al. obtained images of FACS upper-face AUs varying in intensity from 20 people. Processed by a two-layer neural network, a hybrid classification system combining holistic spatial analysis, facial feature measurement, and analysis of motion flow fields was able to correctly classify 92% of the six facial actions [of which three (AUs 4, 6, 7) had been implicated in studies of pain], outperformed naïve human judges, and approximated the performance of human experts. Cohn et al. used video frames of 15 FACS AUs or AU combinations as training stimuli. After alignment, facial landmarks were marked and then automatically tracked using an algorithm to estimate optical flow across images. minant function analysis produced 92% or higher agreement with the classifications of a human coder in a training set and between 81 and 91% (depending on facial region) in a cross-validation set. These studies strongly suggested that advances in computer vision methods combined with advanced statistical analysis could, in principle, make automated analysis of facial expression possible.

The advent of techniques to automatically measure facial expressions naturally stimulated interest in extending the technology to the measurement of facial expressions of pain. Effective automated assessment held promise to overcome barriers to more widespread scientific and practical applications of facial expression measurement. In principle, it could reduce or eliminate the need for human coders thereby managing the problem of observer burden. Once tested sufficiently and validated, an automated system could potentially be more reliable than measurement by human observation because it could reduce variability and human error. Early work on automating measurement of human emotional expressions began to reveal properties of facial action that had been impractical to study. For example, using an automated facial analysis technique, Ambadar et al. (33) showed that different categories of smiling (polite, amused, embarrassed) differed in terms of velocity, duration, and association with head movements. From a scientific perspective, the prospect of an automated system opened the tantalizing possibility of measuring momentary dynamic changes in pain-related facial expressions to draw similar inferences about its meaning and underlying determinants. From a practical perspective, an automatic, objective, reliable, and efficient assay of the occurrence and intensity of pain could improve clinical pain assessment, allowing health-care personnel to provide better treatment to patients, with little to no increase in cost (2). It could also support pharmaceutical therapies by providing an objective quantitative tool for evaluating the efficacy of current and new analgesics and serve as an objective complement to self-reported pain measures in clinical trials of drug or device interventions to reduce pain.

Methodological Foundations

To learn the association between pain occurrence or intensity and facial behavior, recordings of participants responding to painful conditions are needed in order to train and test classifiers. Samples of sufficient size to estimate training parameters and perform validation analyses are necessary. The number of participants should be motivated by two factors. One is the number needed to achieve saturation in the performance of the predictive models (i.e., automatic classifiers). The other is the number needed to enable sufficient power in the statistical models for quantifying the contribution of the used variables in the predictive models. For instance, in prior work on a related problem (training automated classifiers for facial action units), it was found that automatic classifier performance saturates at about 60 participants in the training set (34). With 25 participants in the UNBC-McMaster Shoulder Pain Expression Archive Database, the number of available participants is far lower than that minimum number needed. Additionally, independent criteria for establishing the absence, presence, or intensity of pain (i.e., "ground truth") must be present. Although they have not been as widely tested in pain studies, ground truth in automated pain assessment has mostly been derived from annotations by expert observers (using FACS or a variant of FACS) of video recordings of facial expression of pain. However, there must be sampling in conditions in which it is reasonable to assume that pain has occurred (such as during a clinical test, or during exposure to artificially induced painful conditions, such as noxious heat), and in conditions when pain is unlikely. As an alternative or supplement, judgment studies can be performed in which observers (who might vary in expertise) rate recordings on an appropriate scale of pain intensity. Another alternative that has only recently come to be explored is the subjective judgments of participants undergoing the potentially painful procedure. Finally, known conditions can serve as ground truth, such as when, in one experimental condition, a participant is exposed to a stimulus known to cause pain and in another, they are not. If ground truth is based on annotations or ratings by human observers, they must also meet criteria for acceptable reliability. Meeting the aforementioned criteria is a challenging task but has been achieved by several groups [(35–39)].

Because of the power requirements of machine learning and classification procedures, there is an issue related to the density and precision of annotations. Analyses are based on the recordings made in the aforementioned clinical or experimental conditions. A behavioral sample can be annotated at the level of the overall sequence using a single observation or a summary, which yields one measure per sequence. Alternatively, depending on the annotation method, it can be annotated at the level of the individual frame. Whereas, annotation at the level of the frame provides considerable amounts of data for training and validation purposes, annotation at the level of the sequence provides but one per participant and condition, with obvious implications for sampling in the pain recording phase of any study. In either case, but more particularly for studies in which annotation is frame-by-frame, at least in data collected to date, the distribution of pain intensities is problematic, with there usually being a much higher number of frames in which annotations suggest no pain than pain, with implications for training models.

In part because of the resources required to meet the forgoing criteria, but also because experimentation with different CVML methods benefits from comparison and calibration against extant work, databases that can be shared for model testing are desirable. The UNBC-McMaster Shoulder Pain Expression Archive Database [(40, 41)] was the first to address this need. The archive contains video recordings of people with shoulder pain taken during active abduction, flexion, internal and external rotation of their affected and unaffected shoulders (41). It comprises 200 video sequences from 25 different participants (66% female). For each sequence, the distribution includes 66 Active Appearance Model (AAM) tracked landmarks (fiducial points around the eyes, eyebrows, and mouth) at the frame level and per-frame and per-video pain score annotations. Expert labeled FACS codes were scored using a 0–5 ranking of the intensity of the facial actions in most cases. Intercoder agreement as calculated by the Ekman–Friesen formula (42) was 0.95. the participants' self-reported pain intensity and an independent observer's ratings of pain intensity (OPI) were annotated at the sequence level. Offline observer ratings were performed on a 6-point Likert-type scale that ranged from 0 (no pain) to 5 (strong pain). To assess inter-observer reliability of the OPI pain ratings, a second rater independently rated 210 randomly selected videos. The Pearson correlation between the observers' OPIs was 0.80, which represents high inter-observer reliability.

Since being made available to qualified researchers, the Pain Archive has been the most widely used dataset for exploring automatic pain assessment from facial expression, accounting for approximately 41% of the literature published in this field according to a 2019 systematic review (3).

A smaller number of studies (43–49) have made use of BioVid (50), a heat pain database. BioVid contains recordings of 87 people exposed to four intensities of experimental heat pain and a no pain baseline. Each intensity (including no pain) was presented 20 times in a random sequence. Each video excerpt has a duration of 5.5 s. Unlike the UNBC-McMaster Shoulder Pain Expression Archive Database, ground truth is based on stimulus intensity, rather than a measure of pain expression.

A third database, EmoPain (4) contains recordings from 22 adults with low back pain. The recordings were taken while the patients engaged in movements resembling common therapeutic tasks for back pain patients. Data streams include audio recordings, 3D motion capture, and electromyographic recordings from the paraspinal muscles in addition to facial expression. Measures available for ground truth include patient pain and anxiety ratings, and offline observer ratings using a joystick method. EmoPain has not yet been publicly released as had been planned.

Proof-of-Concept Studies

One of the earliest efforts to develop an automated system for measuring pain expression appeared in Ashraf et al. (51). The authors employed recordings from the UNBC-McMaster Shoulder Pain Expression Archive Database of shoulder-pain patients described above. They had been quantified at the level of the individual video frame by a FACS-based index of expressive intensity, dubbed the Prkachin Solomon Pain Index [PSPI; (41, 52)], and consisting of the summed scores of AUs that have consistently been associated with pain in observational studies. After transformations to optimize registration of the face, support vector machines (SVMs) were trained to classify full sequences or individual frames as showing pain or no pain. The best combination of representations resulted in hit rates of 77 and 82% for sequence level and frame-level classification, respectively, and false acceptance rates of 44 and 30%, showing that it was possible to obtain reasonable differentiation of pain from no pain states when evaluated with respect to the ground truth of direct facial measurement by trained observers. Unsurprisingly, the more granular frame-level approach provided better performance. **Figure 1** displays performance of both approaches for a representative participant.

In another early study of automatic pain detection, Littlewort et al. (36) employed a system for automatic detection of FACS AUs to examine facial changes during exposure to experimental pain produced by immersion of the arm in ice-water and to compare those changes with actions performed when participants pretended to be in pain. Genuine pain was associated with increases in six automatically detected representations of AUs previously associated with cold-pressor pain in studies using human observers. "Faked" pain was associated with 11 automatically coded actions. In a subsequent machine learning phase, automated facial action parameters were processed via a Gaussian SVM in an attempt to discriminate genuine from faked pain. The resultant 2-alternative forced-choice percent correct value of 88% substantially exceeded the performance of naïve human observers at 49%.

Lucey et al. (53), also using the UNBC-McMaster Shoulder Pain Expression Archive Database, applied a system combining Active Appearance Models (AAMs) for tracking face shape and appearance, input to SVM's for pain and AU classification at the level of the individual video frame. Ground truth consisted of expert-coded FACS AUs, including, but not limited to the PSPI. In a test of the system for directly classifying pain (i.e., predicting a PSPI score of >0) the Receiver-Operating-Characteristic (ROC) based A' metric yielded a score of 0.75, indicating performance substantially greater than chance. An indirect classification system, predicting pain from an alternative set of individual FACS AUs that excluded two components of the PSPI and included AU12, performed slightly better, achieving an A' score of 0.77, relative to 0.78 for the PSPI. Building upon those results, Lucey et al. (54) again used a combination of AAM/SVM representations to derive parameters of similarity normalized points (SPTS) and canonical normalized appearance (CAPP). These were trained to detect individual AUs and the PSPI metric. SPTS and CAPP solutions were then used individually and in combination to evaluate performance. With some exceptions, the individual representations performed reasonably at both AU detection and overall PSPI prediction. Combining both parameters yielded an A' value of 0.84 at predicting the PSPI index.

Hammal and Kunz (55) proposed a hybrid machine learning approach to classifying spontaneous expressions of experimental pain, based on the Transferable Belief Model. The model was based on the dynamic fusion of appearance features around the wrinkle areas (the deepening of transient facial features). Video

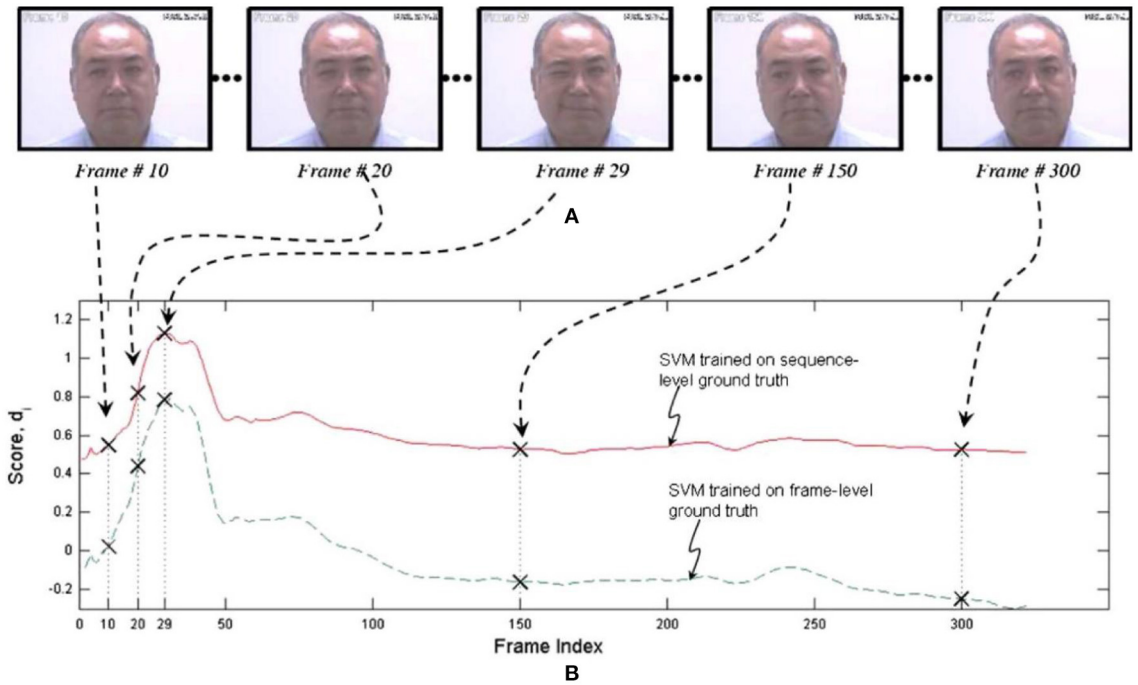


FIGURE 1 | SVM scores for sequence- and frame-level ground truth. The upper pictorial representations **(A)** are the video frames corresponding with the crosses on the respective SVM score plots **(B)** below. Reprinted with permission from Ashraf et al. (51).

sequences of participants responding to painful or non-painful heat stimulation were classified in a 2-alternative forced-choice paradigm, achieving a correct classification rate of 81.2%. A test of the ability of the system to correctly discriminate among pain, posed expressions of six basic emotions, and neutral expressions (an 8-alternative forced choice) achieved a correct classification rate of 84.5%. Automatic classification outperformed untrained human observers. Importantly, these findings demonstrated the feasibility of automatically differentiating pain from other emotional expressions. Unlike approaches that rely exclusively on static information from video recordings, the model incorporated temporal changes in features, thus more closely approximating the perceptual processes of human observers.

Most approaches to pain detection seek to determine only whether pain is present or absent. Hammal and Cohn (56), extended previous efforts by attempting to classify pain intensity (as opposed to presence). Using the UNBC-McMaster Shoulder Pain Expression Archive Database, they defined four pain intensity scores from the PSPI metric: none (PSPI = 0), trace (PSPI = 1), weak (PSPI = 2), and strong (PSPI > = 3). For each video frame, AAMs were first used to track and register rigid and non-rigid face motion. Based on this information, the canonical appearance of the face (CAPP) was extracted for each frame. CAPP features were then rescaled to 96 × 96 pixels and passed through a set of Log-Normal filters of 7 frequencies and 15 orientations. The extracted spatial face representation was then aligned as a vector of 9,216 features and used by four SVMs trained separately to measure the four pain intensity levels. Results showed fair-to-good classification of the intensity levels, depending on the classification accuracy metric and method of validation between training and testing data, with moderate-to-high consistency between automated measurement and the original PSPI metric. Several other researchers have described effective CVML methods for assessing pain intensity from facial expression [(45, 47, 57–62)]. In short, the data suggest that automated assessment of expressed pain intensity is feasible.

These early efforts provided an initial proof-of-concept that the occurrence of pain can be automatically measured from the face. There have since been scores of studies supporting the concept [see Werner et al. (3) for a survey of work to 2019].

Applications in Specific Populations

Interest in evaluating pain by assessment of non-verbal expression has been driven to a significant extent by clinical concerns; in particular, the fact that large cohorts of people cannot report on their pain because of verbal communication deficits. These include infants and young children and people with neurological impairments, especially dementias. There are extensive literatures describing validated techniques for assessing pain via facial expression and other types of non-verbal behavior in neonates and young children (63) and in dementia (64). Many suffer from the same problem of burden associated with observational techniques described above; consequently, there has been a similar interest in development of automated measures for these populations.

Automated Assessment of Pain in Infants and Children

There have been several efforts to develop automated systems for assessing pain in infants and children (65). Most have made use of a publicly available resource, the Classification of Pain Expressions (COPE) database (66). The database consists of 200 still photographs taken of neonates during five conditions, one of which was undergoing blood sampling by lancing of the heel. In an initial study, 88% correct classification in distinguishing the response to heel lancing from pain from rest, crying, air-puff, and friction conditions was achieved with a SVM approach. In a later study, using techniques based on processing of image textures and SVM's, an Area-Under-the-Curve ROC value of 0.93 was obtained discriminating pain from non-pain conditions.

With recordings obtained from neonates undergoing heel-lancing, Zamzmi et al. (67) extracted optical flow strain measures to train a K-nearest neighbor classifier, achieving 96% correct classification distinguishing pain from no pain, as evaluated against the ground-truth of nurses' ratings on an infant pain scale incorporating assessments of facial expression, among other behaviors.

Sikka et al. (37) studied children, aged 5 to 15, during different phases of treatment for appendicitis. An automated procedure—the computer expression recognition toolbox (68)—was used to detect FACS AUs, which were then used in logistic regression to classify pain, achieving Area-Under-the-Curve values of 0.84–0.94 predicting pain.

Automated Assessment of Pain in Aging and Dementia

Kunz et al. (69), using FACS, showed that facial pain expressions were able to document pain among patients with dementia who could not articulate valid verbal pain ratings and that patients with dementia showed a greater pain reaction than controls. As with other applications of behavioral measurement, this knowledge has been slow to affect clinical practice because of the measurement burden problem highlighted above. This has motivated the pursuit of automated systems for evaluating pain expression in dementia.

Progress in this pursuit has recently been documented by Rezaei et al. (70). Using video recordings taken from the UNBC-McMaster Shoulder Pain Expression Archive Database and a new dataset of elderly people with and without dementia undergoing potentially painful physiotherapy maneuvers a computer vision model of fully automated detection of pain expression was developed and evaluated. The model attempted to approximate the perceptual processes of human observers, who take into account temporal changes in expression by pairing target frames and reference frames. The best performing models, when evaluated against a pain/no pain decision based on the PSPI metric, yielded Area Under the Curve values of 0.86, and 0.85 for per-frame detection of people with dementia and those without, respectively. This supports the feasibility of automatically detecting pain-related facial actions in this verbal-communication-impaired population and is all-the-more remarkable when considering the subtlety of the actions evaluated and the presence of perturbing conditions, such as body motion out of plane and variations in lighting.

Automatic Detection of Self-Reported Pain

The bulk of this work has focused on modeling pain as represented in facial expression. More recently, however, some researchers have attempted to model other pain parameters, including sufferers' self-reports. To date, four studies have investigated automatic assessment of self-reports of pain, using video from the UNBC-McMaster Shoulder Pain Expression Archive Database. Lopez-Martinez et al. (45) proposed a two-step learning approach to estimate pain intensity as self-reported on a VAS. The approach began with a Recurrent Neural Network to automatically estimate PSPI scores at the level of individual video frames. The estimated scores were then fed into personalized Hidden Conditional Random Fields, used to estimate the self-reported VAS pain scores at the sequence level. To account for individual differences in facial expressiveness, an individual facial expressiveness score (the ratio of an independent observer's pain intensity rating) to the VAS was introduced.

A limitation of the foregoing technique is that it required retraining on previously acquired VAS ratings and thus could not generalize to previously unseen participants. To overcome this limitation, Liu et al. (59) employed another set of predefined personalized features (i.e., age, gender, complexion) to automatically estimate self-reported VAS ratings. The authors combined facial shape with these features to train an end-to-end combination of Neural Network and Gaussian Regression model (named DeepFaceLIFT), for VAS pain intensity measurement from video.

Szczapa et al. (61), proposed a video-based measurement of pain intensity scores using the dynamics of facial movement. Gram matrices formulation was used for facial point trajectory representations on the Riemannian manifold of symmetric positive semi-definite matrices of fixed rank. Curve fitting and temporal alignment were then used to smooth the extracted trajectories. A Support Vector Regression model was then trained to encode the extracted trajectories into ten pain intensity levels consistent with the VAS pain intensity measurement.

Erekat et al. (57) proposed a spatio-temporal Convolutional Neural Network–Recurrent Neural Network (CNN-RNN) model for automatic measurement of self-reported pain and observed pain intensity, respectively. The authors proposed a new loss function that explored the added value of combining different self-reported pain scales in order to improve the reliability of pain intensity assessment. Using an automatic spatio-temporal architecture, their results showed that enhancing the consistency between different self-reported pain intensity scores enhances self-reported pain estimation.

LIMITATIONS, CONSTRAINTS, AND PERILS OF AUTOMATED ASSESSMENT OF PAIN

Progress toward automated analysis of pain in the past decade has been steady; nevertheless, the field is still in early development.

It is a prudent time to consider some of the limitations of the approaches developed so far and problems that further studies will have to acknowledge or confront.

Alternatives for Ground Truth

Most efforts for automatic assessment of facial expression of pain have focused on frame-level pain intensity measurement such as the FACS-based PSPI metric. The emphasis on frame level scores, from static images or a subset of images, is consistent with approaches to objective AU detection more generally.

An alternative, simpler, approach to assessing facial expression in pain is the judgment study. Using this technique, raters, who may be naïve or could have varying levels of sophistication (e.g., being trained to recognize FACS AUs or having clinical experience with pain), view recordings of subjects who may be in pain and evaluate how much pain they appear to be in by using some kind of rating scale. The number of raters can be adjusted to meet a target reliability criterion for averaged ratings (e.g., intraclass correlation ≥ 0.80) (71). The obtained aggregate scores can then be used as the ground truth of pain intensity score. The judgment study approach is more suitable to evaluating pain intensity at the sequence level because frame-level evaluation is beyond human resolving capacity. It is possible, however, that paradigms that combine slow-motion replay with use of a dial/joystick manipulandum could capture temporal changes in pain action with sufficient reliability and sensitivity to render meaningful measurement. Considering their greater simplicity and reduced burden, it is somewhat surprising that judgment study approaches have not been employed to a greater extent in studies of automated pain assessment. Indeed, because they are based on a holistic analysis that does not assume independence of an expression's component actions and probably represent human perceptual processing more realistically, they likely have advantages over measurement of specific facial actions.

Generalizability

With few exceptions [e.g., (37)], previous efforts in automatic assessment of pain have focused on a single type of pain [shoulder pain, controlled heat; (3, 72)]. Pain comes in a variety of types, differing by modality (heat, electric, chemical), site, nature (clinical vs. artificial), and history (acute vs. chronic) that may produce different behavioral responses both within and across modalities. Given the variety of pain experiences, a variety of procedures, both experimental and observational, participants, and sensors are needed (72). The models and solutions that have shown promise for automatic detection are based on limited sampling. There is considerable evidence from direct facial measurement studies that facial expressions of pain involve a common core of actions (23, 52), but recent findings indicate that those actions come in different clusters (73, 74). This points to a need to collect further databases that sample a broader range of pain types as a way to assess the generality and generalizability of extant and novel models and solutions.

An important related need is to test approaches individually and in head-to-head comparisons across multiple databases. No studies have explicitly trained and tested classifiers on different databases in order to evaluate generalizability of automatic pain assessment across databases. Unless generalizability between separate databases is examined, it remains unknown whether methods developed in one database would be valid in others.

Care needs to be taken to address other issues of generalizability as well. Three crucial dimensions that need to be taken into account are "race," gender, and ethnicity (75). There is an ample literature showing that, apart from facial actions, skin color coding for race has a significant effect on how pain in others is judged (76, 77), and equally abundant literatures showing that race and sex affect pain treatment and outcomes (78, 79). With the exception of the non-publicly available database collected by Sikka et al. (37) demographic information is incomplete or lacking in many instances. In future research, it will be important to systematically collect participants' demographic information to investigate the variance/invariance of pain experience and measurement in order to provide a more comprehensive assessment of pain occurrence and intensity.

Bias

There is recent evidence that algorithms arising from deep-learning approaches to processing the face perform differently as a function of race and sex (for example at facial recognition), sometimes to a considerable degree (80). Likely a consequence of the fact that the datasets used for training largely sample unrepresentatively; i.e., from young, light-skinned, male populations, increasing awareness of the existence and implications of algorithms that are biased raises serious concerns about issues of fairness. The issue has become of sufficient general concern to lead to calls to ban certain applications of artificial intelligence, including work on mental health diagnosis and detection of deception (81).

That the issue of biased behavior of algorithms likely applies to detection of pain was demonstrated by Taati et al. (82), who compared the performance of currently available facial landmark and facial action unit detection algorithms on a dataset consisting of facial expressions showing various degrees of pain in a population of older people with dementia and older people living independently. Ground truth was landmark identification and facial action unit coding by human experts. Performance of the pre-trained algorithms at landmark detection was significantly better for independent-living seniors than for those with dementia. Retraining the algorithms with representative examples of faces of independent-living and seniors with dementia was able to improve performance significantly. With respect to detecting facial action units by available pre-trained algorithms, there was no difference between independent-living seniors and those with dementia, possibly because the algorithms performed poorly in general. The results emphasize the importance of sampling broadly and representatively with respect to subject group and type of pain and highlight the need for extreme caution against overgeneralizing about what the results of automated analysis show, particularly as the field moves inexorably toward implementation in clinical settings.

Fully Automatic Multimodal Pain Assessment

By far, most efforts at automatic analysis of pain have focused on the face. However, pain produces multiple behavioral responses (e.g., facial expressions, head and body movements, vocalizations) both within and across modalities. Various observational systems have been developed for quantifying other behaviors indicative of pain. Some are generic and can be applied or adapted to different types of pain [e.g., (83, 84)]; others have been developed for specific purposes or populations [e.g., the Pain Assessment Checklist for Seniors with Limited Ability to Communicate; (85); the Pain Assessment in Advanced Dementia scale; (86)]. In physical medicine and rehabilitation, body language is an important behavioral index of pain in patients with moderate to severe cognitive impairments, and those who have difficulty communicating verbally (87). Non-verbal (e.g., screaming, sounds of distress) and verbal (e.g., "ouch," "owie") pain vocalizations have proven clinically useful for pain detection in young children and others with limited linguistic abilities (88). There is strong likelihood that automatic analysis of acoustic characteristics of vocal expression can contribute to pain detection and understanding.

There is a nascent literature that has begun to apply the methods of machine learning to these other behavioral indicators of pain [e.g., (4)]. Efforts are needed to extend CVML technologies sensing beyond facial expression to include body and head movement, physiological measures, speech, and paralinguistic communication related to pain experience.

Automatic multimodal measurement affords potentially rich sets of behavioral features to include in automatic measurement of the occurrence and intensity of pain. Newer databases that include multimodal measures, such as EmoPain and BioVid make this development possible. Efforts in this direction will enable the objective measurement and monitoring of pain intensity in clinical, family, and work environments (2).

Links to Concepts of Expression in Pain

For all its technological sophistication, there is a kind of dustbowl empiricism about the corpus of work on automated analysis of pain. Although it builds on prior knowledge and findings–in particular the literature applying fine-grained behavioral analysis to the characterization of expression in pain–for the most part it has not addressed conceptual issues related to its meaning. Behavioral studies suggest that there is considerable complexity in the facial behavior that accompanies pain. Kunz and Lautenbacher (73), for example, provide evidence that the actions that most consistently relate to pain in the literature occur in separable clusters. This is an issue that has not been addressed in the _automated_ assessment literature. Moreover, there is good reason to believe that not all the expression that happens in pain is about pain. For example, the action of zygomaticus major (AU 12 in FACS), which is also the principal movement in a smile, is sometimes found to accompany pain, both in the behavioral literature (89) and the automatic analysis literature. Structural and functional analyses of this action suggest that, although it often does accompany pain, it is likely marking a different process (41, 90). CVML models to date do not seem to have recognized this distinction yet may have analytic potential to advance its understanding. Similarly, there is evidence that different components of the behaviors that correlate with pain are encoding different dimensions of the experience. Kunz et al. (91) found that actions involving movement around the eyes related most closely to sensory features of pain, while movements of the brows and upper lip related most closely to affective features. CVML studies have not addressed such issues to date but could be important in advancing our understanding of them.

Commercial and Other Applications

Commercial tools for pain assessment informed by the existing literature on automated assessment have already been developed and marketed and there is every reason to believe that this trend will continue. For example, Painchek (www.painchek.com) is a smartphone app-based device that combines a facial expression assessment component with input from five other domains (voice, movement, behavior, activity, body) to yield a pain score for application in geriatric and pediatric settings (92). It goes without saying that the development and marketing of tools for clinical assessment should be based on knowledge about automated assessment that is grounded in the empirical literature, consistent with the best-established technological solutions, has been subjected to rigorous validation procedures, and informed by understanding of issues of bias raised above. Importantly, commercial applications must be cognizant of the risks attendant on oversimplified interpretation of the meaning of a pain score derived from automated analysis of the face. An oft-stated rationale for focusing on facial and other behavioral indicators of pain is to improve pain management by improving pain detection. There is a substantial literature, however, showing that observers underestimate behavioral evidence of pain (93). This underestimation bias is paradoxical given that significant proportions of subjects in empirical studies show no behavioral evidence of pain (94). Facial expressions of pain have been characterized as a "late signaling system" (95), which implies that, if facial evidence of pain is present, it is likely very significant and needs to be taken seriously. Conversely, if it is not present, the possibility of its significance should not be discounted, a risk that is present with oversimplified interpretation of pain scores, however rendered.

A related concern arises from what appears to be widespread interest in the idea of pain simulation and empirical work implying that genuine pain can be distinguished from dissimulated pain. The idea lends itself to considerations that there may be forensic applications of automated assessment technology. It is true that perceptual (52), behavioral (96), and now automated assessment studies (36) have shown evidence that facial expressions during genuine and simulated pain have certain identifiable differences; however, the differences that have been documented have occurred under highly artificial conditions and appear, for the most part, to be small. Foreseeable application of forensic products based on automatic analysis appear open to abuse and unlikely to be probative.

WHITHER AUTOMATED ASSESSMENT OF PAIN?

That automated analysis of pain may be feasible has been demonstrated in the proof-of-concept studies reviewed above. The numerous studies that make up the corpus of the field since then have mainly added to the field by exploring alternative artificial intelligence systems. Ultimately, the value of this work is most likely to be realized in basic science and clinical research. In particular, the prospect of a form of assessment that can automatically yield reliable, valid and continuous information about how and when people (and animals) are expressing pain holds promise to enable detailed studies of pain modulation that are prohibitively difficult to perform with human observers who are subject to inherent limitations in their ability to resolve changes in behavior that sometimes occur in milliseconds, fatigue, and error. This could include evaluations of the time-course of pain reducing or augmenting influences but it could also extend to studies of how intrapersonal variables and the interpersonal, social, and environmental context influence pain over momentary differences in time. There is evidence from extant studies that automated detection techniques can give insight into momentary changes at or near the level of a frame of video [(51); see **Figure 1**]. In principle, valid measurement at that level of sensitivity could yield important information about dose-response relationships in evaluations of analgesic medications. A system that combined automated detection of pain with detection of other affective states (e.g., anger) and also permitted time-series analysis could facilitate greater understanding of the interplay of the states. To date, no attention has been applied to how automatic pain detection may vary between men and women, people of different racial and ethnic backgrounds, or context, to name just a few factors. Of particular interest in would be studies of pain expression in interactions in health-care settings or in families.

The work performed to date for automated pain measurement has been interesting, progress has been rapid and has generated the kind of buzz commonly associated with new technologies. But numerous current controversies over unforeseen consequences about how these new algorithms have been developed (for example, errors that have been "baked in" to the data on which facial recognition systems were trained, leading to wrongful arrest), or how they work highlight the need to proceed cautiously, mindful that "move fast and break things" is not a slogan that augurs well for the careful and safe development of a tool to advance understanding of pain in particular and other health related applications in general. The existing approaches are built on a very limited sample of participants, pain types, annotation procedures, conditions of observation, ages, "racial"/ethnic categories, and regions of the world. Careful expansion of audiovisual pain databases that sample more broadly and representatively across these dimensions will be necessary to establish confidence in the quality and meaning of the measurement obtained and to manage foreseeable and unforeseeable perils of using this technology to improve patients' outcomes. Particular concern arises around the prospect of developing and commercializing technologies geared to clinical, medico-legal, and forensic applications, especially around the idea of proprietary knowledge. Practical applications of automatic pain assessment need to be based on rigorous science that meets standards of professional peer review and public accountability, including verification that the CVML processes on which they are based validly produce assessments that are consistent with the claims being made of them.

AUTHOR CONTRIBUTIONS

KP and ZH contributed to conceptualization and writing of this paper. Both authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the National Institutes of Health under Award Number R01NR018451. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ACKNOWLEDGMENTS

Development of the UNBC-McMaster Shoulder Pain Expression Archive Database was supported by grant MOP 53301 from the Canadian Institutes of Health Research.

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) 161:1976–82.
2. Hammal Z, Cohn JF. Automatic, objective, and efficient measurement of pain using automated face analysis. In: Vervoort T, Karos K, Trost Z, Prkachin KM, editors. *Social and Interpersonal Dynamics in Pain: We Don't Suffer Alone*. New York, NY: Springer (2018). p. 121–48. doi: 10.1007/978-3-319-78340-6_7
3. Werner P, Lopez-Martinez D, Walter S, Al-Hamadi A, Gruss S, Picard RW. Automatic recognition methods supporting pain assessment: a survey. In: *IEEE Transactions on Affective Computing*. (2019). doi: 10.1109/TAFFC.2019.2946774
4. Aung MSH, Kaltwang S, Romera-Paredes B, Martinez B, Singh A, Cella M, et al. The automatic detection of chronic pain-related expression: requirements, challenges and a multimodal dataset. *IEEE Trans Affect Comput*. (2015) 7:435–51. doi: 10.1109/TAFFC.2015.2462830
5. Egede JO, Song S, Olugbade TA, Wang C, Williams A, Meng H, et al. EMOPAIN challenge 2020: multimodal pain evaluation from facial and bodily expressions. *arXiv*. (2020) arXiv.2001.07739v07733. doi: 10.1109/FG47880.2020.00078
6. Gracely RH, McGrath P, Dubner R. Ratio scales of sensory and affective verbal pain descriptors. *Pain*. (1978) 5:5–18. doi: 10.1016/0304-3959(78)90020-9

7. Gracely RH, Wolskee PJ. Semantic functional measurement of pain: integrating perception and language. *Pain.* (1983) 15:389–98. doi: 10.1016/0304-3959(83)90074-X
8. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain.* (1983) 17:45–56. doi: 10.1016/0304-3959(83)90126-4
9. Williams ACC, Davies HTO, Chadury Y. Simple pain rating scales hide complex, idiosyncratic meanings. *Pain.* (2000) 85:487–63. doi: 10.1016/S0304-3959(99)00299-7
10. Beecher HK. *Measurement of Subjective Responses: Quantitative Effects of Drugs*. London: Oxford University Press (1959).
11. Craig KD, Weiss SM. Vicarious influences on pain threshold determinations. *J Pers Soc Psychol.* (1971) 19:53–9. doi: 10.1037/h0031097
12. Hilgard ER, Hilgard JR, Macdonald H, Morgan AH, Johnson LS. Covert pain in hypnotic analgesia: its reality as tested by the real-simulator design. *J Abnorm Psychol.* (1978) 87:655–63. doi: 10.1037/0021-843X.87.6.655
13. Spanos NP, Hewitt EC. The hidden observer in hypnotic analgesia: discovery or experimental creation. *J Pers Sock Psychol.* (1980) 39:1201–14. doi: 10.1037/h0077730
14. Mogil J. Sex differences in pain and pain inhibition: multiple explanations of a controversial phenomenon. *Nat Rev Neurosci.* (2012) 13:859–66. doi: 10.1038/nrn3360
15. Tursky B. Physical, physiological, and psychological factors that affect pain reaction to electric shock. *Psychophysiology.* (1974) 11:95–112. doi: 10.1111/j.1469-8986.1974.tb00830.x
16. Worley A, Fabrizi L, Boyd S, Slater R. Multi-modal pain measurements in infants. *J Neurosci Methods.* (2012) 205:252–7. doi: 10.1016/j.jneumeth.2012.01.009
17. Koenig J, Jarczok MN, Ellis RJ, Hillecke TK, Thayer JF. Heart rate variability and experimentally induced pain in healthy adults: a systematic review. *Euro J Pain.* (2014) 18:301–14. doi: 10.1002/j.1532-2149.2013.00379.x
18. Zaslansky R, Sprecher E, Katz Y, Rozenberg B, Hemli JA, Yarnitsky D. Pain-evoked potentials: what do they really measure. *Electroencephalogr Clin Neurophysiol.* (1996) 100:384–92. doi: 10.1016/0168-5597(96)95654-3
19. Robinson ME, Staud R, Price DD. Pain measurement and brain activity: will neuroimages replace pain ratings? *J Pain.* (2013) 14:323–7. doi: 10.1016/j.jpain.2012.05.007
20. Fordyce WE. *Behavioral Methods for Chronic Pain and Illness*. St. Louis, MO: Mosby (1976).
21. Darwin C. *The Expression of the Emotions in Man and Animals*. London: John Murray (1872). doi: 10.1037/10001-000
22. Ekman P, Sorenson ER, Friesen WV. Pan-cultural elements in facial displays of emotion. *Science*. (1969) 164:86–8.
23. Kunz MM, Meixner D, Lautenbacher S. Facial muscle movements encoding pain—a systematic review. *Pain.* (2019) 160:535–49. doi: 10.1097/j.pain.0000000000001424
24. Grunau RVE, Craig KD. Pain expression in neonates: facial action and cry. *Pain.* (1987) 28:395–410. doi: 10.1016/0304-3959(87)90073-X
25. Chambers CT, Cassidy KL, McGrath PJ, Gilbert CA, Craig, KD. *Child Facial Coding System: A Manual*. Halifax: Dalhousie University and University of British Columbia (1996).
26. Prkachin KM. Assessing pain by facial expression: facial expression as nexus. *Pain Res Manage.* (2009) 14:53–8. doi: 10.1155/2009/542964
27. Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, et al. Coding of facial expressions of pain in the laboratory mouse. *Nat Methods.* (2010) 7:447–9. doi: 10.1038/nmeth.1455
28. Rash JA, Prkachin KM, Solomon PE, Campbell TS. Assessing the efficacy of a manual-based intervention for improving the detection of facial pain expression. *Euro J Pain.* (2019) 23:1006–19. doi: 10.1002/ejp.1369
29. Picard RW. *Affective Computing*. Cambridge: MIT Press (2007).
30. Picard RW. Affective computing: from laughter to IEEE. *IEEE Trans Affect Comput.* (2010) 1:11–7. doi: 10.1109/T-AFFC.2010.10
31. Bartlett MS, Hager JC, Ekman P, Sejnowski TJ. Measuring facial expressions by computer image analysis. *Psychophysiology.* (1999) 36:253–63. doi: 10.1017/S0048577299971664
32. Cohn JF, Zlochower AJ, Lien J, Kanade T. Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. *Psychophysiology.* (1999) 36:35–43. doi: 10.1017/S0048577299971184
33. Ambadar Z, Cohn JF, Reed LI. All smiles are not created equal: morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *J Nonverbal Behav.* (2009) 33:17–34. doi: 10.1007/s10919-008-0059-5
34. Girard JM, Cohn JF, Jeni LA, Lucey S, De la Torre F. How much training data for facial action unit detection? In: *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition.* Ljubljana (2015). doi: 10.1109/FG.2015.7163106
35. Bartlett MS, Littlewort GC, Frank MG, Lainscsek C, Fasel I, Movellan J. Recognizing facial expression: machine learning and application to spontaneous behavior. In: *IEEE International Conference on Computer Vision and Pattern Recognition*. (2005). p. 568–73. doi: 10.1109/CVPR.2005.297
36. Littlewort GC, Bartlett MS, Lee K. Automatic coding of facial expressions displayed during posed and genuine pain. *Image Vis Comput.* (2009) 27:1797–803. doi: 10.1016/j.imavis.2008.12.010
37. Sikka K, Ahmed AA, Diaz D, Goodwin MS, Craig KD, Bartlett M, et al. Automated assessment of children's postoperative pain using computer vision. *Pediatrics.* (2015) 136:e124–31. doi: 10.1542/peds.2015-0029
38. Hammal Z, Kunz M, Arguin M, Gosselin F. Spontaneous pain expression recognition in video sequences. In: *Proceedings of the BCS International Conference on Visions of Computer Science*. London: BCS Vision (2008).
39. Hammal Z, Kunz M. Pain monitoring: a dynamic and context-sensitive system. *Pattern Recognit*. (2012) 45:1265–80.
40. Lucey P, Cohn JF, Prkachin KM, Solomon PE, Matthews I. Painful data: the UNBC-McMaster shoulder pain expression archive database. In: *2011 IEEE International Conference on Automatic Face and Gesture Recognition*. (2011). doi: 10.1109/FG.2011.5771462
41. Prkachin KM, Solomon PE. The structure, reliability and validity of pain expression: evidence from patients with shoulder pain. *Pain.* (2008) 139:267–74. doi: 10.1016/j.pain.2008.04.010
42. Ekman P, Friesen WV. *The Facial Action Coding System*. Palo Alto, CA: ThConsulting Psychologists Press (1978). doi: 10.1037/t27734-000
43. Kachele M, Thiam P, Amirian M, Werner P, Walter S, Schwenker F, et al. Multimodal data fusion for person-independent, continuous estimation of pain intensity. In: Iliadis L, Jayne C, editors. *Engineering Applications of Neural Networks EANN 2015. Communications in Computer and Information Science, Vol 517*. Cham: Springer (2015). doi: 10.1007/978-3-319-23983-5_26
44. Kachele M, Amirian M, Thiam P, Werner P, Walter S, Palm G, et al. Adaptive confidence learning for the personalization of pain intensity estimation systems. *Evolv Syst.* (2017) 8:71–83. doi: 10.1007/s12530-016-9158-4
45. Martinez DL, Rudovic O, Picard R. Personalized automatic estimation of self-reported pain intensity from facial expressions. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Workshops*. (2017). p. 318–27. doi: 10.1109/CVPRW.2017.286
46. Lopez-Martinez D, Picard R. Continuous pain intensity estimation from autonomic signals with recurrent neural networks. In: *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2018*. (2018). p. 5624–7. doi: 10.1109/EMBC.2018.8513575
47. Werner P, Al-Hamadi A, Niese R, Walter S, Gruss S, Traue HC. Automatic pain recognition from video and biomedical signals. *Int Conf Pattern Recognit*. (2014) 2014:4582–7. doi: 10.1109/ICPR.2014.784
48. Werner P, Al-Hamadi A, Limbrecht-Ecklundt K, Walter S, Gruss S, Traue HC. Automatic pain assessment with facial activity descriptors. *IEEE Trans Affect Comput*. (2017) 8:286–99. doi: 10.1109/TAFFC.2016.2537327
49. Yang R, Tong S, Bordallo M, Boutellaa E, Peng J, Feng X, et al. On pain assessment from facial videos using spatio-temporal local descriptors. In: *Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. (2016). p. 1–6. doi: 10.1109/IPTA.2016.7820930
50. Walter S, Gruss S, Ehleiter H, Tan J, Traue H, Werner P, et al. The BioVid heat pain database: data for the advancement and systematic validation of an automated pain recognition system. In: *2013 IEEE International Conference on Cybernetics (CYBCO)*. (2013). p. 128–31. doi: 10.1109/CYBConf.2013.6617456

51. Ashraf AB, Lucey S, Cohn JF, Chen T, Ambadar Z, Prkachin KM, et al. The painful face—pain expression recognition using active appearance models. *Image Vis Comput.* (2009) 27:1788–96. doi: 10.1016/j.imavis.2009.05.007
52. Prkachin KM. The consistency of facial expressions of pain: a comparison across modalities. *Pain.* (1992) 51:297–306. doi: 10.1016/0304-3959(92)90213-U
53. Lucey P, Cohn JF, Matthews I, Lucey S, Sridharan S, Howlett J, et al. Automatically detecting pain in video through facial action units. *IEEE Trans Syst Man Cybernet Part B Cybernet.* (2011) 41:664–74. doi: 10.1109/TSMCB.2010.2082525
54. Lucey P, Cohn JF, Solomon P, Matthews I. Painful data: the UNBC-McMaster shoulder pain expression archive database. *Image Vis Comput J.* (2012) 30:197–205. doi: 10.1016/j.imavis.2011.12.003
55. Hammal Z, Kunz MM. Pain monitoring: a dynamic and context-sensitive system. *Pattern Recognit.* (2012) 45:1265–80. doi: 10.1016/j.patcog.2011.09.014
56. Hammal Z, Cohn JF. Automatic detection of pain intensity. *Int Conf Multimodal Interact*. (2012) 2012:47–52. doi: 10.1145/2388676.2388688
57. Erekat D, Hammal Z, Siddiqui M, Dibeklioglu H. Enforcing multilabel consistency for automatic spatio-temporal assessment of shoulder pain intensity. In: *Adjunct Proceedings of the ACM International Conference on Multimodal Interaction (ICMI), Workshop Face and Gesture Analysis for Health Informatics (FGAHI)*. Utrecht (2020). doi: 10.1145/3395035.3425190
58. Kaltwang S, Rudovic O, Pantic M. Continuous pain intensity estimation from facial expressions. In: Bebis G, editor. *Advances in Visual Computing.* Heidelberg: Springer (2012). p. 368–77. doi: 10.1007/978-3-642-33191-6_36
59. Liu D, Peng F, Rudovic O, Picard RW. DeepFaceLIFT: interpretable personalized models for automatic estimation of self-reported pain. In: *Proceedings of the 1st IJCAI Workshop on Artificial Intelligence in Affective Computing (Proceedings of Machine Learning Research).* Melbourne, VIC (2017). p. 1–16.
60. Rudovic O, Pavlovic V, Pantic M. Automatic pain intensity estimation with heteroscedastic conditional random fields. In: Bebis G, editor. *International Symposium on Visual Computing, 2013: Part II*. (2013). p. 234–43. doi: 10.1007/978-3-642-41939-3_23
61. Szczapa B, Daoudi M, Berretti S, Pala P, Del Bimbo A, Hammal Z. Automatic estimation of self-reported pain by interpretable representations of motion dynamics. In: *Proceedings of 25th International Conference on Pattern Recognition (ICPR).* Milan (2020).
62. Sikka K, Dhall A, Bartlett M. Weakly supervised pain localization and classification with multiple segment learning. *Image Vis Comput*. (2014) 32:659–70.
63. Craig KD, Prkachin KM, Grunau RVE. *The Facial Expression of Pain*. In: Turk DC, Melzack R, editors. *Handbook of Pain Assessment*. 3rd edn. New York: Guilford (2011).
64. Hadjistovropoulos T, Herr K, Prkachin KM, Craig KD, Gibson SJ, Lukas A, et al. Pain assessment in elderly adults with dementia. *Lancet Neurol.* (2014) 13:1216–27. doi: 10.1016/S1474-4422(14)70103-6
65. Zamzmi G, Kasturi R, Goldgof D, Zhi R, Ashmeade T, Sun Y. A review of automated pain assessment in infants: features, classification tasks, and databases. *IEEE Rev Biomed Eng.* (2019) 11:77–96. doi: 10.1109/RBME.2017.2777907
66. Brahnam S, Chuang CF, Shih FY, Slack MR. Machine recognition and representation of neonatal facial displays of acute pain. *Int J Artif Intell Med.* (2006) 36:211–22. doi: 10.1016/j.artmed.2004.12.003
67. Zamzmi G, Ruiz G, Goldgof D, Kasturi R, Sun Y, Ashmeade T. Pain assessment in infants: towards spotting pain-expression based on infants' facial strain. In: *Proceedings of the 11th IEEE International Conference. Workshops on Automatic Face and Gesture Recognition*. (2015). p. 1–5. doi: 10.1109/FG.2015.7284857
68. Littlewort G, Whitehill J, Wu T, Fasel I, Frank M, Movellan J, et al. The computer expression recognition toolbox (CERT). In: *2011 International Conference on Automatic Face and Gesture Recognition (FG)*. (2011). p. 298–305. doi: 10.1109/FG.2011.5771414
69. Kunz MM, Scharmann S, Hemmeter U, Schepelmann K, Lautenbacher S. The facial expression of pain in patients with dementia. *Pain.* (2007) 126:64–71. doi: 10.1016/j.pain.2007.09.007
70. Rezaei S, Zhao. S, Moturu I, Prkachin KM, Hadjistavropoulos T, et al. Unobtrusive pain monitoring in older adults with dementia using pairwise and contrastive training. *IEEE J Biomed Health Informat.* (2021) 25:1450–62. doi: 10.1109/JBHI.2020.3045743
71. Rosenthal R. Conducting judgment studies. In: Harrigan JA, Rosenthal R, Scherer KR, editors. *New Handbook of Methods in Nonverbal Behavior Research*. New York, NY: Oxford University Press (2008). p. 199–237. doi: 10.1093/acprof:oso/9780198529620.003.0005
72. Hammal Z, Cohn JF. Automatic, objective, and efficient measurement of pain using automated face analysis. In: Vervoort T, Karos K, Trost Z, Prkachin KM, editors. *Social And Interpersonal Dynamics In Pain: We Don't Suffer Alone*. Cham: Springer (2008). p.113–20.
73. Kunz MM, Lautenbacher S. The faces of pain: a cluster analysis of individual differences in facial activity patterns of pain. *Euro J Pain.* (2014) 18:813–23. doi: 10.1002/j.1532-2149.2013.00421.x
74. Kunz MM, Prkachin KM, Solomon PE, Lautenbacher S. Faces of clinical pain: interindividual facial activity patterns in shoulder pain patients. *Euro J Pain.* (2020) 25:529–40. doi: 10.1002/ejp.1691
75. Green CR, Anderson KO, Baker TA, Campbell LC, Decker S, Fillingim RB, et al. The unequal burden of pain: confronting racial and ethnic disparities in pain. *Pain Medicine.* (2003) 4:277–94 doi: 10.1046/j.1526-4637.2003.03034.x
76. Wandner LD, Stutts LA, Alquday AF, Craggs JG, Scipio CD, Hirsh AT, et al. Virtual human technology: patient demographics and healthcare training factors in pain observation and treatment recommendations. *J Pain Res.* (2010) 3:241–7. doi: 10.2147/JPR.S14708
77. Wandner LD, Scipio CD, Hirsh AT, Torres CA, Robinson ME. The perception of pain in others: how gender, race, and age influence pain expectations. *J Pain.* (2012) 13:220–7. doi: 10.1016/j.jpain.2011.10.014
78. Drwecki BB. Race and pain: a dual injustice. In: Vervoort T, Karos K, Trost Z, Prkachin KM, editors. *Social and Interpersonal Dynamics in Pain: We Don't Suffer Alone*. New York, NY: Springer (2018). p. 455–80. doi: 10.1007/978-3-319-78340-6_21
79. Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley JL. Sex, gender, and pain: a review of recent clinical and experimental findings. *J Pain.* (2009) 10:447–85. doi: 10.1016/j.jpain.2008.12.001
80. Buolamwini J, Gebru T. Gender shades: intersectional accuracy disparities in commercial gender classification. *PMLR*. (2018) 81:1–15.
81. Pasquali F, Malgieri G. *If You Don't Trust AI Yet, You're Not Wrong*. New York, NY: Times Opinion Section (2021).
82. Taati B, Zhao S, Ashraf AB, Asgarian A, Browne ME, Prkachin KM, et al. *Algorithmic Bias in Clinical Populations – Evaluating and Improving Facial Analysis Technology in Older Adults With Dementia*. IEEE Access (2019). doi: 10.1109/ACCESS.2019.2900022
83. Keefe FJ, Somers TJ, Williams DA, Smith SJ. Assessment of pain behaviors. In: Turk DC, Melzack R, editors. *Handbook of Pain Assessment*. 3rd edn. New York, NY: Guilford (2011). p. 134–50.
84. Prkachin KM, Hughes EA, Schultz I, Joy P, Hunt D. Real-time assessment of pain behavior during clinical assessment of low-back pain patients. *Pain.* (2002) 95:23–30. doi: 10.1016/S0304-3959(01)00369-4
85. Fuchs-Lacelle S, Hadjistavropoulos T. Development and preliminary validation of the pain assessment checklist for seniors with limited ability to communicate (PACSLAC). *Pain Manage Nurs.* (2004) 5:37–49. doi: 10.1016/j.pmn.2003.10.001
86. Warden V, Hurley AC, Volicer L. Development and psychometric evaluation of the pain assessment in advanced dementia (PAINAD) scale. *J Am Med Direct Assoc.* (2003) 4:9–15. doi: 10.1097/01.JAM.0000043422.31640.F7
87. De Knegt NC, Pieper MJC, Lobbezoo F, Schuengel C, Evenhuis HM, Passchier J, et al. Behavioral pain indicators in people with intellectural disabilities: a systematic review. *J Pain.* (2013) 14:885–96. doi: 10.1016/j.jpain.2013.04.016
88. DuBois A, Bringuier S, Capdevilla X, Pry R. Vocal and verbal expression of postoperative pain in pre-schoolers. *Pain Manage Nurs.* (2008) 9:160–5. doi: 10.1016/j.pmn.2007.10.003

89. Kunz MM, Prkachin K, Lautenbacher S. The smile of pain. *Pain.* (2009) 145:273–5. doi: 10.1016/j.pain.2009.04.009
90. Kunz MM, Prkachin KM., Lautenbacher S. Smiling in pain: explorations of its social motives. *Pain Res Treat*. (2013) 2013:1–8. doi: 10.1155/2013/128093
91. Kunz MM, Lautenbacher S, LeBlanc N, Rainville P. Are both the sensory and the affective dimensions of pain encoded in the face? *Pain.* (2012) 153:350–8. doi: 10.1016/j.pain.2011.10.027
92. Atee M, Hoti K, Hughes JD. A technical note on the painchek system: a web portal and mobile medical device for assessing pain in people with dementia. *Front Aging Neurosci.* (2018) 10:117.
93. Prkachin KM, Solomon PA, Ross AJ. The underestimation of pain among health-care providers. *Can J Nurs Res.* (2007) 39:88–106.
94. Kunz MM, Karos K, Vervoort T. When, how, and why do we express pain. In: Vervoort T, Karos K, Trost Z, Prkachin KM, editors. *Social and Interpersonal Dynamics in Pain: We Don't Suffer Alone.* New York, NY: Springer (2018). p. 101–20. doi: 10.1007/978-3-319-78340-6_6
95. Prkachin KM. Facial pain expression. *Pain Manage.* (2011) 1:367–76. doi: 10.2217/pmt.11.22
96. Craig KD, Hyde SA, Patrick CJ. Genuine, suppressed, and faked facial behavior during exacerbation of chronic low back pain. *Pain.* (1991) 46:161–71. doi: 10.1016/0304-3959(91)90071-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Pain Research

CORRECTION
published: 25 February 2022
doi: 10.3389/fpain.2022.849950

Corrigendum: Computer Mediated Automatic Detection of Pain-Related Behavior: Prospect, Progress, Perils

Kenneth M. Prkachin[1]* *and Zakia Hammal*[2]

[1] *Department of Psychology, University of Northern British Columbia, Prince George, BC, Canada,* [2] *The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, United States*

Keywords: pain, measurement, facial expression, automation, assessment

A Corrigendum on

Computer Mediated Automatic Detection of Pain-Related Behavior: Prospect, Progress, Perils
by Prkachin, K. M., and Hammal, Z. (2021). Front. Pain Res. 2:788606. doi: 10.3389/fpain.2021.788606

In the original article, the subheading **Assessment Based on Fine-Grained Facial Observation** should have been deleted. The correct subheading is: **Pain Assessment Based on Fine-Grained Facial Observation**.

In the original article, the subsection titled **Limitations of Physiological Assessment of Pain** appeared in an illogical place, immediately after a new section and topic, **Pain Assessment Based on Fine-Grained Facial Observation**, is introduced. The correct location of the subsection is directly after the section titled **Physiological Assessment of Pain**.

In the original article, the authors were inconsistent in the labeling used to refer to the UNBC-McMaster database. Because of how it was designated in the original article describing it, how central it is to the field, and so that researchers in the field can cite and track work based on it accurately, the term "UNBC-McMaster" has been corrected to be referred to uniformly as the "UNBC-McMaster Shoulder Pain Expression Archive Database."

In several places in the original article, where the UNBC-McMaster Shoulder Pain Expression Archive Database is described, the sentence construction was awkward for lack of the preceding definite article, "the". This has been corrected on pages 6, 7, and 8.

In the original article, there was an error in the **Funding** statement. By official US National Institutes of Health policy, the **Author Disclaimer** statement must follow the funding acknowledgement. The **Author Disclaimer** has been moved to the **Funding** statement.

In the original article, acknowledgment of support was inaccurately not contained in an **Acknowledgments** statement. The acknowledgment of Canadian Institutes of Health Research funding gives the correct grant information and does not require the disclaimer.

The corrected **Funding** statement and **Acknowledgments** statement appear below.

FUNDING

This work was supported in part by the National Institutes of Health under Award Number R01NR018451. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

OPEN ACCESS

Approved by:
Frontiers Editorial Office, Frontiers Media SA, Switzerland

***Correspondence:**
Kenneth M. Prkachin
kmprk@unbc.ca

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *06 January 2022*
Accepted: *07 January 2022*
Published: *25 February 2022*

Citation:
Prkachin KM and Hammal Z (2022) Corrigendum: Computer Mediated Automatic Detection of Pain-Related Behavior: Prospect, Progress, Perils. Front. Pain Res. 3:849950. doi: 10.3389/fpain.2022.849950

ACKNOWLEDGMENTS

Development of the UNBC-McMaster Shoulder Pain Expression Archive Database was supported by grant MOP 53301 from the Canadian Institutes of Health Research.

The authors apologize for these errors and state that they do not change the scientific conclusions of the article in any way. The original article has been updated.

frontiers in Pain Research

ORIGINAL RESEARCH
published: 23 December 2021
doi: 10.3389/fpain.2021.729860

The Influence of Examiner Gender on Responses to Tonic Heat Pain Assessments: A Preliminary Investigation

Jessica F. McDougall[1,2], *Nicole G. N. Bailey*[1], *Rohan Banga*[1], *Lukas D. Linde*[1,3,4*†] and *John L. K. Kramer*[1,3,4†]

[1] International Collaboration on Repair Discoveries, University of British Columbia, Vancouver, BC, Canada, [2] Department of Rehabilitation Sciences, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada, [3] Department of Anesthesiology, Pharmacology and Therapeutics, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada, [4] Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, BC, Canada

Background: The influence of examiner gender on pain reporting has been previously explored in both research and clinical settings. However, previous investigations have been limited, with the majority of studies employing single, static assessments of pain (e.g., cold pressor test, verbal pain ratings). The impact of examiner gender on both static and dynamic heat-based pain assessments is currently unknown.

Methods: Thirty eight participants (20 females aged 24.1 ± 4.44, and 18 males, aged 24.8 ± 4.54) completed two identical testing sessions, randomized to a male and female examiner in a cross-over design. Pain sensitivity was examined using heat pain thresholds, verbal pain ratings to tonic heat, computerized visual analog scale (CoVAS) rating to tonic heat, and participant-controlled temperature (PCT) heat pain assessments.

Results: Female participants reported higher verbal pain to tonic heat with a female examiner compared to male participants, with similar trends for CoVAS responses to tonic heat. Conversely heat pain thresholds and PCT were not significantly influenced by experimenter gender.

Conclusions: Overall, verbal ratings were the most impacted by examiner gender, with temperature-based methods such as PCT and pain thresholds showing little to no examiner gender effects. While the gender of the examiner may be an important consideration in the measurement of sex and gender differences in pain research, the choice of pain assessment method may be of similar consequence.

Keywords: quantitative sensory testing, sex differences, gender differences, participant-controlled temperature, thermal pain

OPEN ACCESS

Edited by:
Loren Martin,
University of Toronto, Canada

Reviewed by:
Wiebke Gandhi,
University of Reading, United Kingdom
Kai Karos,
Open University of the Netherlands, Netherlands

****Correspondence:***
Lukas D. Linde
lukas.linde@ubc.ca

†These authors share senior authorship

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *23 June 2021*
Accepted: *06 December 2021*
Published: *23 December 2021*

Citation:
McDougall JF, Bailey NGN, Banga R, Linde LD and Kramer JLK (2021) The Influence of Examiner Gender on Responses to Tonic Heat Pain Assessments: A Preliminary Investigation. Front. Pain Res. 2:729860. doi: 10.3389/fpain.2021.729860

INTRODUCTION

The role of sex and gender on pain has been the source of substantial scientific and public discourse (1–5). In clinical settings, females experience acute and chronic pain with more frequency and to a greater intensity compared to males (2). Experimental studies employing pain sensitivity quantitative sensory testing [QST; a

battery of tests which examines noxious and non-noxious somatosensory sensitivity (6)] outcomes (e.g., cold pressor tests, pain pressure thresholds) have provided complimentary support that females may be more sensitive to noxious stimuli than males (2, 4). Heterogeneity among these QST outcomes are commonplace, however, challenging the notion of the aforementioned robust sex or gender-related differences in pain perception (1).

A number of social factors have been proposed to contribute to variation in QST outcomes between experimental pain studies. These include individual and interpersonal factors, as well as environmental factors such as time of day (7, 8). Related to interpersonal factors, the social, gender context of the pain experience appears to influence pain report. Opposing examiner gender effects have been reported, with male participants tending to verbally report significantly less pain in the presence of a female examiner (and vice versa for female participants) (7, 9, 10) [note: gender is used in this regard given that these effects are *social* as opposed to *biological* (5)]. This follows the Gender Context Model of Pain, which suggests men will be less likely to express pain, especially if the examiner is perceived as being threatening to masculine gender roles, whereas women will be more likely to express pain. However, this difference may be dependent on how pain is expressed. Verbal pain report may be more susceptible to these gender differences than non-verbal expressions (11, 12). Indeed, individual factors add complexity. One possible explanation for these reported gender specific examiner effects may be differences in catastrophizing—a negative cognitive-affective response to pain (13). Catastrophizing is associated with increased pain across a variety of pain measures and may be influenced by the presence of others (4, 13). Moreover, sex differences in catastrophizing have been reported, insofar as women tend to catastrophize more than men (4). As such, catastrophizing may also modulate the interaction between sex and social interaction of pain measurements.

A major limitation of previous experimenter/participant gender investigations has been a narrow focus on pain tolerance, measured chiefly by way of the cold pressor test (7). Advances in QST techniques have led to the development of various static and dynamic outcomes, which have been widely employed to investigate sex/gender differences in pain perception (14). Painful thermal dynamic and static QST measures have shown significant differences between male and female participants (4) and may be differently susceptible to experimenter gender influence, and to gender stereotypes. For example, verbal pain ratings of heat pain involve direct verbal communication with experimenters in response to a noxious stimulus, conversely, automated metrics of pain assessment, such as participant controlled temperature (PCT) (15), require less direct communication with examiners. Verbal pain report has been shown to be susceptible to the gender context in which the report occurs (7, 9, 10), however it is not known how susceptible PCT—a non-verbal form of pain expression—is to these gendered influences. It stands to reason that such differences in participant/experimenter interactions within QST assessments may influence the effect of experimenter gender on pain perception. Including both verbal and non-verbal pain reports to both a male and female examiner allows us to tease apart the impact of social context on the apparent sex/gender differences in pain. To our knowledge, no previous studies have explored the influence of experimenter gender on pain outcomes assessed using multiple painful heat QST techniques.

Our aim was to determine the extent to which modern QST heat-pain measures are influenced by the gender of the examiner. To this end, we employed verbal and non-verbal rating and temperature-based (non-verbal) methods of reporting sensitivity to heat pain, with both static and dynamic outcomes. A secondary, exploratory aim explored role of psychosocial factors, specifically the effect of pain catastrophizing on experimenter gender effects on pain outcomes. We expected to see greater gender differences in verbal ratings-based measurements of pain compared to temperature-based measurements, such that males would verbally rate pain as lower in the presence of a female examiner, and females would demonstrate opposite and smaller effect. We anticipate temperature-based methods to show smaller or non-significant effect, as these rely on less direct social interaction during pain reporting.

MATERIALS AND METHODS

Participants

We determined 40 participants (females aged 24.1 $\pm$ 4.44, and males, aged 24.8 $\pm$ 4.54) would provide a partial eta-squared (η_p^2) = 0.05, with a power of 0.8 and an alpha of 0.5 (calculation completed in G*Power 3.1) (16). This η_p^2 was estimated from previous studies that have compared the interaction of experimenter and participant gender on pain outcomes (17–19). Exclusion criteria included presence or history of chronic pain (i.e., pain persisting longer than 3 months), determined from a self-reported health history questionnaire. All participants were over 18 years of age and provided informed consent. Participants were recruited from the local university and hospital communities through flier advertisements.

Experimenters

The experimenters were a cis-female aged 22 and a cis-male aged 19. Both wore a lab coat over jeans and a shirt, and both identified as cis-gendered [i.e., indicated that their gender (man/woman) did not differ from their sex (male/female)]. We did not control for other experimenter characteristics (e.g., height, weight, or race), and these characteristics were not collected from participants. Scripts were created to standardize interactions with the participants, including instructions for all pain tests.

Procedure

Participants were randomly assigned to a male or female examiner on day 1 in a counterbalanced design, such that half of the participants began with the male examiner, while the other half began with the female examiner (**Figure 1**). Sessions were at least 24 h apart. Each testing day was designed to be approximately 1-h long. The true nature of the study was withheld from participants, who were led to believe that the purpose was to compare two measures of testing heat pain. Given

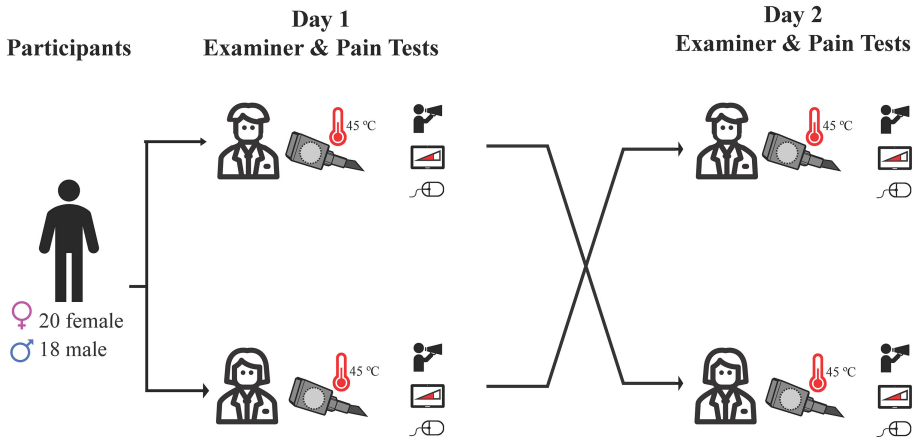


FIGURE 1 | Outline of study protocol. Participants were randomly assigned to a male or female examiner on day 1 in a counterbalanced design, and completed heat pain testing using three different methods: verbal pain rating, computerized visual analog scale, and participant-controlled temperature. Day 2 testing was identical, and was conducted by the opposite gender examiner.

the blinded nature of our study, the experimenter followed a script that introduced them as the research assistant for the study, with no mention of their gender or the true nature of the study. Experimenters stayed close beside the participant for all tests, standing beside the participant and alternating between watching a computer screen (where the test results were being shown), making an arbitrary note on a clipboard, and glancing at the participant to ensure protocols were being followed. The switch of experimenters was explained to participants as a "scheduling conflict," and the other experimenter was filling in due to the absence. At the end of day 2, participants were fully debriefed. This involved the experimenter outlining the need for deception and offering participants the opportunity to withdraw their data from the study. All participants were then asked if they suspected or knew the true purpose of the study. All study procedures were approved by the Behavioral Research Ethics Board at the University of British Columbia (approval number H19-00944), and were conducted in accordance with the Declaration of Helsinki (20) involving research on human participants. Our study protocol was not pre-registered, due the required deception of participants (i.e., pre-registering planned statistical comparisons could give away the true nature of the study).

Heat Pain Measurements

Heat pain thresholds and responses to prolonged heat pain were performed using a calibrated thermode (Medoc Advanced Medical Systems, Ramat Yishai, Israel, CHEPs thermode, 27 mm diameter) applied on the palmar aspect of the forearm. Heat pain thresholds were performed first on the distal 1/3 of a randomly chosen forearm, followed by either PCT or continuous visual analog scale (CoVAS) heat tests performed on the proximal 2/3 of the same forearm—the order of the PCT and CoVAS test presentation was randomized. Prior to the presentation of each heat test, a familiarization test took place to introduce participants to the pain-rating method. A 5-min break separated the three tests (heat pain thresholds, familiarization, and tonic heat test). Following the first tonic heat test, a 10-min break took place. Heat pain thresholds were then performed on the distal 1/3 of the other forearm, followed by the PCT or CoVAS, whichever was randomized to be performed second. Another familiarization test was performed prior to the introduction of the second tonic heat test. Again a 5-min break separated each of the three tests (heat pain thresholds, familiarization, and tonic heat test).

Heat Pain Thresholds

For heat pain thresholds, the thermode temperature was increased at a rate of 1°C/s from a baseline of 32°C to a maximum of 55°C. Participants were instructed to press a button when the first sensations of pain were perceived (i.e., when the original impression of warmth or heat turned into the feeling of "burning," "stinging," "aching," or "drilling") (6). Upon button press, the heat thermode returned to the baseline temperature of 32°C at a rate of 70°C/s. Four trials were conducted consecutively with at least 5 s between each trial. The main outcome measure from pain threshold assessments was the average temperature of the initial pain sensations over the four trials.

Tonic Heat Pain

Participants continuously rated their pain perception throughout a 2-min application of tonic heat (45°C) *via* CoVAS (Medoc Advanced Medical Systems, Ramat Yishai, Israel). The initial temperature of the thermode increased at a rate of 70°C/s, and reached 45°C from a baseline of 32°C, then was maintained at 45°C for 2 min of tonic heat. We chose 45°C for tonic heat pain to maintain similar sensations to the participant-controlled temperature assessment described below (15). At the end of the 2 min, participants also reported their pain verbally to the experimenter (0-10, 0—"no pain at all," 10—"worst pain imaginable"). Participants were instructed to rate their pain using a slider on the CoVAS machine, which has a visual of a linear increasing graph, indicating no pain on one end and the maximal amount of pain they could tolerate on the opposite end.

Participants were asked to rate their pain continuously, moving the slider as desired. The rating was recorded every 20 ms. The average pain rating from the CoVAS readings was recorded as average pain rating to tonic heat.

Participant Controlled Temperature

For participant controlled temperature (PCT), participants continuously adjusted the temperature of the thermode to maintain their initial perception (15). For example, if at the beginning of the 2-min trial (at 45°C) participants rated the pain as a 4/10, they were instructed to either increase or decrease the temperature in order to maintain the 4/10 sensation over the 2 min. Participants were provided a computer mouse to control temperature, whereby left and right button clicks changed the temperature by ±0.1°C, respectively. Participants were informed that the temperature "may feel as though it is increasing or decreasing," and were asked to maintain their initial perception by raising or lowering the heat through clicking the mouse. To confirm participants maintained their pain rating throughout the 2 min, each was asked to verbally report their pain at the beginning and end of the protocol. The protocol was identical to that presented by Jutzeler et al. (15). Average temperature across the 2 min of PCT was taken as the primary outcome.

Familiarization to Heat Pain Assessments

Familiarization trials for both CoVAS and PCT were conducted on a neutral test site. Participants were exposed to 1 min of heat, beginning at a baseline of 40°C. Then, the temperature oscillated by ±2°C at rate of 0.5°C/s. During this time, participants were instructed to rate their pain for CoVAS or to maintain consistent pain sensations via button clicks for PCT. This oscillation in temperature provided participants the opportunity to become accustomed to both heat sensations and the CoVAS and PCT apparatus in response to multiple temperatures. The familiarization trials also helped to reinforce the concept that the temperature in the PCT trials also could be perceived as though it was increasing or decreasing, supporting the blinding of participants to the nature of the PCT trials.

Questionnaires

At the conclusion of the second day of testing, the pain catastrophizing scale (PCS) questionnaire was administered. The PCS involves the participant rating 13 statements regarding the types of thoughts and feelings that occur when they are in pain from 0 ("not at all") to 4 ("all the time"). There are three subscales in the PCS; magnifying (three items, "*I become afraid that the pain will get worse*"), rumination (four items, "*I keep thinking about how badly I want the pain to stop*"), and helplessness (six items, "*It's terrible and I think it's never going to get any better*"). Higher PCS scores have been associated with greater levels of pain and pain-focused experiences (21). PCS scores also tend to be higher in females (1).

A demographics questionnaire was also delivered on the first day of testing, asking participants to report their sex, gender, and age. For gender, participants were asked "*What is your gender?*" with options for "female," "male," "non-binary/third gender," "prefer to self describe," or "prefer not to say."

Statistical Analysis

Cohen's d effect sizes were calculated for differences in pain outcomes between male and female participants separately by examiner. This was done to simply model pain outcomes measured by a single examiner of one sex, as would be commonplace in previous studies. The primary outcomes were verbal pain rating following 2 min of tonic heat, average CoVAS rating over 2 min of tonic heat pain, heat pain thresholds, and average temperature over 2 min of PCT assessment. Descriptive statistics were assessed using histograms, box plots, and Q-Q plots to confirm normal distributions of pain outcomes. A preliminary analysis revealed that all pain outcomes were normally distributed (Shapiro-Wilk test range: 0.05-0.29). To formally and comprehensively test our study design, we adopted a repeated measure ANOVAs approach with participant gender as a between-subject variable, and examiner gender as the within-subject variable. Order of testing (i.e., day 1 or day 2) was considered as a covariate to confirm effects were due to the examiner gender and not the repeat-testing nature of the study design. Significant interaction effects were further explored with *post hoc* Bonferroni corrected pairwise comparisons.

Relationships between PCS and pain outcomes were explored using bivariable Pearson correlations, with a Bonferroni correction for multiple comparisons. We examined relationships between pain scores and PCS across both testing sessions as well as explored associations between PCS scores and relative differences in pain scores between testing sessions (i.e., examiners).

RESULTS

Forty participants were recruited, 38 of which completed both sessions (20 females and 18 males). Missing data from the two subjects was due to technical issues with the heat stimulator—they were unable to complete either day of testing. All other subjects completed both experimental sessions. No subjects withdrew their data after debriefing. Upon debrief, all participants confirmed no knowledge of the true purpose of the study. All participants identified as cis-gendered.

Rating Based Methods

There was a significant main effect of participant gender on verbal pain rating to tonic heat [$F_{(1,36)} = 5.77$, $p = 0.02$, $\eta_p^2 = 0.14$]. This suggests that female participants verbally reported heat as more painful than men. Examiner gender had no main effect on verbal pain ratings [$_{F(1,36)} = 0.93$, $p = 0.34$, $\eta_p^2 = 0.03$]. However, there was a significant interaction effect for participant and experimenter gender on verbal pain rating [$F_{(1,36)} = 5.61$, $p = 0.02$, $\eta_p^2 = 0.14$]. Bonferroni corrected *post-hoc* analysis revealed that female participants verbally reported higher tonic heat pain than males in the presence of a female examiner ($t = 3.21$, $p = 0.01$). Order of day of testing did not influence the gender effect [$F_{(1,36)} = 0.01$, $p = 0.91$]. For average CoVAS ratings, there were no significant main effects of participant [$F_{(1,36)} = 1.20$, $p = 0.28$, $\eta_p^2 = 0.03$] or examiner gender [$F_{(1,36)} = 3.88$, $p = 0.06$, $\eta_p^2 = 0.10$]. There was also no significant

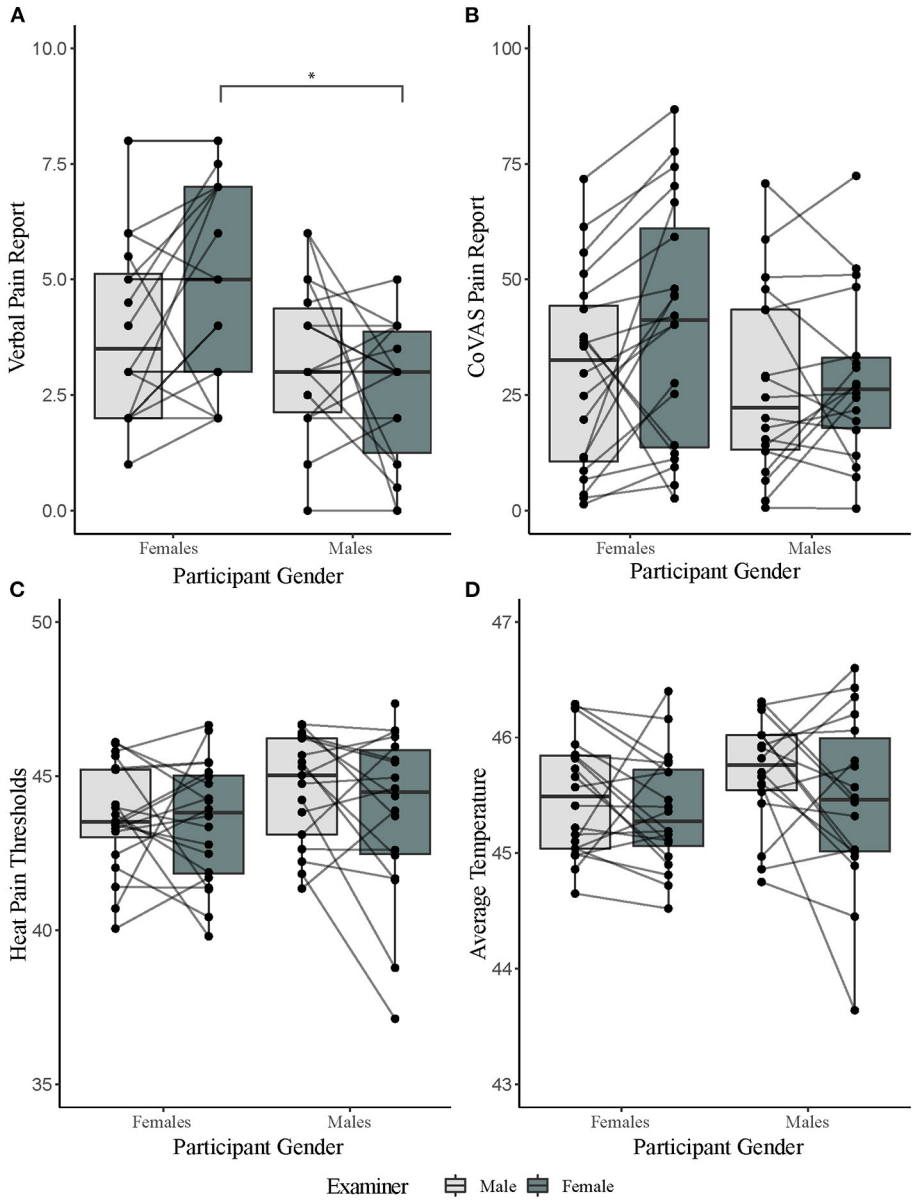


FIGURE 2 | Results of pain tests separated by participant and examiner gender. **(A)** Verbal pain reports for a 2-min tonic heat test separated by gender. **(B)** computerized visual analog scale (CoVAS) reports for a 2-min tonic heat test separated by gender. **(C)** Average heat pain threshold reports separated by gender. **(D)** PCT reports (as calculated by average temperature over a for a 2-min tonic heat test) separated by gender. *denotes significance level of $p < 0.05$ from Bonferroni corrected *post hoc* analysis.

interaction effect [$F_{(1,36)} = 2.70$, $p = 0.11$, $\eta_p^2 = 0.07$] (**Figure 2**; **Table 1**). Order of session did not influence CoVAS ratings [$F_{(1,36)} = 3.08$, $p = 0.09$].

Temperature Based Methods

There was no significant main effect of participant [heat pain thresholds: $F_{(1,36)} = 1.80$, $p = 0.19$, $\eta_p^2 = 0.05$; PCT: $F_{(1,36)} = 1.02$, $p = 0.32$, $\eta_p^2 = 0.03$] or examiner gender [heat pain thresholds: $F_{(1,36)} = 2.64$, $p = 0.11$, $\eta_p^2 = 0.07$; PCT: $F_{(1,36)} = 3.31$, $p = 0.08$, $\eta_p^2 = 0.08$]. There was also no significant interaction effect between participant and examiner gender [heat pain thresholds: $F_{(1,36)} = 1.59$, $p = 0.22$, $\eta_p^2 = 0.04$; PCT: $F_{(1,36)} = 0.45$, $p = 0.51$, $\eta_p^2 = 0.01$] (**Figure 2**; **Table 1**). Order of session did not influence PCT scores [$F_{(1,36)} = 0.56$, $p = 0.46$] or pain thresholds [$F_{(1,36)} = 0.66$, $p = 0.42$].

We also ran a repeat measures ANOVA on the initial rating of the PCT stimulus to investigate if there was a gender difference in this initial perception. There was no significant difference between genders [$F_{(1,36)} = 2.417$, $p = 0.129$], nor was there an

TABLE 1 | Means and standard deviations and results of the repeated measures ANOVA tests, separated by male and female participants and examiners. ANOVA output for interaction effect presented.

	Female examiner Mean (SD)		Male examiner Mean (SD)		RM-ANOVA[a]
	Females	**Males**	**Females**	**Males**	***F* (*p*)**
PCT	45.36 (0.49)	45.48 (0.51)	45.45 (0.75)	45.70 (0.47)	0.45 (0.51)
CoVAS	29.83 (21.23)	40.34 (26.25)	28.45 (18.16)	27.50 (20.53)	2.70 (0.11)
Heat pain thresholds	43.64 (1.73)	43.5 (1.97)	43.84 (2.74)	44.93 (2.32)	1.59 (0.22)
Verbal rating	4.00 (2.03)	4.93 (2.25)	3.33 (1.64)	2.94 (1.52)	5.61 (0.02)

[a] *df = 19, interaction effect of participant gender x examiner gender.*

RM-ANOVA, repeated measure ANOVA; PCT, participant controlled temperature; CoVAS, computerized visual analog scale.

TABLE 2 | Correlations coefficients (R) between pain catastrophizing subscales and pain measurements adjusted for multiple comparisons (Bonferroni).

	Male examiner				Female examiner				Difference scores between male and female examiners			
	CoVAS	**PCT**	**PT**	**Verbal**	**CoVAS**	**PCT**	**PT**	**Verbal**	**CoVAS**	**PCT**	**PT**	**Verbal**
Rumination	0.16	−0.16	−0.23	0.12	0.10	−0.23	−0.16	0.22	0.05	0.10	−0.05	−0.13
Magnification	0.14	−0.14	−0.21	0.01	0.09	−0.05	−0.11	0.08	0.05	−0.06	−0.08	−0.09
Helplessness	0.23	−0.32	−0.28	0.20	0.20	−0.03	−0.09	0.28	0.00	−0.22	−0.16	−0.13
Total	0.21	−0.24	−0.27	0.14	0.16	−0.11	−0.13	0.23	0.03	−0.08	−0.11	−0.13

effect of examiner gender [$F_{(1,36)} = 1.490$, $p = 0.230$]. Average initial rating for female participants was 5.55 ± 1.56 when tested by the male examiner and 5.80 ± 1.64 when tested by the female examiner. Average initial rating for male participants was 4.83 ± 1.58 when tested by the male examiner and 5.00 ± 1.57 when tested by the female examiner. Additionally, 33/38 participants reported the same pain rating at the beginning and end of the PCT test, 4 were within ±1/10 on an NRS, and 1 participant was within ±2/10 on an NRS. This is in contrast to the CoVAS test, where the range was ±3/10 on the NRS.

PCS Correlations to Pain Outcomes

PCS subscales were not correlated to any pain outcomes in both males and females, and were also not correlated to relative difference in pain outcomes between examiners (**Table 2**).

DISCUSSION

The impact of examiner characteristics on study outcomes have been attributed a causal role in the ongoing scientific replication crisis (3). Among concerns is that the gender of the examiner contributes to heterogeneous outcomes between studies. As predicted by the Gender Context Model of Pain, we observed that sex differences in tonic heat pain perception may be exaggerated by verbal rating-based methods when the examiner is female. CoVAS pain ratings demonstrated similar trends, albeit not significant. In contrast, temperature-based methods of assessing heat pain were not significantly affected by the gender of the examiner.

To our knowledge, the effect of examiner gender on pain outcomes has been explicitly tested in six previous studies (see **Table 3** for description) (9, 10, 17–19, 22). For subjective pain ratings, our observations correspond with those reporting an opposing examiner gender effect (9, 10, 17, 18) as well as social theories of pain which propose the gender context in which pain is expressed influences pain report (12). The former was evidenced in our reported verbal ratings in women, which were significantly higher in the presence of a female compared to a male examiner. Similar, albeit more variable results were observed for CoVAS ratings to heat pain. Our findings support the notion that pain communication may be more affected by gender interactions as compared to the actual pain experience. For example, when comparing verbal pain ratings to CoVAS ratings, the pain experience (CoVAS) was comparable, while the act of reporting to the experimenter verbally was influenced my experimenter gender. The notion that pain communication, but not experience, is influenced by gender is supported by a previous study that showed biological responses to pain (e.g., autonomic changes) are unaffected by examiner gender (17). Taken together, our findings provide evidence for a dissociation between pain experience and pain reporting, which is influenced by examiner gender. Overall, this lends support to the Gender Context Model of Pain (12), in that outcomes with the most social communication were more influenced by experimenter gender.

The modernization of QST assessments has seen a shift to temperature-based methods, including standardized methods of measuring heat pain thresholds (6). Previous studies exploring experimenter gender effects (**Table 3**) have not incorporated temperature-based methods of assessing pain, relying instead on

TABLE 3 | Summary of studies examining the effect of examiner gender on pain outcomes.

References	Test stimuli	Rating method	Gender effects	Additional measures	Study design
Levine and De Simone (9)	**Cold pressor** Both hands in 0-1°C ice bucket	**Pain intensity** Numeric rating scale, given every 15 s for 180 s	**Intensity** Male participants reported lower pain intensity to a female experimenter	**Pain Affective scale** • Males reported less negative affective words to female experimenter	Parallel experimental design **Participants** *n* = 68 (33 female, 35 male) Ages 17-29 (M = 19.13)
Kallai et al. (18)	**Cold pressor** Non-dominant hand in circulating −1°C ice bucket	**Pain intensity** 10-point rating scale, given immediately after CPT **Pain threshold** Seconds **Pain tolerance** Seconds	**Intensity** Both male and female participants reported higher pain intensity to a female experimenter **Tolerance** Female participants had higher pain tolerances with a male experimenter Male participants had higher pain tolerances with a female experimenter **Threshold** No experimenter gender effect found for pain threshold	Participants rated the examiner's authority, competence, likeability and masculinity/femininity on seven-point rating scales	Parallel experimental design **Participants** *n* = 160 (80 female, 80 male) Female ages 17-36 (M = 23.19, SD 3.59) Male ages 19-59 (M = 24.55, SD 5.79)
Gijsbers and Nicholson (10)	**Pressure** Pressure algometer with 0-9 kg force range on upper sternum	**Pain threshold** kilograms	**Threshold** Male participants had higher pain thresholds with a female examiner	**Anxiety** • Measured with 10 cm VAS • Anxiety was low for both female and male participants • No correlation with pain thresholds **McGill Pain Questionnaire** • No significant examiner gender effect on pain scores • Indicated low emotional concern in participants	Parallel experimental design **Participants** *n* = 64 (32 females, 32 males) Female ages 18-36 (M = 21.0, SD 4.4) Male ages 18-49 (M = 23.0, SD 8.1)
Weisse et al. (19)	**Cold pressor** Non-dominant hand in 0-2°C ice bucket	**Pain intensity** 0–20 rating scale every 15 s for a total of 300 s	**Intensity** No main effect found for pain reporting and examiner gender. However, an interaction was found with participant race and examiner gender: Black participants reported higher pain intensities than white participants to a female examiner	**Pain unpleasantness scale** • No main effect for pain reporting and examiner gender • An interaction found for participant race and examiner gender: black participants reported more unpleasantness than white participants to a female examiner	Parallel experimental design **Participants** *n* = 343 (187 females, 156 males) Ages 17-43 (M = 20.27)
Aslaksen et al. (17)	**Heat** TSA II Neurosensory Analyzer (Medoc, Israel): 30 • 30 mm aluminium contact thermode with a 10°C/s change rate on right forearm	**Pain intensity** 100 mm VAS **Physiological pain response** Heartrate variability and skin conductance levels	**Intensity** Male participants reported lower pain intensity to a female examiner Physiological pain response: No examiner gender effect found for physiological responses	**Pain unpleasantness scale** • No significant examiner gender effect • Short **Adjective Check List and Self-Assessment Manikin scale** • Male participants reported lower arousal to female experimenters • No significant examiner gender effect with subjective stress or mood scales	Parallel experimental design **Participants** *n* = 64 (32 females, 32 males) Female ages 19-40 (M = 23.61, SD 3.99) Male ages 19-35 (M = 23.3, SD 2.49)

(Continued)

TABLE 3 | Continued

References	Test stimuli	Rating method	Gender effects	Additional measures	Study design
Vigil et al. (22)	**Cold pressor** One of two CPT protocols used on left hand: (1) 5°C ± 1°C circulating ice bucket, or (2) Isotemp 6200R28 (Fisher Scientific, USA) electromechanical CPT device at 5°C ± 0.1°C	**Pain intensity** 10-point VAS, 30 s into CPT **Pain threshold** Seconds **Pain tolerance** Seconds	**Intensity** Both male and female participants reported higher pain intensity to a female examiner **Tolerance** Subjects had higher pain tolerances with a male examiner **Threshold** No examiner gender effect found for pain threshold	No additional measures performed	Parallel experimental design **Participants** $n = 352$ (48% males) Ages 18-30 (M = 19.8, SD 2.1)

CPT, cold pressor test.

verbal ratings or time-based approaches that assess tolerance (e.g., cold pressor). To address this limitation, we assessed examiner gender effects on pain threshold determined by method of limits and PCT. The latter, a revitalized approach based a method originally established by Hardy and Greene (23), involves participants continuously adjusting the temperature of the thermode over 2 min in order to maintain their initial perception of noxious heat (15). The concept of PCT is similar to CoVAS, but dynamic aspects of pain (i.e., the fluctuations in the perception of a constant painful stimuli over time) are reflected by changes in temperature as opposed to continuous ratings (15). Compared to CoVAS and verbal pain ratings, PCT provides pain reporting with the least obvious social context. Where verbal pain ratings involved direct communication with examiners and CoVAS involved the perceived communication of digital 0-10 scale, PCT involves button clicks to maintain sensation. To that end, PCT was more resilient to gender effects compared to verbal or CoVAS outcomes, as examiner gender did not significantly influence PCT. These findings provides further support for the social context of pain model, as PCT is less clearly a "rating" of pain to an examiner, and thus less influenced by the social context (12).

To consider a potential psychosocial factor, we aimed to explore the relationship between participants' PCS scores and variability introduced by the gender of the examiner. For pain catastrophizing, we observed no significant associations between PCS scores and any pain outcomes, for both raw scores as well as evaluating relative differences in pain outcomes between examiners. This suggests that pain catastrophizing does not have a significant influence on our observed gender effects on pain outcomes.

Limitations

Our findings are limited to a relatively homogenous population (i.e., undergraduate and graduate students). The extent our results are generalizable to other populations (e.g., older, community dwelling adults) requires further study. We also did not collect or report relationships between the race, ethnicity, height, or weight of our participants or examiners and the possible effects on pain ratings. This was beyond the scope of our current study and represents another avenue for further exploration. To that end, we did not control for experimenter ethnicity, or other examiner characteristics (e.g., hair color, eye color etc.). We sought to maintain ecological validity in our selection of a male and female examiner, rather than overly constrain various aspects of personal appearance/characteristics. To that end, our findings are based on the effect of clearly male and clearly female examiners.

In comparison to previous investigations of experimenter gender effects on pain perception (9, 10, 17–19, 22), our study is limited to a relatively small sample size. However, as a seminal study to explore experimenter gender effects on multiple heat-pain outcomes, our sample size was chosen pragmatically and in accordance with a sample size calculation related to quantitative pain assessments previously used in similar experimenter gender comparisons. We were unable to collect data on two participants due to technical issues, resulting in a fewer number of participants than reported in our a priori power calculation. We reported η_p^2 values for all repeated measures ANOVA analyses along with Cohen's d values to highlight within experimenter effects. Finally, our findings are also limited to our included heat pain-based assessment methods. Future studies should continue to explore experimenter gender effects in other pain outcomes making use of differing modalities, such as mechanical pinpricks and more modern cold pain assessments.

We did not have our examiners conform to stereotypical gender roles, which may have muted examiner effects. Studies whose examiners dressed in "stereotypical gender conforming" ways (9, 10, 17, 18) appear more likely to see significant examiner effects compared to those that did not control for dress (24–31). Status of the examiner may also matter—participants of both genders report higher pain tolerance to "high status" (i.e., professionally dressed, used formal names) examiners (18). In the present study we attempted to control for gender stereotypes through recruiting peer examiners that wore a uniform—lab coat over pants and a t-shirt—and that used the same script. This moderate "de-gendering" of the examiners and reduction of potential power imbalances through using peers may have reduced gender differences in the heat pain assessments.

Also, our study and those previous have focused on participants that conformed to gender norms. It is not clear if those who do not conform to gender norms may report

pain differently or have different examiner-participant gender interaction effects in the reporting of pain. Examining pain in transgendered and non-binary individuals represents an important and understudied area of pain science—an area that would not only shed light on a marginalized populations' pain experience, but would also extend our understanding of the interaction between gender and pain.

Conclusions and Future Implications

Overall, our findings are aligned with the Gender Context Model of Pain, insofar as those outcome measures that were most likely to be influenced by social factors (i.e., verbal pain ratings) were more susceptible to experimenter gender effects, while outcomes less likely to be influenced by social factors (i.e., PCT) were not significantly influenced by experimenter gender. The examiner and participant gender can both influenced pain reporting, with the perceived level of examiner-participant interaction appearing to mediate these effects. Researchers should consider the social environment of their experiments, the pain measurement used, and the gender of their experimenters as these factors all play a role in detecting sex/gender differences in pain measurements. The use of non-verbal pain measures, with little to no examiner influence (e.g., coded temperature information *via* PCT) may be a potential solution to circumvent the effects of experimenter gender on pain related outcomes.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Behavioral Research Ethics Board at the University of British Columbia. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JM, LL, and JK designed and directed the project. NB and RB collected the data and ran the statistics, with assistance and supervision by JM and LL. Figures were created by JM and LL. JM wrote the manuscript, with all authors contributing to the final version.

FUNDING

JK was supported by a NSERC discovery grant.

REFERENCES

1. Racine M, Tousignant-Laflamme Y, Kloda LA, Dion D, Dupuis G, Choinire M. A systematic literature review of 10 years of research on sex/gender and pain perception - part 2: do biopsychosocial factors alter pain sensitivity differently in women and men? *Pain.* (2012) 153:619–35. doi: 10.1016/j.pain.2011.11.026
2. Mogil JS. Sex differences in pain and pain inhibition: multiple explanations of a controversial phenomenon. *Nat Rev Neurosci.* (2012) 13:859–66. doi: 10.1038/nrn3360
3. Chapman CD, Benedict C, Schiöth HB. Experimenter gender and replicability in science. *Sci Adv.* (2018) 4:e1701427. doi: 10.1126/sciadv.1701427
4. Bartley EJ, Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings. *Br J Anaesth.* (2013) 111:52–8. doi: 10.1093/bja/aet127
5. Bernardes SF, Keogh E, Lima ML. Bridging the gap between pain and gender research: a selective literature review. *Eur J Pain.* (2008) 12:427–40. doi: 10.1016/j.ejpain.2007.08.007
6. Rolke R, Magerl W, Andrews Campbell K, Schalber C, Caspari S, Birklein F, et al. Quantitative sensory testing: a comprehensive protocol for clinical trials. *Eur J Pain.* (2006) 10:77–88. doi: 10.1016/j.ejpain.2005.02.003
7. Daniali H, Flaten MA. A qualitative systematic review of effects of provider characteristics and nonverbal behavior on pain, and placebo and nocebo effects. *Front Psychiatry.* (2019) 10:242. doi: 10.3389/fpsyt.2019.00242
8. Turk D, Monarch E. Biopsychosocial perspective on chronic pain. In: Turk DC, Gatchel RJ, editors. *Psychological Approaches to Pain Management: A Practitioner's Handbook.* 3rd ed. New York, NY: The Guildford Press (2018). p. 3–29.
9. Levine FM, Simone LL De. The effects of experimenter gender on pain report in male and female subjects. *Pain.* (1991) 44:69–72. doi: 10.1016/0304-3959(91)90149-R
10. Gijsbers K, Nicholson F. Experimental pain thresholds influenced by sex of experimenter. *Percept Mot Skills.* (2005) 101:803–7. doi: 10.2466/pms.101.3.803-807
11. Keogh E. Gender differences in the nonverbal communication of pain: a new direction for sex, gender, and pain research? *Pain.* (2014) 155:1927–31. doi: 10.1016/j.pain.2014.06.024
12. Keogh E. The gender context of pain. *Health Psychol Rev.* (2020) 15:454-81. doi: 10.1080/17437199.2020.1813602
13. Quartana PJ, Campbell CM, Edwards RR. Pain catastrophizing: a critial review. *Expert Rev Neurother.* (2009) 9:745–58. doi: 10.1586/ern.09.34
14. Arendt-Nielsen L, Yarnitsky D. Experimental and clinical applications of quantitative sensory testing applied to skin, muscles and viscera. *J Pain.* (2009) 10:556–72. doi: 10.1016/j.jpain.2009.02.002
15. Jutzeler CR, Sirucek L, Scheuren PS, Bobo T, Anenberg E, Ortiz O, et al. New life for an old idea: assessing tonic heat pain by means of participant controlled temperature. *J Neurosci Methods.* (2019) 321:20–7. doi: 10.1016/j.jneumeth.2019.04.003
16. Faul F, Erdfelder E, Lang AG, BuchnerA. A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods.* (2007) 39:175–91. doi: 10.3758/BF03193146
17. Aslaksen PM, Myrbakk IN, Høifødt RS, Flaten MA. The effect of experimenter gender on autonomic and subjective responses to pain stimuli. *Pain.* (2007) 129:260–8. doi: 10.1016/j.pain.2006.10.011
18. Kallai I, Barke A, Voss U. The effects of experimenter characteristics on pain reports in women and men. *Pain.* (2004) 112:142–7. doi: 10.1016/j.pain.2004.08.008
19. Weisse CS, Foster KK, Fisher EA. The influence of experimenter gender and race on pain reporting: does racial or gender concordance matter? *Pain Med.* (2005) 6:80–7. doi: 10.1111/j.1526-4637.2005.05004.x
20. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *J Am Med Assoc.* (2013) 310:2191–4. doi: 10.1001/jama.2013.281053
21. Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess.* (1995) 7:524–32. doi: 10.1037/1040-3590.7.4.524
22. Vigil JM, Ba LNR, Alcock J, Maestes R. Laboratory personnel gender and cold pressor apparatus affect subject pain reports. *Res Pain Manag.* (2014) 19:e13–8. doi: 10.1155/2014/213950

23. Greene LC, Hardy JD. Adaptation of thermal pain in the skin. *J Appl Physiol.* (1962) 17:693–6. doi: 10.1152/jappl.1962.17.4.693
24. Feine JS, Bushnell MC, Miron D, Duncan GH. Sex differences in the perception of noxious heat stimuli. *Pain.* (1991) 44:255–62. doi: 10.1016/0304-3959(91)90094-E
25. Bush FM, Harkins SW, Harrington WG, Price DD. Analysis of gender effects on pain perception and symptom presentation in temporomandibular pain. *Pain.* (1993) 53:73–80. doi: 10.1016/0304-3959(93)90058-W
26. Otto MW, Dougher MJ. Sex differences and personality factors in responsivity to pain. *Percept Mot Skills.* (1985) 61:383–90. doi: 10.2466/pms.1985.61.2.383
27. Fillingim RB, Edwards RR, Powell T. The relationship of sex and clinical pain to experimental pain responses. *Pain.* (1999) 83:419–25. doi: 10.1016/S0304-3959(99)00128-1
28. Carter LE, Mcneil DW, Ma KEV, Ma JTS, Turk CL, Ries BJ, et al. Effects of emotion on pain reports, tolerance and physiology. *Pain Res Manag.* (2002) 7:21–30. doi: 10.1155/2002/426193
29. Thorn BE, Clements KL, Ward LC, Dixon KE, Kersh BC, Boothby JL, et al. Personality factors in the explanation of sex differences in pain catastrophizing and response to experimental pain. *Pain.* (2004) 20:275–82. doi: 10.1097/00002508-200409000-00001
30. Essick G, Guest S, Martinez E, Chen C, Mcglone F, Essick G, et al. Site-dependent and subject-related variations in perioral thermal sensitivity. *Somatosens Mot Res.* (2004) 21:159–75. doi: 10.1080/08990220400012414
31. Jackson T, Iezzi T, Chen H, Ebnet S, Eglitis K. Gender, interpersonal transactions, and the perception of pain: an experimental analysis. *J Pain.* (2005) 6:228–36. doi: 10.1016/j.jpain.2004.12.004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pain Research

ORIGINAL RESEARCH
published: 24 December 2021
doi: 10.3389/fpain.2021.784362

The Potential Clinical Utility of Pressure-Based vs. Heat-Based Paradigms to Measure Conditioned Pain Modulation in Healthy Individuals and Those With Chronic Pain

Rima El-Sayed[1,2], *Camille Fauchon*[2], *Junseok A. Kim*[1,2], *Shahrzad Firouzian*[1,2], *Natalie R. Osborne*[1,2], *Ariana Besik*[2], *Emily P. Mills*[2], *Anuj Bhatia*[2,3] *and Karen D. Davis*[1,2,4]*

[1] *Institute of Medical Science, University of Toronto, Toronto, ON, Canada,* [2] *Krembil Brain Institute, Division of Brain, Imaging, and Behaviour, University Health Network, Toronto, ON, Canada,* [3] *Department of Anesthesia and Pain Medicine, Toronto Western Hospital, University of Toronto, Toronto, ON, Canada,* [4] *Department of Surgery, University of Toronto, Toronto, ON, Canada*

OPEN ACCESS

Edited by:
Kai Karos, Open University of the Netherlands, Netherlands

Reviewed by:
Elena Enax-Krumova, BG University Hospital Bergmannsheil GmbH, Germany
Wacław M. Adamczyk, University of Luebeck, Germany

***Correspondence:**
Karen D. Davis
karen.davis@uhnresearch.ca

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *27 September 2021*
Accepted: *06 December 2021*
Published: *24 December 2021*

Citation:
El-Sayed R, Fauchon C, Kim JA, Firouzian S, Osborne NR, Besik A, Mills EP, Bhatia A and Davis KD (2021) The Potential Clinical Utility of Pressure-Based vs. Heat-Based Paradigms to Measure Conditioned Pain Modulation in Healthy Individuals and Those With Chronic Pain. Front. Pain Res. 2:784362. doi: 10.3389/fpain.2021.784362

Conditioned pain modulation (CPM) is a physiological measure thought to reflect an individual's endogenous pain modulation system. CPM varies across individuals and provides insight into chronic pain pathophysiology. There is growing evidence that CPM may help predict individual pain treatment outcome. However, paradigm variabilities and practical issues have impeded widespread clinical adoption of CPM assessment. This study aimed to compare two CPM paradigms in people with chronic pain and healthy individuals. A total of 30 individuals (12 chronic pain, 18 healthy) underwent two CPM paradigms. The heat CPM paradigm acquired pain intensity ratings evoked by a test stimulus (TS) applied before and during the conditioning stimulus (CS). The pressure CPM paradigm acquired continuous pain intensity ratings of a gradually increasing TS, before and during CS. Pain intensity was rated from 0 (no pain) to 100 (worst pain imaginable); Pain50 is the stimulus level for a response rated 50. Heat and pressure CPM were calculated as a change in TS pain intensity ratings at Pain50, where negative CPM scores indicate pain inhibition. We also determined CPM in the pressure paradigm as change in pressure pain detection threshold (PDT). We found that in healthy individuals the CPM effect was significantly more inhibitory using the pressure paradigm than the heat paradigm. The pressure CPM effect was also significantly more inhibitory when based on changes at Pain50 than at PDT. However, in individuals with chronic pain there was no significant difference in pressure CPM compared to heat or PDT CPM. There was no significant correlation between clinical pain measures (painDETECT and Brief Pain Inventory) and paradigm type (heat vs. pressure), although heat-based CPM and painDETECT scores showed a trend. Importantly, the pressure paradigm could be administered in less time than the heat paradigm. Thus, our study indicates that in healthy individuals, interpretation of CPM findings should consider potential modality-dependent effects. However, in individuals with chronic pain, either heat or pressure paradigms

can similarly be used to assess CPM. Given the practical advantages of the pressure paradigm (e.g., short test time, ease of use), we propose this approach to be well-suited for clinical adoption.

Keywords: conditioned pain modulation (CPM), stimulus modality, cuff algometry, heat thermode, chronic pain, antinociception

INTRODUCTION

Conditioned pain modulation (CPM) is a behavioral phenomenon that reflects an individual's inherent capacity to modulate their pain. CPM can be evoked experimentally using "pain inhibits pain" type psychophysical tests (1). Numerous studies have demonstrated the potential clinical utility of CPM to predict the effectiveness of therapeutic approaches that target mechanisms of CPM (2–4).

The CPM effect [a term coined by Yarnitsky et al. (5)] refers to any change in the intensity of pain that is evoked by a test stimulus (TS) applied to one area of the body due to the presence of a concurrent conditioning stimulus (CS) applied to another area of the body (6). This psychophysical measure of CPM designed for testing in humans was motivated by the discovery of the diffuse noxious inhibitory control (DNIC) effect observed in animal electrophysiological single neuronal recordings. Decades of DNIC studies have shown that a noxious stimulus activates a spino-bulbar-spinal feedback loop such that spinal nociceptive projection neurons activate neurons in the brainstem subnucleus reticularis dorsalis (SRD) (7–9). The SRD then activates descending projections through the dorsolateral funiculus, that ultimately inhibits ipsilateral wide dynamic range (WDR) spinal dorsal horn neurons, and thus attenuates their response to a second incoming noxious stimulus (9, 10). However, unlike the inhibitory DNIC effect in animals, the CPM effect in human can be inhibitory or facilitatory. It is now clear that CPM can vary across a wide spectrum, from reduced pain due to the presence of a CS (inhibitory CPM) to increased pain (facilitatory CPM), and in some cases CPM may not occur at all (no-CPM) (6, 11, 12).

Individual factors contribute to the variability of CPM across the population. A systematic meta-analysis in many chronic pain conditions found that on average, people with chronic pain exhibit a weaker inhibitory CPM effect compared to healthy individuals (13). For example, weaker inhibitory CPM has been reported in studies of people with neuropathy, fibromyalgia, irritable bowel syndrome, osteoarthritis, tension-type headache and whiplash-associated disorders (6, 13). Furthermore, there is evidence that an individual's CPM may be used as a clinical measure to guide personalized treatment selection. For example, in a study of people undergoing treatment for painful diabetic neuropathy with the serotonin-noradrenaline reuptake inhibitor (SNRI) duloxetine, patients with weaker inhibitory CPM (thought to reflect a weaker anti-nociceptive pathway) benefited more than those with a stronger inhibitory CPM (3). Furthermore, the improvement in clinical pain was observed alongside an improvement of post-treatment CPM. Thus, this patient-specific treatment outcome was thought to be due to the action of this SNRI to strengthen the descending anti-nociceptive serotonergic and adrenergic neurotransmission that is part of the spino-bulbar-spinal loop. A link between CPM and pain treatment outcome was also found in two studies of osteoarthritis, where patients' CPM shifted to more closely resemble that of the healthy group following a successful knee or hip surgery treatment (14, 15).

Studies of CPM in pain-free individuals are also important not only to glean insight into basic mechanisms of pain modulation, but also to determine its utility in risk assessment for the potential development of chronic pain. For example, compared to the quantitative sensory tests for pain thresholds and suprathreshold pain assessed before a thoracotomy, stronger inhibitory CPM was the only measure that predicted the lower risk of developing chronic post-surgery pain (4). A similar finding was also reported for patients undergoing cesarean and major abdominal surgeries (16, 17). Therefore, assessing CPM has potential clinical utility to predict the risk of persistent post-operative pain, as well as to predict the efficacy of therapeutic approaches that target endogenous pain modulation, which can ultimately guide treatment plans for chronic pain management.

Despite decades of research in the field of DNIC and CPM, there remains challenges to adopting a CPM test for clinical use. Practical issues can be major factors that impact translating CPM testing from an experimental research tool into a clinical tool. Thus, it is important to establish methodology that is easy to administer and conducive to a clinical setting. For example, there have been recent pursuits to establish a new simple pressure pain stimulator that can induce CPM for bed-side testing (18). Additionally it has been suggested that clinical translation of CPM could be helped by increasing clinical experimental data that assesses the dependency of CPM on stimulus test modalities (19). In the past, CPM has been assessed with paradigms that use different types of stimulus modalities (e.g., heat, cold, electrical and pressure) and there are also different metrics used to quantify the CPM effect (e.g., a change in suprathreshold pain ratings vs. pain detection thresholds). The assumption in the field has been that different stimulus modalities produce basically the same CPM effect, however this has not been definitively established. In 2015, the growing need to reduce variability and standardize the CPM paradigm led a group of experts to recommend the use of either heat or pressure stimulus based paradigms (20). However, since that time, the field has continued to evolve without any particular paradigm being established as a gold standard. Therefore, the aim of the current study was to use a within-subject analysis to assess a commonly used heat-based paradigm with a presumptive simpler pressure-based

paradigm in healthy individuals and those with chronic pain. We hypothesized that CPM based on a heat vs. a pressure paradigm would not differ significantly in an individual (healthy or with chronic pain).

MATERIALS AND METHODS

Participants

The study consisted of two groups: 1) healthy individuals recruited through advertisements posted throughout the University Health Network, Toronto, Canada and through word of mouth, and 2) people with chronic pain who were recruited as part of a larger, ongoing study of chronic pain. All study participants provided informed consent for the procedures approved by the University Health Network Research Ethics Board. All study participants underwent evaluation of CPM using both a heat pain-based paradigm and a pressure pain-based paradigm, allowing for both within-subject and group evaluations. The CPM data in this study were collected as part of a large battery of psychophysical tests for studies of acute and chronic pain. Healthy participants were excluded if they had 1) current ongoing pain or a history of chronic pain (pain lasting >3 months) 2) any major chronic health condition, or 3) a psychiatric disorder, neurological disorder, or a Beck inventory Depression (BDI) score (range 0–63) >13 (indicating greater than minimal self-reported depression). The chronic pain group consisted of people with chronic pain who were awaiting a spinal cord stimulation trial for pain management due to failed back surgery syndrome with back and/or lower limb pain ($n = 7$), complex regional pain syndrome in the lower limbs ($n = 3$), post-traumatic neuropathic pain in the lower limb ($n = 1$), and occipital neuralgia ($n = 1$).

Evaluation of Conditioned Pain Modulation

In the heat paradigm, stimuli were delivered to the volar forearms through two 30 × 30 mm contact thermodes (QSense device; Medoc Ltd, Israel) (**Figure 1**). In the pressure paradigm, stimuli were delivered to the calves through two inflatable 10 × 61 cm pressure cuffs (CPAR, NociTech Inc., Denmark) (**Figure 1**). In individuals with chronic pain, the cuff was applied to the upper arm bicep if their chronic pain included the leg. This was to ensure that CPM was tested in both paradigms at a body region that was not affected by the chronic pain condition. Stimulus-evoked pain intensity was rated on a scale from 0 to 100 (0 being no pain at all and 100 being the worst pain imaginable) in both paradigms. Participants provided these pain intensity ratings verbally during the heat paradigm and manually using a visual analog scale (VAS) slider during the pressure paradigm.

Conditioned Pain Modulation Calculation

The test stimulus (TS) and conditioning stimulus (CS) were set individually for each participant at an intensity that evoked a pain intensity rating of ∼50 out of 100 (known as Pain50). The CPM paradigm used was a parallel sequence paradigm where the CS was given concurrently with the second TS as follows: (1) pain intensity is rated during a TS (TS1), (2) a sustained CS is applied to the contralateral body region, (3) during the CS, the pain intensity of the second test stimulus (TS2) is rated. The CPM effect was calculated as a percentage using the following formula:

$$CPM\ Effect\ \% = \frac{TS2\ Pain\ rating - TS1\ Pain\ rating}{TS1\ Pain\ rating} \times 100\%$$

Therefore, a negative CPM effect is indicative of inhibitory CPM where a concurrent CS results in a lower pain rating of the second TS. A positive CPM effect is indicative of facilitatory CPM where a concurrent CS results in a higher pain rating of the second TS. Lastly, 0% indicates no CPM effect, where the concurrent CS did not change the pain rating of the TS.

Pain ratings at Pain50 were determined for both heat and pressure paradigms. In the pressure paradigm, in addition to the Pain50 measure used to calculate CPM, we determined the pressure pain detection threshold (PDT) and pain tolerance threshold (PTT) because previous studies have used these metrics to calculate the CPM effect. To be consistent with designating a negative CPM effect as reflecting inhibitory CPM, we calculated CPM from the pressure pain detection threshold (PDT) with the formula:

$$CPM\ Effect\ \% = \frac{TS1\ PDT - TS2\ PDT}{TS1\ PDT} \times 100\%$$

Heat-Based CPM Paradigm

For each participant, prior to the CPM test, a familiarization paradigm was used to determine their Pain50. In this paradigm, participants rated the pain intensity that was evoked by each of the six heat stimuli in the following order: 44, 45, 43, 46, 42, and 47°C. Since the aim was to find a temperature that evokes a pain intensity rating of 50/100, if any of the first five stimuli evoked a pain intensity rating >75/100, then the last 47°C did not need to be tested. Each of these familiarization test stimuli were delivered from a baseline temperature and interstimulus temperature of 35°C for 15 s and a ramp-up rate of 2°C/s to reach the target temperature which was held for 6 s. After the temperature was at the target temperature for 3 s, participants were prompted to rate the evoked pain intensity and the thermode temperature returned to baseline at a rate of 1°C/s. The temperature that evoked Pain50 was estimated from the familiarization paradigm. We then confirmed that this stimulus did evoke a pain rating of 50/100 during several TS that were part of a habituation paradigm (TS had identical timing and ramp rates to the TS in the CPM paradigm below). The Pain50 TS and CS temperatures were then manually set based on the result of the familiarization and habituation paradigm.

To test CPM, one thermode delivered the TS at 2°C/s from a 35°C baseline to the target Pain50 temperature. The temperature was held at this target for 7 s at which point the participant verbally provided a rating of their pain intensity, and then the temperature decreased back to baseline at 1°C/s (**Figure 1**). The second thermode delivered the CS to the contralateral forearm, the temperature increased from baseline at 1°C/s to the Pain50, was held there for 100 s, and then returned to baseline at 1°C/s

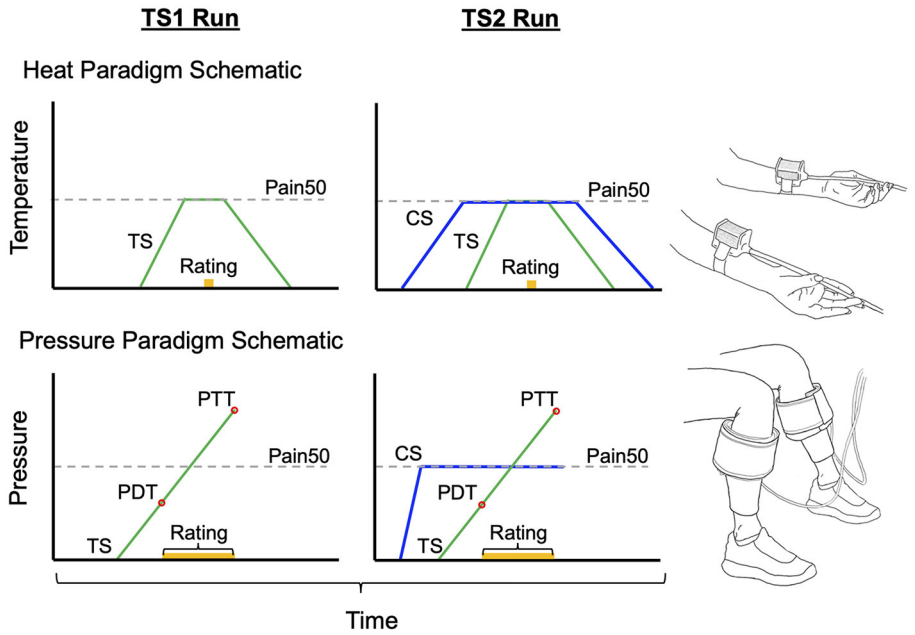


FIGURE 1 | Schematics to represent the stimulus order given in each CPM paradigm and the standard setup. In the heat paradigm the test stimulus (TS) is held at the Pain50 temperature only while the pain rating is obtained. In the pressure paradigm the TS continues to rise after Pain50 until they reach their pain tolerance. However, the pain rating at Pain50 pressure is extracted from TS1 and TS2 in order to calculate CPM similarly to the heat paradigm. Both are parallel sequence paradigms [conditioning stimulus (CS) in blue overlaps with TS2 when the second pain rating is obtained]. Pain50 is the stimulus intensity that evokes a pain rating of 50/100, where 0 is no pain and 100 is the worst pain imaginable. PDT, pain detection threshold; PTT, pain threshold tolerance.

(**Figure 1**). The TS1 started after 5 s of baseline and the TS2 was delivered 69 s after TS1 (i.e., during the CS). The full heat-based protocol (familiarization, habituation, and CPM) required ∼16 min to complete.

Pressure-Based CPM Paradigm

In the pressure paradigm, pain was rated continuously from the moment the TS starts to feel painful (i.e., the threshold for detecting pressure pain, PDT) until the tolerance level is reached (i.e., the threshold for pain tolerance, PTT). Previous studies using pressure paradigms commonly calculate the CPM effect as a change in PDT between TS1 and TS2 (21–23). The PDT is recorded when the VAS slider is moved from 0 to 0.1 cm. In addition to evaluating CPM using PDT, we evaluated CPM effect as a change in pain rating at Pain50 between TS1 and TS2 (similarly to the heat paradigm). To determine these pain ratings from the continuous pain ratings, first we found the initial pressure during TS1 that evoked a pain intensity rating of 50/100 (Pain50, indicated when the slider was at 5.0 cm along the 10 cm length). Then the TS2 pain rating used to calculate the CPM effect was the pain intensity rated when the TS2 pressure was at Pain50.

In each participant, prior to the CPM test, a threshold paradigm was used three times (separated by 1 min) for familiarization, determining the CS pressure, and to deliver the TS1 (**Figure 1**). In each trial, the pressure in one cuff continuously increased at a rate of 1 kPa/s and the participant used a VAS slider to continuously rate the evoked pain intensity. The VAS slider scale was labeled with words and numbers; "No pain" at 0 and "Most intense pain imaginable" at 100. The participants were instructed to press a button on the slider when they reached their PTT; pressing this button then deflated the cuff. The maximum cuff pressure allowable was 100 kPa, with the cuff automatically deflating if it reached this level. The first trial on the right limb was only used to familiarize the participant with the protocol. For the second trial, the stimulus was delivered to the opposite limb. The software for the pressure system set the CS pressure level for the CPM test at 70% of the PTT from the second trial. The third trial on the right limb was the TS1.

The CPM assessment consisted of determining the TS2 pain on the right limb in the presence of a concurrent CS delivered to the opposite limb (**Figure 1**). To do this, the CS rapidly increases to the set pressure and is held at that level for 100 s. At the same time the TS (TS2) pressure gradually increased with the same protocol given as the TS1 test (participants continuously rated the pain intensity evoked by the TS until the TS reached their PTT where they then press the button). Pressing the button deflates both cuffs. This paradigm takes ∼10 min.

Statistical Analyses

All correlation and group statistical analyses were performed using R (version 3.6.0; https://www.r-project.org) in RStudio (version 1.0.44; https://www.rstudio.com). GraphPad Prism (version 7.03, https://www.graphpad.com) was used to create the figures and Microsoft Excel (version 2010; microsoft.com/excel) was used for some descriptive statistics (mean and standard deviation). The Shapiro-Wilk test was used to assess the

normality of the data distribution that is required to subsequently run a parametric 2-tailed test. If the distribution passed the normality test a paired *t*-test (*t* statistic) was used to evaluate within-subject differences, otherwise the Wilcoxon signed rank test (*W* statistic) was used. For the paired *t*-test analyses the common measure of effect size Cohen's *d* is reported which had the same conclusion (small, medium, or large effect size) when assessed using Hedges' g (24). The effect size for the Wilcoxon signed rank test analyses is *r* (*z* statistic divided by the square root of the sample size) (25). For between group comparisons if both groups were normally distributed, an independent *t*-test (welch two-sample *t*-test in *R*, *t* statistic) was used otherwise the Mann-Whitney Wilcoxon test (*W* statistic) was used. The comparison of sex difference proportions between the healthy and chronic pain group was assessed using two-proportions *z*-test (X^2 statistic). Data passed the Shapiro-Wilk normality test before being correlated using the Pearson correlation test (*r* statistic). In the results section, bracketed values followed by a ± symbol represent the mean ± the standard deviation.

RESULTS

Demographics and Descriptive Statistics

Data were collected from a total of 30 participants (18 healthy controls, 12 people with chronic pain). There were no significant differences in the proportion of females and males across the healthy group (9F, 9M) and the chronic pain group (7F, 5M) ($X^2 = 0.006$, $p > 0.05$). However, there was a significant difference in mean age across the chronic pain group (55.3 ± 15.6 years old) and healthy control group (31.8 ± 11.0 years old; $W = 26$, $p < 0.01$). The average BDI scores were also significantly higher ($W = 39$, $p < 0.01$) in the chronic pain group (9.7 ± 5.1) compared to the healthy group (3.9 ± 4.0).

Of the three CPM measures, the heat CPM and the PDT pressure CPM were not significantly different between the healthy and chronic pain group ($p < 0.05$). However, the pressure CPM (at Pain50) was significantly more inhibitory in the healthy group compared to the chronic pain group ($t = -2.23$, $p = 0.04$). Additional descriptive statistics for each CPM paradigm and group can be found in **Table 1**. The following result sections highlight within-subject comparisons.

TABLE 1 | Group demographics and CPM descriptive statistics.

Variable	Healthy group	Chronic pain group
N (F, M)	18 (9, 9)	12 (7, 5)
Age (Y)	31.8 ± 11.0*	55.3 ±15.6*
BDI	3.9 ± 4.0*	9.7 ± 5.1*
PDT Pressure CPM Effect (% change)	−14.6 ± 32.4	−26.6 ± 48.7
PDT Pressure CPM Effect (absolute change)	−3.1 ± 6.6	−3.6 ± 9.3
TS1 pressure (kPa)	22.3 ± 9.7	22.7 ± 9.0
TS2 pressure (kPa)	25.4 ± 13.4	26.3 ± 8.8
Pain50 Pressure CPM effect (% change)	−50.1 ± 33.0*	−21.4 ± 35.8*
Pain50 Pressure CPM effect (absolute change)	−25.2 ± 16.6	−10.8 ± 17.9
TS1 pressure pain rating (0–100)	50.3 ± 0.4	50.2 ± 0.4
TS2 pressure pain rating (0–100)	25.1 ± 16.5	39.4 ± 17.8
TS Pain50 pressure (kPa)	36.9 ± 14.1	46.8 ± 15.0
CS pressure (70% PTT) (kPa)	33.7 ± 12.5	50.3 ± 12.9
CS pressure Pain50 (kPa)	34.2 ± 13.4	47.9 ± 14.4
Pain50 Heat CPM effect (% change)	−6.5 ± 32.3	−11.1 ± 33.6
Pain50 Heat CPM effect (absolute change)	−3.6 ± 16.4	−6.5 ± 16.4
TS1 heat pain rating (0–100)	50.5 ± 9.3	49.6 ± 16.3
TS2 heat pain rating (0–100)	46.9 ± 18.5	43.1 ±17.9
TS Pain50 temperature (°C)	45.7 ± 1.4	44.6 ± 3.0
CS Pain50 temperature (°C)	45.4 ± 1.4	44.3 ± 2.4
painDetect score (NNP, MNP, NP)	NA	19.7 (2, 2, 8) ± 8.5
BPI Pain Severity score	NA	6.3 ± 1.0
BPI Interference score	NA	6.0 ± 1.8

Group data are shown as mean ± standard deviation. N, Number of participants; F, Female; M, Male; BDI, Beck Depression Inventory; PDT, Pain Detection Threshold; TS1, first test stimulus; TS2, second test stimulus; CS, conditioning stimulus; Pain50, stimulus evoking pain rating of 50/100; CPM, conditioned pain modulation; PTT, pain tolerance threshold; NP, Neuropathic Pain; MNP, Mixed-NP; NNP, non-NP; BPI, Brief Pain Inventory. Note that the CS pressure used during pressure-based CPM was at 70% PTT. Asterisks denote statistically significant difference between healthy and chronic pain group (p < 0.05).

Relationship Between CPM Effect and Stimulus Modality

The TS pain ratings during the heat paradigm were only collected at Pain50. Therefore, the following comparisons of CPM between the heat and pressure paradigm are all from CPM calculated as a change in TS pain ratings at Pain50; with inhibitory CPM being a negative % and facilitatory CPM being a positive %.

Overall, in the healthy individuals, the CPM effect was significantly different between heat and pressure paradigms (*t*-test: $t = -3.41$, $p = 0.004$, Cohen's $d = -1.34$; see **Figure 2**); where the CPM effect in the pressure paradigm (−50.1 ± 33.0%) is on average more inhibitory compared to the heat paradigm (−6.5 ± 32.3%). In the chronic pain group, the CPM effect in the pressure paradigm on average was more inhibitory (−21.4 ± 35.8%) than the heat paradigm (−11.1 ± 33.6%), but there was no significant difference between modalities at the individual level (*t*-test: $t = -1.05$, $p = 0.32$, Cohen's $d = -0.30$; **Figure 2**).

The within-individual data plots in **Figure 2** reveal that subjects either exhibited the same (i.e., modality independent) or opposite (i.e., modality dependent) type of CPM effect in the heat and pressure paradigms. Overall, CPM in most of the healthy individuals was modality-dependent but most of the individuals with chronic pain had modality-independent CPM effects (**Figure 3**). In the healthy group, only five individuals exhibited modality-independent CPM (inhibitory CPM effect regardless of paradigm). However, modality-dependent CPM effects were found for nine individuals: eight had exhibited a facilitatory heat CPM effect and an inhibitory pressure CPM effect while one had an inhibitory heat CPM effect and a facilitatory pressure CPM effect. Four individuals did not exhibit CPM from the heat paradigm but had an inhibitory CPM effect

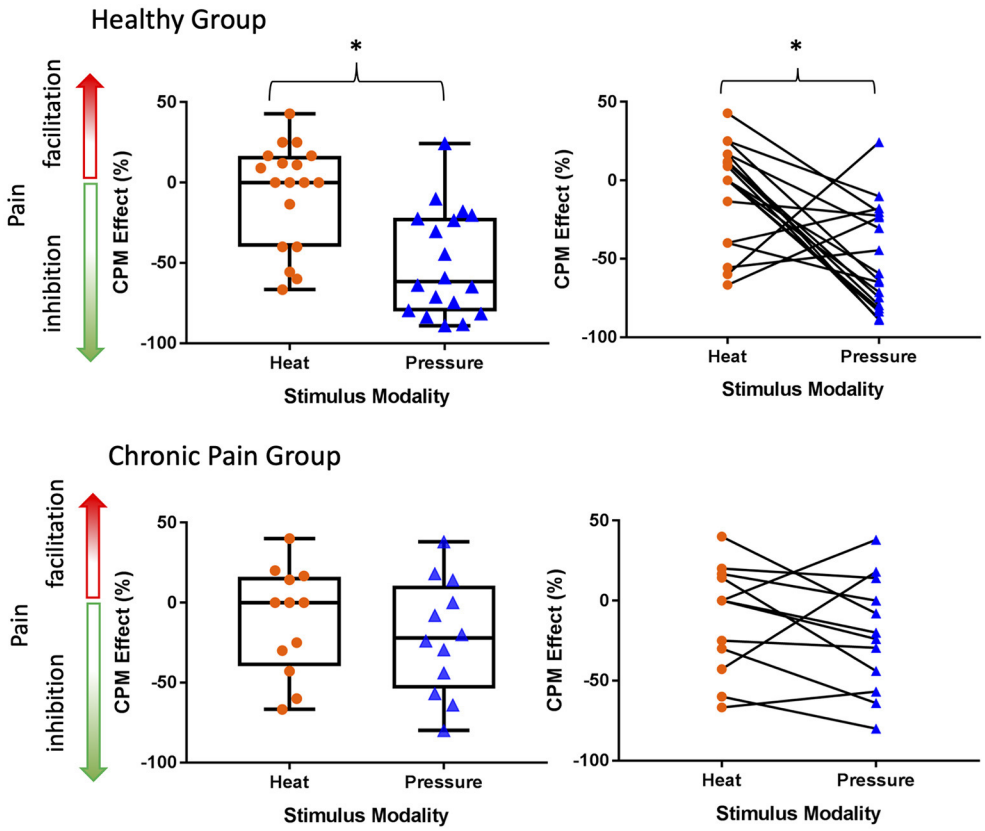


FIGURE 2 | The CPM effect only significantly differs between stimulus modalities in the healthy group. In both the healthy group and chronic pain group the inhibitory CPM effect is more pronounced using pressure than heat stimuli. Line within box is the median, edges of box are 25th to 75th percentiles, and the whiskers denote the maximum and minimum. Asterisks denote significant difference within-subjects ($p < 0.05$).

from the pressure paradigm. There was no clear modality-dependent pattern for healthy males and females (**Figure 3**).

In the chronic pain group, five individuals had modality-independent CPM effect: four with inhibitory and one with facilitatory CPM effects. Amongst the chronic pain group, three individuals had a modality-dependent CPM effect: two with facilitatory heat CPM and inhibitory pressure CPM, while one had inhibitory heat CPM and facilitatory pressure CPM. Three individuals exhibited no heat CPM, of which two had inhibitory pressure CPM and one had facilitatory pressure CPM. One individual with chronic pain exhibited no pressure CPM with facilitatory heat CPM. There was no clear modality-dependent pattern for males and females with chronic pain (**Figure 3**).

Pressure CPM Calculated Using Pain50 Pain Ratings vs. PDT

A within-subject analysis was used to assess the difference in CPM effects based on a change in pressure PDT vs. TS pain intensity ratings at Pain50 (**Figure 4**). One healthy participant was deemed to be an outlier in terms of their PDT and was excluded from this analysis because they had an extremely low TS1 PDT that was not consistent with other participants or with their own pain thresholds responses from other trials during their psychophysical testing, and thus likely was due to attentional or other effects. Within the healthy participant group, the CPM effect was significantly different between these two measures of CPM ($t = -4.76$, $p = 0.0002$, Cohen's $d = -1.03$). Specifically, the healthy participants exhibited a significantly more inhibitory CPM effect ($-48.4 \pm 33.1\%$) when measured as Pain50 CPM compared to PDT CPM ($-14.6 \pm 32.4\%$). In contrast, in the chronic pain individuals there was no significant difference ($t = 0.46$, $p = 0.66$, Cohen's $d = 0.12$) in the Pain50 CPM ($-21.4, \pm 35.8\%$) compared to the PDT CPM ($-26.6 \pm 48.7\%$).

Difference Between CS Pressure at Pain50 and 70% PTT

The standard software that drives the NociTech Inc. pressure cuff system sets the CS pressure at 70% PTT for the CPM test. While we compared the CPM heat and pressure paradigms with TS pain intensity ratings at Pain50, the CS in the heat paradigm was set to Pain50 while in the pressure paradigm the CS was set to 70% PTT. Therefore, this analysis assesses whether the pressure at Pain50 was significantly different from the 70% PTT used to set the CS.

The healthy group data was not normally distributed ($W = 0.87$, $p = 0.02$) and so was assessed using non-parametric statistics. This indicated that there was no significant difference

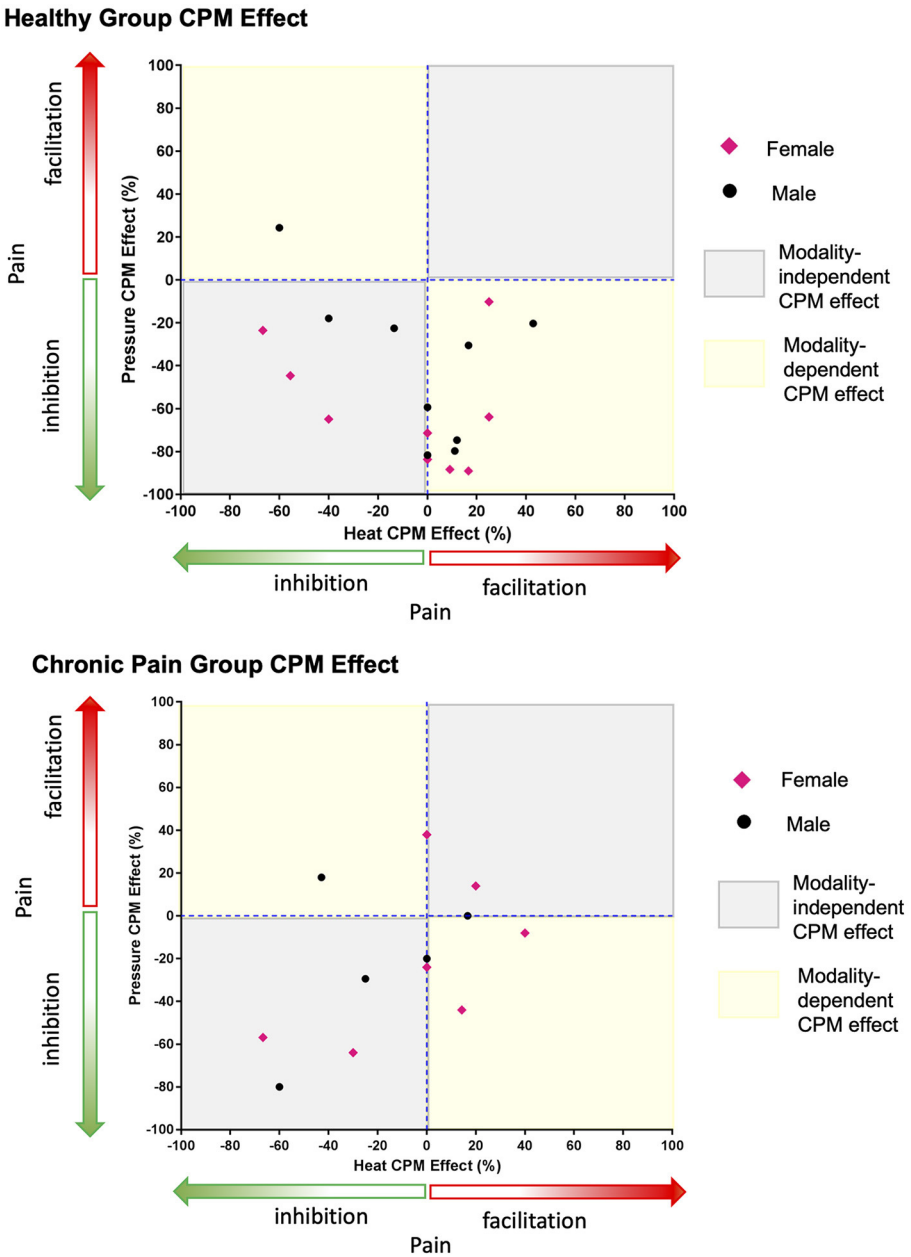


FIGURE 3 | Greater number of healthy individuals demonstrate a modality-dependent CPM effect compared to individuals with chronic pain. Points that fall in the modality-dependent yellow quadrants reflect individuals who could be classified on the opposite ends of the CPM effect spectrum (facilitatory vs. inhibitory) depending on whether the paradigm is heat or pressure based. Modality-independent gray quadrants reflect individuals who would have the same type of CPM effect regardless of the paradigm.

(Wilcoxon signed rank test: $V = 105$, $p = 0.42$, r effect size $= 0.20$) between the CS pressure at Pain50 (34.2 ± 13.4 kPa, median $= 35.9$ kPa) and the CS pressure at 70% PTT (33.7 ± 12.5 kPa, median $= 34.6$ kPa). Similarly, in the chronic pain group, there was no significant difference ($t = -0.68$, $p = 0.51$, Cohen's $d = -0.18$) between the CS pressure at Pain50 (47.9 ± 14.4 kPa, median $= 49.6$ kPa) and the CS pressure at 70% PTT (50.3 ± 12.9 kPa, median $= 52.7$ kPa). For comparisons purposes this was also checked using the non-parametric test, and a similar conclusion was found using the Wilcoxon signed rank test ($V = 35$, $p = 0.79$).

Relationship Between CPM and Clinical Pain Parameters

The relationships between CPM and measures of clinical pain (i.e., painDETECT, BPI pain interference, and BPI pain severity)

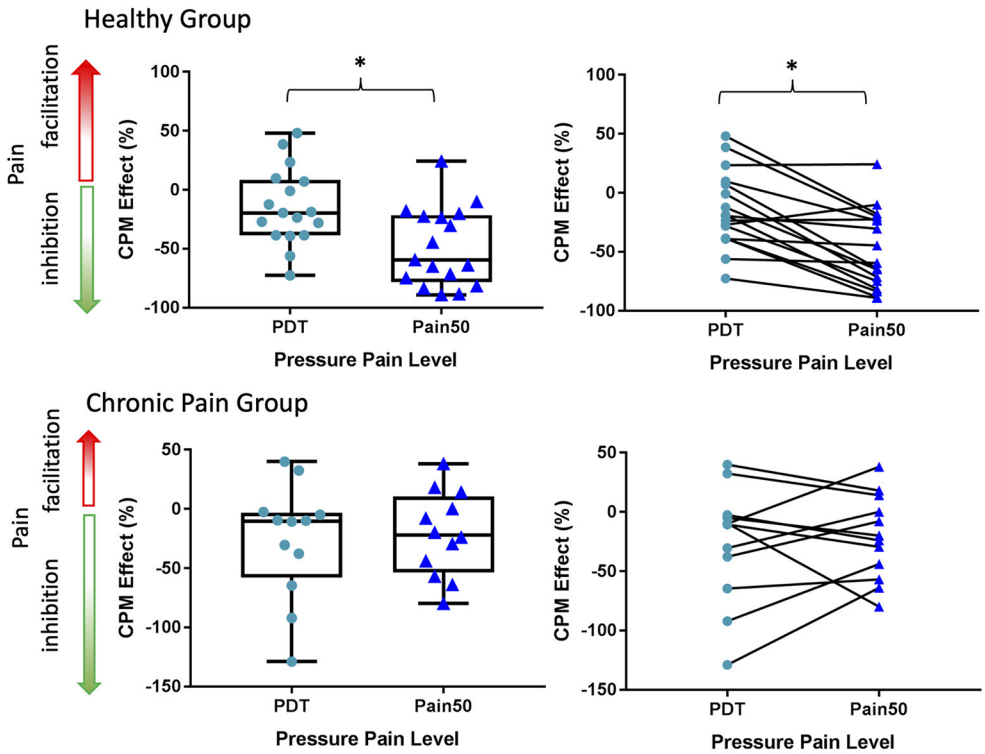


FIGURE 4 | A comparison of CPM effects based on changes in Pain50 vs. the PDT. The CPM effect at PDT and Pain50 only significantly differs in the healthy group. In both the healthy and chronic pain group the inhibitory CPM effect is more pronounced when calculated as a change in pain ratings at Pain50 than a change in the pressure pain detection threshold (PDT). Pain50 is the stimulus intensity that evokes a pain rating of 50/100, where 0 is no pain and 100 is the worst pain imaginable. Line within box is the median, edges of box are 25th to 75th percentiles, and the whiskers denote the maximum and minimum. Asterisks denote significant difference within-subjects ($p < 0.05$).

are shown in **Figure 5**. In general, both the heat-based ($r = -0.55$), and pressure-based ($r = -0.32$) CPM effects were not significantly correlated with painDETECT scores (heat: $t = -2.06$, $p = 0.066$; pressure: $t = -1.10$, $p = 0.32$). In addition, the heat-based CPM ($r = -0.28$) and pressure-based CPM ($r = -0.28$) were not significantly correlated with BPI average interference scores (heat: $t = -0.93$, $p = 0.38$; pressure: $t = -0.92$, $p = 0.38$). Similarly, the heat-based CPM ($r = -0.23$) and pressure-based CPM ($r = -0.21$) were not significantly correlated with BPI average pain severity scores (heat: $t = -0.73$, $p = 0.48$; pressure: $t = -0.68$, $p = 0.51$). Overall, all the clinical pain parameters showed non-significant trends toward being negatively correlated with both CPM paradigms, with the heat-based CPM and painDETECT scores showing the strongest correlation that approached significance ($r = -0.55$, $p = 0.066$).

DISCUSSION

This study sought to interrogate paradigms that could potentially be used to evaluate CPM consistently and easily for research and clinical investigations. Toward this goal, our study aim was to compare heat and pressure paradigms that can be used to evaluate CPM in healthy individuals and those with chronic pain. Our main findings were that 1) in healthy individuals, the pressure-based paradigm produced a stronger inhibitory CPM compared to the heat-based paradigm, and most participants exhibited modality-dependent CPM effects (i.e., inhibitory vs. facilitatory), 2) in people with chronic pain, there was no significant difference in the CPM evoked by the pressure- and heat-based paradigms, with the majority exhibiting modality-independent CPM effects, 3) the pressure paradigm evoked a similar (in chronic pain individuals) or more inhibitory CPM effect (in healthy individuals) when calculating CPM based on the Pain50 level compared to the pain detection threshold (PDT), 4) The healthy group had a significantly more inhibitory CPM than the chronic pain group in the pressure paradigm but not the heat paradigm. Overall, our findings indicate that the interpretation of CPM effects in healthy individuals needs to consider the stimulus modality and metric used to calculate CPM. Importantly though, given that the heat and pressure CPM paradigms evoke similar CPM magnitudes in chronic pain, our findings provide support for the adoption of a faster, more robust and more easily administered pressure-based paradigm to assess CPM in clinical populations and in research environments.

Overall, across subjects, the average pain evoked by a pressure stimulus induced a more inhibitory CPM effect compared to pain evoked by a heat stimulus. Possible factors underlying this

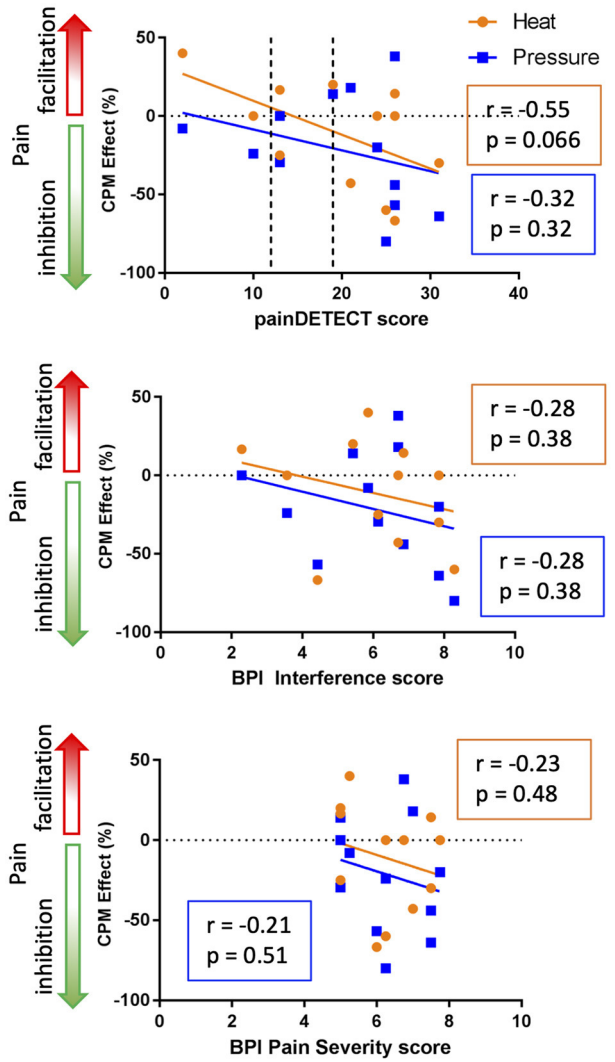


FIGURE 5 | Relationship between clinical pain and CPM. The top graph indicates that there is a trend toward greater inhibitory heat-based CPM effects in those people with higher levels of neuropathic pain based on PainDetect ($r = -0.55$, $p = 0.066$). In the painDETECT questionnaire scores 12 or lower to indicate non-neuropathic pain, 13–18 mixed neuropathic pain, and 19 or greater indicates the presence of neuropathic pain is likely (categories indicated by the dotted lines). The middle and bottom graphs depict the relationship between CPM and the Brief Pain Inventory (BPI) interference and pain severity scores. The BPI interference and pain severity scores are negatively correlated with CPM effect but neither of these correlations are significant ($p < 0.05$).

finding could include that the pressure stimulus can recruit A-beta fibers and the pressure cuffs can induce additional recruitment of nociceptors from the deep-somatic stimulus and from the larger surface area compared to thermodes and this could have triggered a greater response of the descending anti-nociceptive control system (26–28). Curiously though, we found that it was only in the healthy individuals that the pressure paradigm induced a significantly more robust inhibitory CPM effect than the heat paradigm, whereas there was no significant difference between the CPM effect induced by these paradigms in individuals with chronic pain. This suggests that the assumption that a similar CPM effect can be achieved by any stimulus modality holds true for those with chronic pain but not healthy individuals.

A large systematic review of CPM studies highlighted that regardless of stimulus modality (electrical, mechanical, etc.) the majority of studies did not find any significant correlations between CPM and clinical attributes and manifestations of pain (e.g., pain intensity, severity disability, interference, and duration) (29). This aligns with our findings that BPI and PainDetect scores did not significantly correlate with CPM from the pressure paradigm. Interestingly though, this systematic review noted an exception to their overall conclusion was for the studies that used thermal TS and CS, 55% of these studies had negative significant correlations between CPM and clinical pain (29). They also found across all the studies that the clinical manifestations of pain and the CPM effect were not significantly correlated in any of the studies in idiopathic pain and most of the studies with nociceptive pain, although, about half of the studies of neuropathic pain did report a significant correlation between CPM and pain (29). Taken together these findings support our finding that the strongest correlation of heat-based CPM with a clinical pain measure was with the painDETECT scores (an assessment of neuropathic pain). A link between neuropathic pain and CPM particularly in heat-based paradigm but not a pressure paradigm, was also recently found in a study of CPM effects using heat and pressure test paradigms in people with painful and non-painful diabetic polyneuropathy (DPN) (30). This study reported that the inhibitory CPM evoked by heat was significantly correlated with greater neuropathic pain (30). This suggests that in our population of people with chronic pain, the severity of their clinical pain and its interference on activities of daily living may not greatly impact their ability to exhibit an experimentally-induced CPM. However, the neuropathic nature of their pain may impact the heat-based CPM, but not the pressure-based CPM. This finding suggests that the experimental pressure-based CPM test can be used to assess CPM ability in patients with chronic pain regardless of the severity, neuropathic aspect, and interference of their chronic pain.

Despite CPM not necessarily being related to the magnitude of clinical pain, it nonetheless holds great potential clinical utility because a weaker inhibitory CPM in individuals with chronic pain can distinguish them from healthy individuals with stronger inhibitory CPM (13, 29). This is thought to be due to individuals with weak inhibitory CPM being more at risk to develop chronic pain and/or once chronic pain develops it exhausts the pain inhibition capacity leading a weaker inhibitory CPM than healthy individuals (2). This aligns with our findings in the pressure paradigm where the healthy group had a significantly more inhibitory CPM than the chronic pain group, however this was not the case in the heat-paradigm. Clinical translation of a CPM paradigm could guide treatment selection for individuals in that it could assess the risk of post-operative chronic pain and predict treatment efficacy (4). Since CPM in the chronic pain individuals was not significantly different between pressure and

heat-based paradigms, either could be translated to the clinics. The main limitation of translating the pressure cuff-based CPM paradigm is that it cannot be used on those with cardiovascular health concerns. However, the limitation of the heat paradigm is that despite the familiarization and habituation step used to find Pain50 in the heat paradigm, often the first TS1 is still not rated at a pain intensity of 50/100 and therefore extra time was then needed to re-calibrate the Pain50 before running the CPM protocol. In the pressure paradigm this was never necessary because the pain tolerance threshold could be reached in under 100s, we often could run the paradigm in under 10 min. An additional practical challenge with the heat paradigm is that care had to be taken so that the thermode lays flat on the skin, but the straps are not too tight to avoid pressure confounding the pain intensity rating of the heat stimulus. Thus, overall, the pressure paradigm was faster to run, simpler to administer, and the equipment is somewhat more straightforward, and thus may be more adaptable to a clinical setting.

One limitation of comparing a heat to a pressure paradigm is that elements of the protocols were different for the two approaches which limits an exact head-to-head comparison. For example, the heat paradigm was based on the administering rapidly increasing intensity of stimuli for blocks of time and asking a participant to provide a single rating of the evoked pain intensity. In contrast, in the pressure paradigm, stimuli were delivered continuously at a slowly increasing intensity and participants provided continuous ratings of pain intensity. However, it is of note that in the pressure protocol, we calculated CPM using the change in pain rated at Pain50 rather than the change in pressure pain detection thresholds that has been used by some research labs (21–23). We chose this approach so that we could more directly compare the pressure paradigm to heat paradigm which used Pain50 to avoid floor effects (20, 31). Additionally, much of what is known of the underlying mechanisms of CPM are based on animal studies of DNIC that used suprathreshold stimuli, and this is similar to using the suprathreshold metric of Pain50 stimuli. In contrast a determination of CPM as a threshold change (i.e., PDT) may be due to a different underlying mechanism. For example, a change in pain evoked by a suprathreshold stimulus may be more reflective of a mechanism related to hyperalgesia rather than a measure of a threshold change which may be more akin to an allodynia mechanism. Although assessing pressure CPM at Pain50 is a departure from the approach used by other labs, we note that in individuals with chronic pain there was no significant difference in pressure CPM calculated using Pain50 vs. PDT. However, in healthy individuals using Pain50 resulted in a significantly stronger inhibitory CPM compared to using PDT. Unexpectantly, the mean Pain50 pressure was higher in the chronic pain group than in the healthy control group. The reason for this is not clear given that the PDT pressures were similar in the healthy and chronic pain groups, suggesting that sensory loss is not a factor. However, healthy individuals unfamiliar with deep intense pain may be more sensitive to the pressure stimulus and thus have a lower Pain50 pressure compared to those with chronic pain whose ratings to the stimulus may be perceived as relatively less painful given their experience from their condition.

The relatively small sample size and diversity in types of chronic pain conditions in this study may have precluded the ability to detect significant correlations between clinical measures and CPM in the chronic pain group. This sample size limitation also did not allow us to carry out an extensive sex and age effect analysis in either group. Although the proportion of males and females in our study was not significantly different between groups, the healthy group was significantly younger than the chronic pain group. Interestingly, a recent study in healthy individuals found the age effect was larger than the sex effect; across all the different stimulus modality paradigms younger males had the strongest inhibitory CPM (32). This aligns with previous findings showing that younger and healthy individuals on average have a stronger inhibitory CPM than older individuals and those with chronic pain (13). Here in our study as well, the somewhat younger healthy group had a significantly more inhibitory CPM than the older chronic pain group in the pressure paradigm. Therefore, the CPM we found to be evoked at Pain50 in the pressure paradigm was more consistent with age and health effects in the literature than the heat paradigm.

In conclusion, this study found that a pressure-based CPM paradigm evoked an inhibitory CPM effect in most participants (healthy individuals and those with chronic pain), but was significantly stronger than the heat-based CPM only in the healthy individuals. Similarly, CPM in the pressure-based CPM paradigm at Pain50 was significantly more inhibitory than at pressure pain detection threshold in healthy individuals. In individuals with chronic pain the paradigm type is not critical because the pressure CPM does not significantly differ from heat CPM and PDT CPM. Given our finding that the pressure paradigm may be less impacted from clinical neuropathic pain and can be carried out using a relatively simple system with a shorter test time than heat-based paradigm, we propose that the pressure-based CPM paradigm offers a good option for use in research studies and is particularly well-suited for clinical use, potentially to aid in assessing predictive factors of treatment outcome.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University Health Network Research Ethics Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KD and RE-S conceived this study. Data collected from participants with chronic pain was part of a larger study of

treatment effects, clinically led by ABh. RE-S collected data with the help from CF, ABe, EM, JK, SF, and NO. All authors contributed to the drafting of the manuscript and approved the final version.

FUNDING

This research was supported by the Mayday Fund, and by the Canadian Institute of Health Research (operating grant PJT-156409 to KD) and Strategy for Patient-Oriented Research (SPOR) funding of the Canadian Chronic Pain Network (SCA-145102).

ACKNOWLEDGMENTS

The author thanks Jamal Kara for his help in patient recruitment and also thanks Dr. Lee Kisler for help in data collection and in the early development of the project and thank Dr. Mills for serving as a model for the drawing of the pressure cuffs in **Figure 1**. Finally, the author thanks the healthy participants and those with chronic pain who participated in this study.

REFERENCES

1. Pud D, Granovsky Y, Yarnitsky D. The methodology of experimentally induced diffuse noxious inhibitory control (DNIC)-like effect in humans. *Pain.* (2009) 144:16–9. doi: 10.1016/j.pain.2009.02.015
2. Yarnitsky D. Role of endogenous pain modulation in chronic pain mechanisms and treatment. *Pain.* (2015) 156(Suppl. 1):S24–31. doi: 10.1097/01.j.pain.0000460343.46847.58
3. Yarnitsky D, Granot M, Nahman-Averbuch H, Khamaisi M, Granovsky Y. Conditioned pain modulation predicts duloxetine efficacy in painful diabetic neuropathy. *Pain.* (2012) 153:1193–8. doi: 10.1016/j.pain.2012.02.021
4. Yarnitsky D, Crispel Y, Eisenberg E, Granovsky Y, Ben-Nun A, Sprecher E, et al. Prediction of chronic post-operative pain: pre-operative DNIC testing identifies patients at risk. *Pain.* (2008) 138:22–8. doi: 10.1016/j.pain.2007.10.033
5. Yarnitsky D, Arendt-Nielsen L, Bouhassira D, Edwards RR, Fillingim RB, Granot M, et al. Recommendations on terminology and practice of psychophysical DNIC testing. *Eur J Pain.* (2010) 14:339. doi: 10.1016/j.ejpain.2010.02.004
6. Yarnitsky D. Conditioned pain modulation (the diffuse noxious inhibitory control-like effect): its relevance for acute and chronic pain states. *Curr Opin Anaesthesiol.* (2010) 23:611–5. doi: 10.1097/ACO.0b013e32833c348b
7. Bars D, Dickenson A, Besson J. Diffuse noxious inhibitory controls (DNIC). II. Lack of effect on non-convergent neurones, supraspinal involvement and theoretical implications. *Pain.* (1979) 6:305–27. doi: 10.1016/0304-3959(79)90050-2
8. Le Bars D, Dickenson A, Besson J-M. Diffuse noxious inhibitory controls (DNIC). I. Effects on dorsal horn convergent neurones in the rat. *Pain.* (1979) 6:283–304. doi: 10.1016/0304-3959(79)90049-6
9. Villanueva L, Bouhassira D, Le Bars D. The medullary subnucleus reticularis dorsalis (SRD) as a key link in both the transmission and modulation of pain signals. *Pain.* (1996) 67:231–40. doi: 10.1016/0304-3959(96)03121-1
10. Villanueva L, Chitour D, Le Bars D. Involvement of the dorsolateral funiculus in the descending spinal projections responsible for diffuse noxious inhibitory controls in the rat. *J Neurophysiol.* (1986) 56:1185–95. doi: 10.1152/jn.1986.56.4.1185
11. Firouzian S, Osborne NR, Cheng JC, Kim JA, Bosma RL, Hemington KS, et al. Individual variability and sex differences in conditioned pain modulation and the impact of resilience, and conditioning stimulus pain unpleasantness and salience. *Pain.* (2020) 161:1847–60. doi: 10.1097/j.pain.0000000000001863
12. Potvin S, Marchand S. Pain facilitation and pain inhibition during conditioned pain modulation in fibromyalgia and in healthy controls. *Pain.* (2016) 157:1704–10. doi: 10.1097/j.pain.0000000000000573
13. Lewis GN, Rice DA, McNair PJ. Conditioned pain modulation in populations with chronic pain: a systematic review and meta-analysis. *J Pain.* (2012) 13:936–44. doi: 10.1016/j.jpain.2012.07.005
14. Graven-Nielsen T, Wodehouse T, Langford RM, Arendt-Nielsen L, Kidd BL. Normalization of widespread hyperesthesia and facilitated spatial summation of deep-tissue pain in knee osteoarthritis patients after knee replacement. *Arthritis Rheum.* (2012) 64:2907–16. doi: 10.1002/art.34466
15. Kosek E, Ordeberg G. Lack of pressure pain modulation by heterotopic noxious conditioning stimulation in patients with painful osteoarthritis before, but not following, surgical pain relief. *Pain.* (2000) 88:69–78. doi: 10.1016/S0304-3959(00)00310-9
16. Wilder-Smith OH, Schreyer T, Scheffer GJ, Arendt-Nielsen L. Patients with chronic pain after abdominal surgery show less preoperative endogenous pain inhibition and more postoperative hyperalgesia: a pilot study. *J Pain Palliat Care Pharmacother.* (2010) 24:119–28. doi: 10.3109/15360281003706069
17. Landau R, Kraft JC, Flint LY, Carvalho B, Richebe P, Cardoso M, et al. An experimental paradigm for the prediction of post-operative pain (PPOP). *J Vis Exp.* (2010) 35:1671. doi: 10.3791/1671
18. Arendt-Nielsen L, Larsen JB, Rasmussen S, Krogh M, Borg L, Madeleine P. A novel clinical applicable bed-side tool for assessing conditioning pain modulation: proof-of-concept. *Scand J Pain.* (2020) 20:801–7. doi: 10.1515/sjpain-2020-0033
19. Yarnitsky D, Granot M, Granovsky Y. Pain modulation profile and pain therapy: between pro- and antinociception. *Pain.* (2014) 155:663–5. doi: 10.1016/j.pain.2013.11.005
20. Yarnitsky D, Bouhassira D, Drewes AM, Fillingim RB, Granot M, Hansson P, et al. Recommendations on practice of conditioned pain modulation (CPM) testing. *Eur J Pain.* (2015) 19:805–6. doi: 10.1002/ejp.605
21. Kennedy DL, Kemp HI, Wu C, Ridout DA, Rice ASC. Determining real change in conditioned pain modulation: a repeated measures study in healthy volunteers. *J Pain.* (2020) 21:708–21. doi: 10.1016/j.jpain.2019.09.010
22. McPhee M, Graven-Nielsen T. Alterations in temporal summation of pain and conditioned pain modulation across an episode of experimental exercise-induced low back pain. *J Pain.* (2019) 20:264–76. doi: 10.1016/j.jpain.2018.08.010
23. Graven-Nielsen T, Izumi M, Petersen KK, Arendt-Nielsen L. User-independent assessment of conditioning pain modulation by cuff pressure algometry. *Eur J Pain.* (2017) 21:552–61. doi: 10.1002/ejp.958
24. Ho J, Tumkaya T, Aryal S, Choi H, Claridge-Chang A. Moving beyond P values: data analysis with estimation graphics. *Nat Methods.* (2019) 16:565–6. doi: 10.1038/s41592-019-0470-3
25. Maciej T, Ewa T. The need to report effect size estimates revisited. 652 An overview of some recommended measures of effect size. *Trends Sport Sci.* (2014) 1(21):19–653.
26. Imai Y, Petersen KK, Morch CD, Arendt Nielsen L. Comparing test-retest reliability and magnitude of conditioned pain modulation using different combinations of test and conditioning stimuli. *Somatosens Mot Res.* (2016) 33:169–77. doi: 10.1080/08990220.2016.1229178
27. Keay K, Bandler R. Parallel circuits mediating distinct emotional coping reactions to different types of stress. *Neurosci Biobehav Rev.* (2001) 25:669–78. doi: 10.1016/S0149-7634(01)00049-5
28. Melzack R, Wall DP. Pain mechanisms a new theory. *Science.* (1965) 150:971–9. doi: 10.1126/science.150.3699.971

29. Fernandes C, Pidal-Miranda M, Samartin-Veiga N, Carrillo-de-la-Pena MT. Conditioned pain modulation as a biomarker of chronic pain: a systematic review of its concurrent validity. *Pain.* (2019) 160:2679–90. doi: 10.1097/j.pain.0000000000001664
30. Granovsky Y, Shafran Topaz L, Laycock H, Zubiedat R, Crystal S, Buxbaum C, et al. Conditioned pain modulation is more efficient in patients with painful diabetic polyneuropathy than those with nonpainful diabetic polyneuropathy. *Pain.* (2021). doi: 10.1097/j.pain.0000000000002434. [Epub ahead of print].
31. Honigman L, Bar-Bachar O, Yarnitsky D, Sprecher E, Granovsky Y. Nonpainful wide-area compression inhibits experimental pain. *Pain.* (2016) 157:2000–11. doi: 10.1097/j.pain.0000000000000604
32. Riley JL III, Cruz-Almeida Y, Staud R, Fillingim RB. Age does not affect sex effect of conditioned pain modulation of pressure and thermal pain across 2 conditioning stimuli. *Pain Rep.* (2020) 5:e796. doi: 10.1097/PR9.0000000000000796

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pain Research

ORIGINAL RESEARCH
published: 22 February 2022
doi: 10.3389/fpain.2021.809351

Parental Pain Catastrophizing, Communication Ability, and Post-surgical Pain Outcomes Following Intrathecal Baclofen Implant Surgery for Patients With Cerebral Palsy

Breanne J. Byiers[1], Caroline L. Roberts[1], Chantel C. Burkitt[1,2], Alyssa M. Merbler[1], Kenneth D. Craig[3] and Frank J. Symons[1]*

[1] Department of Educational Psychology, University of Minnesota, Minneapolis, MN, United States, [2] Gillette Children's Specialty Healthcare, Saint Paul, MN, United States, [3] Department of Psychology, University of British Columbia, Vancouver, BC, Canada

OPEN ACCESS

Edited by:
Mathieu Piché, Université du Québec à Trois-Rivières, Canada

Reviewed by:
Adrienne Harvey, Royal Children's Hospital, Australia
Emma Fisher, University of Bath, United Kingdom

***Correspondence:**
Breanne J. Byiers
byier001@umn.edu

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *04 November 2021*
Accepted: *29 December 2021*
Published: *22 February 2022*

Citation:
Byiers BJ, Roberts CL, Burkitt CC, Merbler AM, Craig KD and Symons FJ (2022) Parental Pain Catastrophizing, Communication Ability, and Post-surgical Pain Outcomes Following Intrathecal Baclofen Implant Surgery for Patients With Cerebral Palsy. Front. Pain Res. 2:809351. doi: 10.3389/fpain.2021.809351

There is strong evidence that psychosocial variables, including pain catastrophizing, influence parental and child ratings of pain, pain expression, and long-term outcomes among children with chronic pain. The role of these factors among children who have communication deficits due to cerebral palsy (CP) and other intellectual and developmental disabilities is currently unclear. In this study, parental pain catastrophizing was assessed before intrathecal baclofen (ITB) pump implantation for spasticity management in 40 children and adolescents with CP, aged 4 to 24 years. Pain was assessed before and after surgery with two methods: a parent-reported pain interference scale, and behavioral pain signs during a standardized range of motion exam. Linear mixed models with clinical/demographic factors and scores from the Pain Catastrophizing Scale for Parents (PCS-P), and child spoken language ability as predictors and the pain variables as the outcomes were implemented. On average, both pain outcomes improved after surgery. Only child spoken language ability predicted change in behavioral reactivity scores, with children with phrase speech showing an increase in reactivity at follow-up compared to pre-surgery levels, on average. A significant interaction between PCS-P scores and spoken language ability on change in pain interference scores over time showed that dyads with children with phrase speech whose parents reported high PCS-P scores reported the least improvement in pain interference at follow-up. Due to the preliminary nature of the study, future work is needed to investigate the parental behaviors that mediate the relationships between parental catastrophizing and pain outcomes in this population.

Keywords: catastrophizing, cerebral palsy, pain measurement, pain behaviors, parent child dyads, communication, observational pain assessment tools

INTRODUCTION

For children with complex communication needs resulting from intellectual and developmental disabilities (IDD) such as cerebral palsy (CP), parents play an essential role with regard to pain assessment and treatment. Cerebral palsy (CP) is the most common motor disability in children and musculoskeletal pain is very common in children and adolescents (1–3), and adults with CP (4). Because many individuals with severe CP cannot self-report their pain, particularly when cognitive impairment is also present, parents or other primary caregivers must interpret their child's behavior when they suspect that their child is experiencing pain to determine the severity and source of that pain, when to seek treatment, and what type of treatment to seek. It is well established that individuals with IDD frequently express pain in idiosyncratic ways (5), such as freezing or even laughing, so people who are unfamiliar may misinterpret these signs. As a result, clinicians often rely on caregivers to serve as proxy reporters for their child's pain, making their judgements a critical component of medical care.

Although parents and other primary caregivers are the most appropriate proxies for judging pain in children with IDD in most cases, many factors can influence both pain expression on the part of the child and judgements about the presence or severity of pain on the part of the caregiver. The Social Communication Model of Pain posits that the experience and expression of pain is the result of an integration of various biological, social, and psychological factors at the intrapersonal and interpersonal levels, involving not just the person experiencing pain, but also others present in the environment (6, 7). The model suggests that pain is not a simple biological construct, but a multifaceted dynamic process that is shaped by an individual's history and social environment.

For children and adolescents with (or without) severe disability, pain expression takes place within family microcultures. Repeated interactions and patterns of communication about pain over time have reciprocal influences on how pain is expressed and managed within the family system. Specific to children and adolescents with complex health care and communication needs who cannot reliably self-report, when considering proxy reports of pain, presumably the caregiver's own biological, affective, cognitive, and social factors influence both how the caregiver makes judgements about the presence and severity of pain, and the caregiver's behavioral responses to that pain. Relevant factors may include the caregiver's personal history of pain experiences, their sensitivity, biases, knowledge about pain and disability, and their relationship and perceived duties toward the child in pain (6). Such individual factors interact with dyadic parent-child variables and family level variables, according to family systems theory (8). Various moderators or mediators may be present, such as the child's age or developmental stage (9).

Among parents of children with chronic pain without IDD, parental behaviors, including modeling of pain behaviors and responses to child pain behaviors, have been shown to affect their child's pain expression, ratings of pain severity, and mental health and functional outcomes (10–13). In these studies, however, the children experiencing pain had no developmental or physical disabilities that affected their ability to communicate about their pain. It is therefore unclear how the psychosocial factors in the family environment affect pain expression and pain assessment among families of children with IDD, for whom conventional forms of communication are more difficult or absent.

Negative cognitive and emotional states with regard to pain have emerged as critical psychosocial variables in understanding individual differences in pain perception and experience. Although many different measures of pain cognitions exist, the self-report and parent-report versions of the Pain Catastrophizing Scale [PCS; (14, 15)] are among the most widely used. Catastrophizing is defined as "an exaggerated negative mental set brought to bear during actual or anticipated painful experience" [(16), p. 220]. People who are high catastrophizes judge pain stimuli as more threatening and express exaggerated pain reactions (17). Vervoort et al. (18) reported that catastrophic thinking moderated the relationship between parental presence/absence and facial expression of pain during an experimental pain paradigm among typically-developing children, suggesting potentially complex relationships between cognitive factors, the social context, and pain expression among typically-developing children.

Although it is typical and even adaptive in many cases for parents to worry about a child's pain, when worry becomes extreme it can become maladaptive, particularly in the context of chronic pain. Most studies have found that child pain catastrophizing is a stronger predictor of pain outcomes than parental pain catastrophizing among typically-developing children and adolescents with chronic pain. There is good evidence, however, that parent catastrophizing has an indirect influence on child outcomes primarily through its impact on child catastrophizing (19). In a triadic study of pain catastrophizing, Kraljevic et al. (20) found a significant positive correlation between the pain catastrophizing of fathers, mothers, and adult children. A systematic review found that parent catastrophizing was significantly related to increased child disability, depression, and parenting stress, yet weakly associated with child-reported pain intensity (21).

When the individual experiencing pain cannot advocate for themselves, the potential of parent catastrophizing to influence the parent's proxy report of their child's pain raises questions about the validity of these reports as accurate reflections of the child's pain experience. Parental psychosocial factors may also indirectly shape how children and adolescents express pain through the parent's behavioral responses to child pain behaviors. Therefore, the present study is a secondary data analysis exploring relationships between scores on the Pain Catastrophizing Scale for Parents in relation to pain outcomes in children and adolescents with CP with varying degrees of communication abilities. This was done in the context of a study primarily designed to evaluate the impact of surgical implantation of intrathecal baclofen (ITB) pumps on pain outcomes among children and adolescents with severe spasticity due to CP [see (22, 23)]. ITB implantation has been shown to reduce spasticity and parent-reported pain among individuals with CP (24, 25), but no studies have examined

the influence of caregiver factors on pain outcomes in this population. We anticipated that the relationships between parental pain catastrophizing and the pain measures, including change over time, might differ by child communication ability. We hypothesized that parental pain catastrophizing would be positively correlated with parental pain interference ratings and behavioral expression of pain at a group level, and that both measures would show significant decreases following ITB implantation.

METHOD

Participants

Parent-child dyads in which the children had clinical diagnoses of CP and were scheduled for ITB implantation at a specialty pediatric hospital were eligible for participation in this prospective cohort study. The sample represents a clinical convenience sample formed through consecutive enrollment based on scheduled ITB pump implant surgery. A total of 63 dyads participated between October 2013 and March 2019. For the current analyses, dyads were excluded if: the Pain Catastrophizing Scale was missing because the caregiver did not speak English fluently enough to complete it ($n = 2$), or due to time constraints ($n = 2$), the adult who attended the surgery and completed the initial questionnaires was not a primary caregiver for the child ($n = 4$), the same caregiver did not complete the questionnaires at all visits ($n = 3$), or the dyad did not complete any follow-up assessments within 90–280 days (∼3 to 9 months) following the implantation ($n = 6$). In addition, preliminary analyses suggested that patterns for both the parent-reported and direct observational measures differed between male and female caregivers. Because the sample of fathers who participated in the study was too small to provide stable estimates ($n = 5$), these dyads were excluded from the analyses, so all participants represent mother-child dyads. Finally, data from visits that occurred within 3 months of a major surgery or procedure, or during which parents reported acute pain, such as due to acute illness or injury, in the previous week were excluded to ensure that pain scores reflected primarily the influence of chronic pain. This resulted in the exclusion of one additional participant with no study follow-up study visits without reported acute pain. A final sample of 40 dyads contributed two valid data points for the pain scores. Of these participants, 32 completed at least one in-person follow-up assessment in the 3 to 9-month post-operative period. Observational data from two participants were not usable due to technical difficulties and/or challenges obtaining clear views of the participant's face during the standardized exam at one or both time periods, and one participant had acute pain at both follow-up visits during the time window. The final sample for the direct observation analyses was therefore 29. Demographic and clinical factors by communication status are reported in **Table 1**.

Procedure

As part of a larger prospective intrathecal baclofen (ITB) outcomes project, parent-reported psychosocial assessments were completed before ITB implantation for spasticity management. Assessments regarding pain and comfort, including the standardized pain exam, were completed prior to surgery and again at ∼3, 6, and 9 months post-surgery. Parents completed the initial measures on an iPad with the assistance of a researcher. Parents completed the follow-up measures on their own *via* an online REDCap survey. The direct observational assessment was completed in clinic areas while the participants were present prior to surgery and at follow-up standard of care appointments. The 3 to 9-month window for follow-up was selected for the current analyses because previous studies have shown that ITB results in decreases in pain within this period (23, 26) the initial 3-month visit was for complete post-surgical recovery, and in general, the time window minimized the risk of intervening surgeries or other health events that could interfere with the results. If participants had two time points with valid data within the selected 3 to 9-month window, the date closest to the 180 days post-surgery was selected for analysis. The average time to follow-up after surgical implant was 172 days (range = 95–251) for parent questionnaires and 175 days (range = 99–251) for the direct observational measure.

Measures

Pain Interference

A modified version of the Pain Interference subscale of the Brief Pain Inventory (BPI) as described by Tyler et al. (27) was used to assess the degree to which parents perceived that ongoing pain interfered with daily living for their child. The scale includes 10 items, each rated on an 11-point scale (0 = pain did not interfere, 10 = pain completely interfered). The items include general activity, mood, mobility, work school or chores, relationships with other people, sleep, enjoyment of life, self-care, recreational activities, and social activities (27). A previous study examining the psychometric properties of this modified version of the BPI reported that scores were significantly correlated to other proxy-completed pain assessments, such as the Dalhousie Pain Inventory (28). In our sample, Cronbach's alpha of the BPI was 0.97 at Time 1 and 0.98 at Time 2.

Pain Catastrophizing Scale for Parents

The PCS-P (assessed at Time 1 only) is a 13-item scale characterizing thoughts and feelings that parents may experience when their child is in pain (14). Parents rate the frequency with which they experience thoughts and feelings on a 5-point scale (0 = "not at all", 4 = "extremely"). In our sample, Cronbach's alpha of the PCS-P was 0.93.

Spoken Language Ability

Participants with CP were grouped into three groups according to their spoken language ability based on parent-reported verbal ability (phrased as a yes/no question), and observation of the child's spoken language during the study visits. Participants who were reported to not use spoken language were categorized as "none", those who were reported to use spoken language were grouped into "some words" or "phrase speech" based on observations of the participants during study visits. Participants

TABLE 1 | Participant dyad demographics by analysis.

Demographic and clinical variables	Behavioral reactivity and proxy report ($N = 29$)			Proxy report only ($N = 40$)		
Categorical variables	***n***	***%***		***n***	***%***	
Child sex						
Male	16	55		24	60	
Female	13	45		16	40	
GMFCS level						
II or III	5	17		8	20	
IV	5	17		9	23	
V	19	66		23	58	
Intellectual disability						
None	2	7		3	8	
Mild/moderate	11	38		18	45	
Severe/Profound	16	55		19	48	
Spoken language						
No phrase speech	23	79		30	75	
Phrase speech	6	21		10	25	
Parent college degree	21	72		28	70	
Race = white, not Hispanic, Latinx	22	76		32	80	
Continuous variables	***Mean***	***SD***	***IQR***	***Mean***	***SD***	***IQR***
Child age (months)	132.83	50.29	99–155	131.73	53.65	96–153
PCS-P scores	37.45	13.11	27–49	38.13	12.00	31–47
Behavioral react. scores						
Before surgery	23.28	7.30	18–28	23.47	7.35	18–28
After surgery	19.93	7.71	16–25	18.71	7.74	16–25
Pain interference scores						
Before surgery	43.59	34.18	12–67	41.68	32.63	8.5–65
Follow-up	25.07	26.47	6–36	22.23	23.91	18–29

who did not speak during the visits, or who spoke in brief utterances (i.e., no more than two words) were categorized as using "some words", and those who uttered at least one three-word utterance during the visit were categorized as having "phrase speech".

Demographic Information

Additional demographic information, including degree of cognitive/intellectual impairment (i.e., no impairment, mild/moderate, or severe/profound), child's date of birth, parental sex and educational attainment, and race/ethnicity were collected *via* a parent survey.

Gross Motor Function Classification Scale

The GMFCS is designed to provide an objective classification of motor disability in children with CP, with an emphasis on sitting and walking. Function is divided into five levels, with children at Level I having the most independent motor function and Level V having the least (29). Because a large majority of the children in the current study were functioning at GMFCS levels IV or V, those in levels II and III were grouped together for analysis purposes.

Pain Examination Procedure

A standardized range-of-motion pain examination procedure (PEP) was completed at both time points. The PEP was designed to identify potential sources of pain or discomfort, such as spasticity or gastrointestinal pain. This exam has been used in previous studies of assessing pain- and discomfort-related nonverbal behavioral reactivity to experimental pain in IDD, including Rett syndrome, a neurodevelopmental disorder with associated motor impairment (30). The PEP involved the examiner slowly moving each joint of the arms and legs through its full range of motion in a standard sequence. The exam also included rotation of the head to each side, but due to the face being difficult to score during the head movements, reactivity was only scored for the arm and leg portions. The procedure was video recorded for later scoring by trained observational coders.

Direct Observational Scoring

Behavioral reactivity was scored for each limb of the PEP using a modified version of the Pain and Discomfort Scale (31–33), a behavioral coding system consisting of observationally-defined nonverbal signs of pain and discomfort derived from the Non-Communicating Children's Pain Checklist–Revised

(34). Behaviors were scored in four categories: upper face behaviors included brow furrows, eyebrow raises, eye squeezes, narrowing of eyes, lip puckering, and rapid blinks; lower face behaviors included parted lips, mouth opening, mouth stretches, smiles, or grimaces, tongue thrusts, teeth grinding, and biting lips; body codes included flinches, movements away from the examiner, and guarding of the limb being touched (e.g., blocking the examiner's attempt with another body part); vocalizations included any vocalizations if the participant was nonverbal (e.g., moaning, grunting, yelling) and only words related to the experience if participants had verbal language (e.g., "that hurts").

To start, coders identified behaviors that would meet the operational definitions for the codes described above but occurred repeatedly or constantly outside the context of the exam. These codes were only scored during the exam if they increased in intensity or frequency during the PEP, such as a slightly open mouth at baseline opening widely during an arm movement. This process was included to minimize the impact of movement disorders or other idiosyncratic non-pain movements on scores. Subsequently, each behavioral category was scored for each limb of the four limbs on a 0–3 scale, with 0 being no observed behaviors from that category and three being three or more seconds or three or more occurrences of a defined behavior in that category. Scores for each category were summed for a total score for each limb from 0 to 12, and a total test score of 0 to 48.

All coders were trained to a 90% or higher interobserver agreement criteria with the lead coding trainer. All videos were first coded independently by two coders, then disagreements between the two coders' score sheets were resolved *via* consensus to create a final score for each PEP. Pre-consensus IOA for this sample was 86.50% (SD = 6.13).

Statistical Analyses

Descriptive statistics, including means and standard deviations, were calculated for all variables. Paired *t*-tests were calculated to evaluate simple change from before surgery to follow-up for each outcome. For descriptive purposes, change scores for each outcome were calculated by subtracting the value for each participant at follow-up from their value prior to surgery. Bivariate correlations were calculated for all continuous predictors and outcomes, including change scores.

Restricted maximum likelihood linear mixed models were used to evaluate change relationships between demographic factors, communication abilities, and PCS-P scores on change in pain interference, and behavioral reactivity scores. All analyses were conducted in R (35) using the lme4 (36), lmerTest (37), ggeffects (38), and cAIC4 (39) packages. *P*-values for fixed effects were estimated *via* the Satterthwaite approximation (37) for descriptive purposes only. For each outcome, a full model was calculated and then backwards elimination was used to remove uninformative terms from the models by comparing AIC values at each step (40). The model with the smallest conditional AIC value was selected as the final model (39). All models included random intercepts for participant dyads. The initial full models included time (i.e., before surgery and at follow-up), child sex, child age, GMFCS level (i.e., Levels II/III and Level V each compared to Level IV), spoken language ability, PCS-P scores, and the interaction between communication score and PCS-P score as predictors. Because preliminary analyses suggested that the "no spoken language" and "some words" groups did not differ from each other in any of the models, spoken language ability was dichotomized into "phrase speech" or "no phrase speech" for the purposes of the analyses. Each predictor was also included as an interaction with time to evaluate its effect on the slope of change. All continuous predictors were mean-centered and scaled to range from −1 to 1 to maximize the likelihood of model convergence. Time to follow-up (in weeks) was not correlated with any of the predictors or outcomes and so was not included in any models. For interpretation of reduced model results, estimated marginal means were calculated at the 25^{th} and 75^{th} percentiles from the current sample for continuous predictors as high and low values.

Although pain interference scores showed significant skewing, residuals plots from the final model showed that the assumptions of heteroscedasticity and normality of the residuals were reasonably met when using Gaussian (identity) link models for both variables. Behavioral reactivity scores were approximately normally distributed.

RESULTS

No significant differences were found by spoken language ability for child age or sex, parental educational attainment, race, or PCS-P, or pain interference scores. Although behavioral reactivity scores did not differ between dyads with and without phrase speech prior to surgery ($t_{27} = 0.537$, $p = 0.596$), the two groups did show significant differences at follow-up ($t_{27} = -3.992$, $p < 0.001$). There was a strong association between verbal ability and reported degree of cognitive impairment (gamma $= -0.763$, $p < 0.001$), and between communication ability and GMFCS (gamma $= -0.701$, $p < 0.001$). Descriptive statistics by spoken language ability are reported in **Table 2**.

Correlations among the pain outcomes and child age are reported in **Table 3**. At the bivariate level, PCS-P scores were most strongly correlated with pain interference scores before surgery ($r = 0.422$, $p = 0.007$), and at follow-up ($r = 0.374$, $p = 0.017$). Behavioral reactivity scores were not associated with parental pain catastrophizing. None of the variables were associated with child age.

Full and reduced model results for pain interference scores are presented in **Table 4**, and estimated marginal means are presented in **Figure 1**. On average, pain interference score decreased from 41.68 (SD = 32.63) prior to surgery to 22.23 (SD = 23.91) at follow-up [$t_{(39)} = 4.00$, $p < 0.001$, Cohen's d = 30.73]. Child age and sex were not maintained in the final model as main effects or interactions with time. GMFCS level was maintained as a main effect only. On average, individuals functioning at Level V (i.e., requires support to sit) had higher reported pain interference scores across both time points. The three-way interaction between spoken language, PCS-P scores, and time was retained in the reduced model. Among dyads

in which the child did not use phrase speech, estimated pain interference scores estimated at the 25th percentile of PCS-P scores decreased were 21.89 [95% CI: (4.58, 39.20)] prior to surgery and 4.29 [95% CI: (−12.95, 21.72)] at follow-up; at the 75th percentile of PCS-P scores, estimated pain interference scores were 45.98 [95% CI: (28.68, 63.27)] and 13.48 (−3.87, 30.84). For dyads in which the child had phrase speech, at the 25th percentile of PCS-P scores, estimated pain interference scores were 25.54 [95% CI: (5.41, 45.67)] and 15.93 [95% CI: (−4.20, 36.06)]; at the 75th percentile of PCS-P scores, estimated pain interference scores were 35.71 [95% CI: (8.69, 62.74)] and 38.63 [95% CI: (11.60, 65.66)].

Full and reduced model results for total behavioral reactivity are presented in **Table 5**. For the total behavioral reactivity during the PEP, scores decreased from an average of 23.28 (SD = 7.30) prior to surgery to 18.93 (SD = 7.71) at follow-up [$t_{(28)} = 2.94$, $p = 0.014$, Cohen's d = 8.92]. Child sex, gross motor function, and PCS-P scores were not retained in the reduced model. Only child age and phrase speech were maintained as main effects and interactions with time. Estimated marginal means over time for dyads in which the child did and did not use phrase speech are presented in **Figure 2**. Overall, dyads in which the child did not use phrase speech showed a decrease in behavioral reactivity from 23.25 [95% CI: (20.26, 26.24)] to 17.44 [95% CI: (14.42, 20.46)], whereas dyads in which the child used phrase speech showed no substantial change in behavioral reactivity, with a mean of 20.78 [95% CI: (14.79, 26.76)] prior to surgery, and 24.66 [95% CI: (18.72, 30.59)] at follow-up.

TABLE 2 | Summary of demographic and clinical characteristics by spoken language status.

Demographic and clinical variables	No phrase speech (*N* = 30)	Phrase speech (*N* = 10)
Categorical variables	***n* (%)**	***n* (%)**
Child sex		
Male	18 (60)	6 (60)
Female	12 (40)	4 (40)
GMFCS level		
II or III	3 (10)	5 (50)
IV	6 (20)	3 (30)
V	21 (70)	2 (20)
Intellectual disability		
None	4 (13)	2 (20)
Mild/moderate	7 (23)	8 (80)
Severe/Profound	19 (63)	0 (0)
Parent college degree	23 (77)	6 (60)
Race = White, not Hispanic/Latinx	24 (80)	8 (80)
Continuous variables	***Mean (SD)***	***Mean (SD)***
Child age (months)	136.47 (57.34)	117.50 (39.51)
PCS-P scores	38.83 (11.44)	38.12 (12.00)
Behavioral react. scores		
Before surgery	23.65 (6.81)	21.83 (9.54)
After surgery	17.43 (7.98)	24.67 (1.75)
Pain interference scores		
Before surgery	45.80 (32.61)	29.30 (30.97)
Follow-up	21.77 (19.85)	23.60 (34.70)

DISCUSSION

The purpose of this exploratory study was to evaluate whether child spoken language ability and parental pain catastrophizing influenced parent-reported pain and child behavioral expression of pain during a standardized pain exam among children and adolescents with CP undergoing ITB implantation. Overall, parents' pain interference ratings showed substantial decreases following ITB pump implantation, but the changes in behavioral expression of pain were modest. The interaction between PCS-P scores and child language ability scores predicted change in pain interference scores over time, with dyads in which the child used phrase speech and the parents reported high PCS-P scores estimated to have the smallest decreases in pain interference scores. As anticipated, however, higher PCS-P scores were associated with higher pain interference scores across levels of communication ability. These results suggest that, although parental negative pain cognitions appear to be related to estimated pain interference prior to surgery, relatively independent of the child's ability to communicate verbally, the change in parents' perceptions of pain interference following surgery may depend, at least in part, on the child's verbal communication ability. Nevertheless, these effects were fairly modest, and confidence intervals for the estimated

TABLE 3 | Bivariate relationships among all predictors and outcomes.

Variable		1	2	3	4	5	6	7
1	Pain interference T1	-						
2	Pain interference T2	0.443 (0.004)	-					
3	Pain interference change (T1-T2)	0.717 (<0.001)	−0.307 (0.054)	-				
4	Behavioral reactivity T1	−0.368 (0.049)	0.181 (0.346)	−0.522 (0.004)	-			
5	Behavioral reactivity T2	−0.237 (0.216)	0.000 (0.999)	−0.243 (0.204)	0.295 (0.120)	-		
6	Behavior reactivity change (T1-T2)	−0.096 (0.619)	0.149 (0.442)	−0.217 (0.259)	0.563 (0.001)	−0.623 (<0.001)	-	
7	Parental pain catastrophizing	0.422 (0.007)	0.374 (0.017)	0.157 (0.335)	−0.070 (0.720)	−0.031 (0.871)	−0.030 (0.878)	-
8	Child age	−0.018 (0.912)	−0.055 (0.735)	0.024 (0.884)	−0.246 (0.199)	−0.086 (0.659)	−0.127 (0.512)	−0.075 (0.647)

TABLE 4 | Full and reduced linear mixed model results for pain interference scores.

	Full model					Reduced model				
Model terms	**Est.**	**SE**	**df**	***t***	***p***	**Est.**	**SE**	**df**	***t***	***p***
Main effects										
Time	−27.34	11.73	31.98	−2.33	0.026	−23.96	5.27	36.19	−4.55	<0.001
Female	−10.86	8.61	57.58	−1.26	0.212					
Age	−1.27	9.66	57.56	−0.13	0.896					
Phrase speech	−3.76	11.25	58.93	−0.33	0.739	−2.34	10.36	57.91	−0.23	0.822
PCS-P	99.39	28.60	57.88	3.48	0.001	97.90	27.19	60.65	3.60	0.001
GMFCS										
Level II or III	−2.67	14.30	59.86	−0.19	0.852	−6.11	11.50	38.16	−0.53	0.598
Level V	18.85	10.99	57.95	1.72	0.092	18.19	8.70	34.79	2.09	0.044
PCS-P*Phrase speech	15.26	18.28	36.18	0.84	0.409	−56.55	48.55	59.99	−1.17	0.249
Interactions with Time										
Sex	10.92	9.90	32.02	1.10	0.278					
Age	−0.12	11.11	31.85	−0.01	0.991					
Phrase speech	21.36	13.32	32.54	1.60	0.119	19.75	10.58	36.03	1.87	0.070
PCS-P	−62.69	32.71	32.11	−1.92	0.064	−60.92	30.38	36.34	−2.01	0.052
GMFCS										
Level II or III	−4.37	17.11	32.90	−0.26	0.800					
Level V	−0.70	12.66	31.99	−0.06	0.957					
PCS-P*Phrase speech	121.49	59.36	31.95	2.05	0.049	111.86	53.92	36.09	2.08	0.045
cAIC = 751.82						cAIC = 741.24				
df = 34.67						df = 30.05				
Conditional log-likelihood = −341.24						Conditional log-likelihood = −340.58				

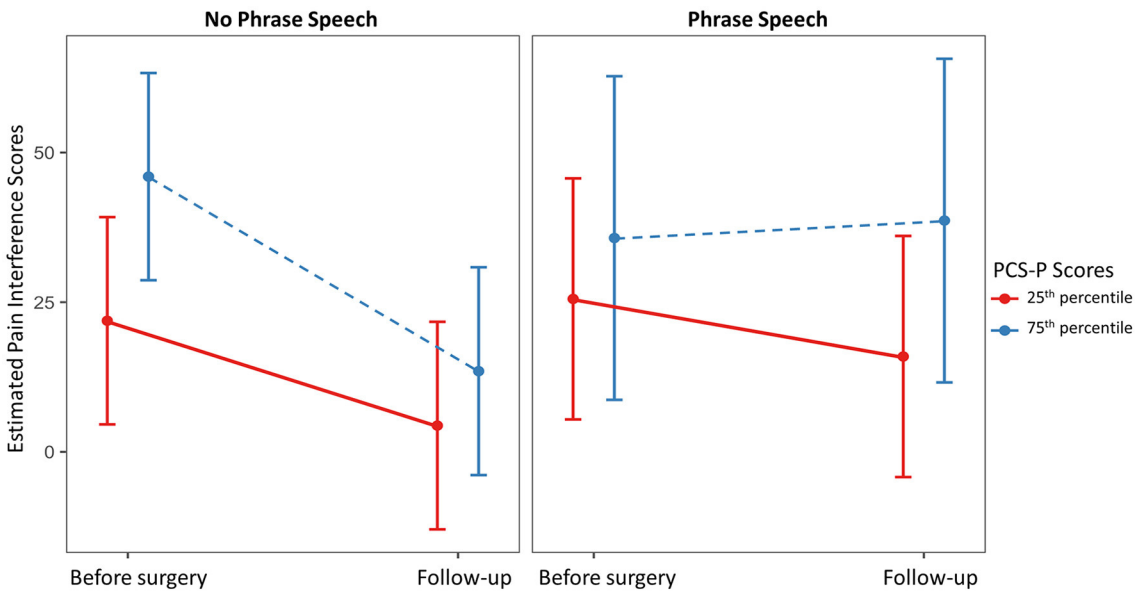


FIGURE 1 | Estimated marginal mean pain interference scores at the 25th and 75th percentile scores on the Pain Catastrophizing Scale for Parents (PCS-P) by child spoken language ability prior to surgery and at follow-up.

pain interference scores were wide, particularly in the group with phrase speech. Further, the findings do not provide information regarding the causality of the relationship between pain interference and PCS-P scores. In the context of chronic pain, this relationship is likely bidirectional, with higher child pain leading to greater parental catastrophizing, which in turn, leads parents to restrict activities or engage in other protective behaviors that ultimately increase child disability and chronic

TABLE 5 | Full and reduced linear mixed model results for behavioral reactivity scores.

	Full model					Reduced model				
Model terms	**Est.**	**SE**	**df**	***t***	***p***	**Est.**	**SE**	**df**	***t***	***p***
Main effects										
Time	−9.63	4.96	20.79	−1.94	0.066	−5.83	1.69	26.34	−3.44	0.002
Sex	−0.08	3.18	36.13	−0.02	0.981					
Age	−7.00	4.20	36.09	−1.67	0.104	−5.30	3.37	45.07	−1.57	0.123
Phrase speech	−3.78	4.10	36.49	−0.92	0.362	−2.48	3.43	45.04	−0.72	0.474
PCS-P	−12.17	9.09	36.53	−1.34	0.189					
GMFCS										
Level II or III	−0.51	5.44	37.98	−0.09	0.926					
Level V	−2.80	4.55	36.39	−0.62	0.541					
PCS-P*Phrase speech	15.26	18.28	36.18	0.84	0.409					
Interactions with time										
Sex	0.86	3.45	20.95	0.25	0.806					
Age	7.71	4.55	20.73	1.70	0.105	5.23	3.71	26.01	1.41	0.170
Phrase speech	11.27	4.52	20.91	2.49	0.021	9.69	3.78	26.00	2.56	0.017
PCS-P	12.31	9.81	20.99	1.26	0.223					
GMFCS										
Level II or III	2.32	6.14	21.28	0.38	0.709					
Level V	4.03	4.96	20.79	0.81	0.426					
PCS-P*Phrase speech	−14.78	19.84	20.76	−0.75	0.465					
cAIC = 406.50						cAIC = 391.53				
df = 30.53						df = 22.98				
Conditional log-likelihood = −172.72						Conditional log-likelihood = −172.78				

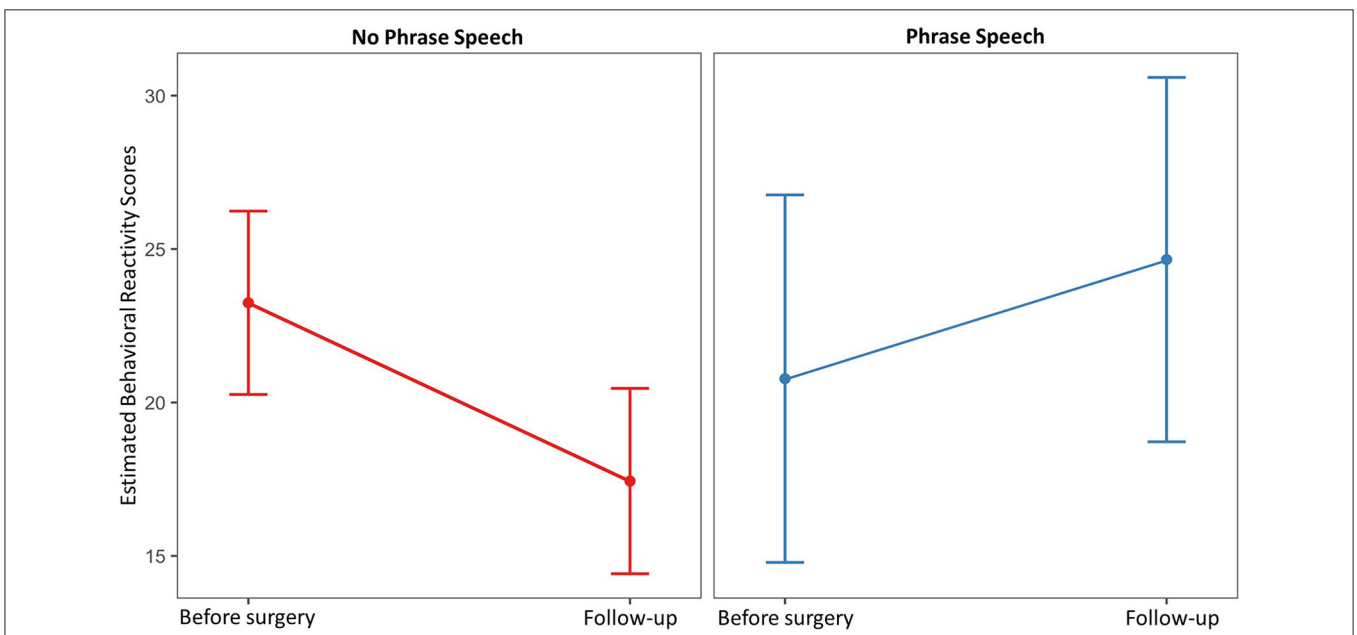


FIGURE 2 | Estimated marginal mean behavioral reactivity scores prior to surgery and at follow-up by child spoken language ability.

pain (9). Still, these results provide preliminary evidence that parent pain-related cognitions and resulting behaviors that are believed to mediate the relationships between cognitions and child pain outcomes may be most pronounced among dyads in which the child has more spoken language abilities. This relationship may be mediated by differences in the child's ability to describe their pain and advocate for pain relief, how parents talk with their children about their pain, and the degree to which

the children are able to imitate and learn from their parents' pain-related behaviors. Prospective longitudinal studies that evaluate parent psychosocial factors prior to the onset of child chronic pain are needed to fully understand these complex relationships.

PCS-P scores were not associated with behavioral reactivity in the current analysis. Although behavioral reactivity did not differ by spoken language ability prior to surgery, on average, individuals without phrase speech showed relatively large decreases in behavioral reactivity across time, whereas individuals with phrase speech showed no substantial change and a trend toward increasing reactivity. The reasons for this finding are unclear, although there are several potential explanations. One plausible explanation is that the finding is attributable, at least in part, to differences in cognitive ability between the two groups. Cognitive function was not formally measured in this study, and as such was not included in the statistical models. Assessing cognitive ability among children with severe communication and motor impairments is extremely challenging, but evidence suggests that communication abilities are often (but not always) correlated with cognitive ability in CP (41, 42). It is therefore plausible that individuals with more language abilities were also more aware of the context in which the study took place, and therefore more likely to exhibit anxiety in anticipation of the potentially painful exam, which may have inflated their reactivity scores at follow-up. Alternatively, it is possible that the behavioral observation scale used, which was designed to evaluate pain and discomfort among individuals with severe IDD, is not a valid measure for use among individuals without cognitive impairment. Behavioral expression of pain varies widely between individuals, and it is possible that individuals with more severe cognitive impairment may be less susceptible to social and environmental influences that may lead individuals to mask or suppress signs of pain. Scores for the two groups did not differ prior to surgery, however, suggesting that the measure captured some pain signs in this group in this context, and making this explanation less likely. Finally, because individuals with phrase speech were also more likely to have better gross motor function, it is possible that the observed differences are due to these factors as opposed to communication ability specifically. GMFCS level was included in the statistical models, however, and did not contribute to model fit for the behavioral reactivity models, suggesting that gross motor function does not account for the finding.

Although the current study was not set up to evaluate longitudinal relationships between the pain outcomes, it is notable that change in pain interference was not significantly correlated with change in behavioral reactivity. One possible explanation for these null results is that the way in which the standardized pain examination was implemented may have obscured improvements in musculoskeletal pain for some participants. Because the protocol indicated that the examiner should move each joint through the full range of motion to the degree possible, many participants likely exhibited a greater range of motion in their joints at follow-up due to the effects of the ITB pump. As a result, it is possible that changes in behavioral reactivity were not observed for some participants because there were minimal changes in the amount of pain elicited, despite likely improvements in pain during activities of daily living. Future work using similar measures should consider ways to control for such variability in the change of range of motion.

To our knowledge, this study was the first to examine the relationships among PCS-P scores, child spoken language ability, and pain outcomes among children and adolescents with IDD and associated chronic health conditions. As such, the study was exploratory in nature and all of the findings should be considered preliminary and specific to the sample. It is unclear whether these results would be replicated among samples of individuals with different etiologies of IDD, or in samples of individuals with CP without intellectual disability. Given the important role of parents as advocates for their children in healthcare contexts, future research is needed to replicate and extend these results.

The small sample size and missing data, particularly for the direct observational measure, are significant limitations of the current study. The dyads with complete data likely were not representative of the population of mother-child dyads undergoing ITB implantation. It is likely that the participants who returned to the hospital for follow-ups (as opposed to seeking follow-up care with other providers) differed from those who did not in terms of geographic location and other demographic factors. Because of the small number of father-child dyads who participated in the study, we were unable to include this group in the analyses. Preliminary analyses showed that pain catastrophizing scores, pain interference scores, and behavioral reactivity scores were all lower among father-child dyads compared to the rest of the sample, suggesting a potential influence of parent sex on parent psychosocial factors and child outcomes. Previous research has shown that mothers typically engage in higher levels of pain catastrophizing than fathers (43). As only one parent completed the measures, we were unable to consider triadic influences in two-parent families, or the role of sibling influence at the sibling subsystem or family level. Future research should consider specifically selecting for differences in family structures and expanding beyond the dyadic level to better understand these relationships.

Another limitation is that the pain measures used in the study did not specifically differentiate between acute and chronic pain. Although we attempted to isolate the effects of chronic musculoskeletal pain in the current analyses, relying on parent report of recent painful events is imperfect. Visits in which the parent reported a recent acute pain were excluded from the analyses, but this likely missed pain events that were not mentioned by or known to the parent. Nevertheless, we believe that the current results primarily reflect the impact of chronic pain in this population.

Several participating dyads did not complete the relevant questionnaire measures because translated versions of the measures were not available and the parents did not speak or read English well enough to complete them. As is the case for many pain studies, the resulting study sample was predominantly white and well-educated. Because pain experience and expression are likely influenced by family ethnocultures and microcultures (44), research in more diverse samples is needed to evaluate whether the impact of parent psychosocial factors on child outcomes varies across ethnic, racial, and socioeconomic groups.

In addition to the limitations already noted, several additional issues should be considered. Although the PCS-P has been used extensively in other populations, there are many other psychosocial variables that likely play a role in how parents rate and respond to their child's pain. For example, parent behavioral responses to pain, parent and child symptoms of anxiety and depression, and specific coping strategies have all been associated with outcomes for typically-developing children and adolescents with chronic pain [e.g., (12, 45–48)]. There is also some disagreement in the literature regarding the validity of the construct of pain catastrophizing as measured through self-report (49). Nevertheless, it remains among the most predictive psychosocial variables examined to date. Categorization of child spoken language ability was also limited; as this study was a secondary data analysis, however, no other standardized communication, cognition, or adaptive behavior measures were collected. Because the phrase speech variable was based primarily on the behavior of the individual with CP during study visits, it is possible that some individuals with the ability to use phrase speech were missed. The measure also did not take into consideration receptive language ability or nonverbal forms of communication, including formal augmentative and alternative communication systems and devices.

Many different measures exist for the assessment of pain and discomfort among individuals with CP and IDD, and selection of measures likely influences their relationships with parental variables. The pain interference scale was selected for the current analysis because it is a global measure of the parents' perceptions of the degree to which their child's quality of life is impacted by pain, as opposed to a more direct measure of pain intensity. We hypothesized that this would be more closely related to parents' levels of pain catastrophizing than other parent-reported pain measures, although there is no work investigating these relationships. The direct observation measure was selected because it was specifically designed to elicit signs of the musculoskeletal pain that was considered to be most relevant in the current context. Although both measures used in the current study have previously been documented as useful pain assessments among children with CP, there is no consensus regarding the most appropriate measures for this population, and development and validation is an ongoing process. Future research is needed to evaluate the relationships between parent psychosocial factors and the various pain assessment methods and to determine the most useful measures for specific purposes.

Despite these limitations, the results of the current study provide preliminary evidence that parent and child factors may influence proxy report measures of pain in children and adolescents with CP, although child factors appear to be more relevant for the direct observational measure. Palermo and Chambers suggest an integrative framework for the role of parent and family factors in a child's pain (2005). In this framework, pain expression in the family is a reciprocal process influenced by a child's developmental status that occurs within an ecological context (9). Parent catastrophizing or child verbal ability could be understood as individual variables that impact the dyadic and family levels. At the dyadic level, the parent's catastrophizing influences the child's catastrophizing and the child's verbal ability impacts the parent's perceptions of the child's pain interference (19). For example, the child's pain expression may be influenced by their parent's modeling of pain expression over time, including facial expressions, vocalizations, and gestures. The parent's perception of pain interference may be influenced by the child's verbal complaints or requests for analgesics, comfort, or rest.

At the family level, the family microculture around pain expression and management may be influenced by pain catastrophizing and/or child verbal ability. Consistent with the communal coping model of pain, Kraljevic et al. describe a family's microculture around pain as a "specific cognitive style for coping with pain, which is associated with a child's responses to pain experiences" (2011, p. 115). Over the long term, catastrophizing may adversely affect family atmosphere; this is likely bidirectional, as there is also evidence that family dysfunction predicts catastrophic thinking (50). When a child has significant disabilities associated with chronic pain, such as cerebral palsy, the risk for adverse effect on family atmosphere likely becomes more pronounced. Child verbal ability is also likely to have an influence on the family microculture as it relates to their pain; the family's communal coping style and pain management patterns may more or less depend on verbal cues.

There is an urgent need to understand the variables that may influence proxy report of pain for individuals with IDD because of the clinical implications of analgesic decision-making in this vulnerable population. This is the first investigation showing that parent-reported pain interference and behavioral reactivity in the context of a standardized pain exam vary according to both parent (pain catastrophizing) and child (communication ability) psychosocial factors. The construct of pain catastrophizing and how it relates to parents serving as proxy pain reporters for their child with a developmental disability needs further investigation, as does the construct of child communication ability, and this investigation should be considered in the context of parent-child dyadic and family level interactions over time.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Minnesota Medical IRB. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

BB was substantially responsible for data analysis and interpretation of the results, drafting, and revision of the manuscript. CR was substantially responsible for the initial draft of the manuscript, supporting interpretation of the results, and revision of the manuscript. CB was substantially

involved in conceptualizing the study, collecting the data, and revision of the manuscript. AM was substantially involved in data collection and analysis, and revision of the manuscript. KC was involved in the initial conceptualization of the study, and in revision of the manuscript. FS was substantially responsible for initial conceptualization of the study, data collection, and revision of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The work was supported, in part, by Eunice Kennedy Shriver NICHD Grant No. 73126.

REFERENCES

1. Mckinnon CT, Meehan EM, Harvey AR, Antolovich GC, Morgan PE. Prevalence and characteristics of pain in children and young adults with cerebral palsy: a systematic review. *Dev Med Child Neurol.* (2019) 61:305–14. doi: 10.1111/dmcn.14111
2. Schwartz L, Engel JM, Jensen MP. Pain in persons with cerebral palsy. *Arch Phys Med Rehabil.* (1999) 80:1243–6. doi: 10.1016/S0003-9993(99)90023-0
3. Turk MA, Geremski CA, Rosenbaum PF, Weber RJ. The health status of women with cerebral palsy. *Arch Phys Med Rehabil.* (1997) 78:S10–7. doi: 10.1016/S0003-9993(97)90216-1
4. van der Slot WMA, Benner JL, Brunton L, Engel JM, Gallien P, Hilberink SR, et al. Pain in adults with cerebral palsy: a systematic review and meta-analysis of individual participant data. *Ann Phys Rehabil Med*. (2021) 64:101359. doi: 10.1016/j.rehab.2019.12.011
5. Foley DC, McCutcheon H. Detecting pain in people with an intellectual disability. *Accid Emerg Nurs*. (2004) 12:196–200. doi: 10.1016/j.aaen.2004.06.002
6. Craig KD. The social communication model of pain. *Can Psychol.* (2009) 50:22. doi: 10.1037/a0014772
7. Craig KD. Social communication model of pain. *Pain.* (2015) 156:1198–9. doi: 10.1097/j.pain.0000000000000185
8. Bowen M. The use of family theory in clinical practice. *Compr Psychiatry*. (1966) 7:345–74. doi: 10.1016/S0010-440X(66)80065-2
9. Palermo TM, Chambers CT. Parent and family factors in pediatric chronic pain and disability: an integrative approach. *Pain.* (2005) 119:1–4. doi: 10.1016/j.pain.2005.10.027
10. Boerner KE, Chambers CT, McGrath PJ, LoLordo V, Uher R. The effect of parental modeling on child pain responses: the role of parent and child sex. *The Journal of Pain.* (2017) 18:702–15. doi: 10.1016/j.jpain.2017.01.007
11. Harrison LE, Timmers I, Heathcote LC, Fisher E, Tanna V, Duarte Silva Bans T, et al. Parent responses to their child's pain: systematic review and meta-analysis of measures. *Journal of Pediatric Psychology*. (2020) 45:281–98. doi: 10.1093/jpepsy/jsaa005
12. Logan DE, Simons LE, Carpino EA. Too sick for school? parent influences on school functioning among children with chronic pain. *Pain.* (2012) 153:437–43. doi: 10.1016/j.pain.2011.11.004
13. Lynch-Jordan AM, Kashikar-Zuck S, Szabova A, Goldschneider KR. The interplay of parent and adolescent catastrophizing and its impact on adolescents' pain, functioning, and pain behavior. *Clin J Pain*. (2013) 29:681–8. doi: 10.1097/AJP.0b013e3182757720
14. Goubert L, Eccleston C, Vervoort T, Jordan A, Crombez G. Parental catastrophizing about their child's pain. The parent version of the pain catastrophizing scale (PCS-P): a preliminary validation. *Pain.* (2006) 123:254–63. doi: 10.1016/j.pain.2006.02.035
15. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess.* (1995) 7:524. doi: 10.1037/1040-3590.7.4.524
16. Sullivan MJL, Adams H, Sullivan ME. Communicative dimensions of pain catastrophizing: social cueing effects on pain behaviour and coping. *Pain.* (2004) 107:220–6. doi: 10.1016/j.pain.2003.11.003
17. Kappesser J. The facial expression of pain in humans considered from a social perspective. *Philos Trans R Soc Lond B Biol Sci.* (2019) 374:20190284. doi: 10.1098/rstb.2019.0284
18. Vervoort T, Goubert L, Eccleston C, Verhoeven K, De Clercq A, Buysse A, et al. The effects of parental presence upon the facial expression of pain: the moderating role of child pain catastrophizing. *Pain*. (2008) 138:277–85. doi: 10.1016/j.pain.2007.12.013
19. Birnie KA, Chambers CT, Chorney J, Fernandez CV, McGrath PJ. Dyadic analysis of child and parent trait and state pain catastrophizing in the process of children's pain communication. *Pain.* (2016) 157:938–48. doi: 10.1097/j.pain.0000000000000461
20. Kraljevic S, Banozic A, Maric A, Cosic A, Sapunar D, Puljak L. Parents' pain catastrophizing is related to pain catastrophizing of their adult children. *Int J Behav Med.* (2012) 19:115–9. doi: 10.1007/s12529-011-9151-z
21. Donnelly TJ, Palermo TM, Newton-John TR. Parent cognitive, behavioural, and affective factors and their relation to child pain and functioning in pediatric chronic pain: A systematic review and meta-analysis. *Pain.* (2020) 161:1401–19. doi: 10.1097/j.pain.0000000000001833
22. Barney CC, Merbler AM, Simone DA, Walk D, Symons FJ. Investigating the feasibility of a modified quantitative sensory testing approach to profile sensory function and predict pain outcomes following intrathecal baclofen implant surgery in cerebral palsy. *Pain Med.* (2020) 21:109–117. doi: 10.1093/pm/pnz114
23. Barney CC, Merbler AM, Stansbury J, Krach LE, Partington M, Graupman P, et al. Musculoskeletal pain outcomes pre- and post intrathecal baclofen pump implant in children with cerebral palsy: a prospective cohort study. *Arch Rehabil Res Clin Transl.* (2020) 2:100049. doi: 10.1016/j.arrct.2020.100049
24. Ramstad K, Jahnsen R, Skjeldal OH, Diseth TH. Characteristics of recurrent musculoskeletal pain in children with cerebral palsy aged 8 to 18 years. *Dev Med Child Neurol.* (2011) 53:1013–8. doi: 10.1111/j.1469-8749.2011.04070.x
25. Vles GF, Soudant DL, Hoving MA, Vermeulen RJ, Bonouvri,é LA, van Oostenbrugge RJ, et al. Long-term follow-up on continuous intrathecal Baclofen therapy in non-ambulant children with intractable spastic Cerebral Palsy. *Eur J Paediatr Neurol.* (2013) 17:639–44. doi: 10.1016/j.ejpn.2013.06.003
26. Hoving MA, van Raak EPM, Spincemaille GH, Palmans LJ, Palmans LJ, Becher JG, et al. Efficacy of intrathecal baclofen therapy in children with intractable spastic cerebral palsy: a randomised controlled trial. *Eur J Paediatr Neurol.* (2009) 13:240–6. doi: 10.1016/j.ejpn.2008.04.013
27. Tyler EJ, Jensen MP, Engel JM, and Schwartz L. The reliability and validity of pain interference measures in persons with cerebral palsy. *Arch Phys Med Rehabil.* (2002) 83:236–9. doi: 10.1053/apmr.2002.27466
28. Barney CC, Stibb SM, Merbler AM, Summers RLS, Deshpande S, Krach LE, et al. Psychometric properties of the brief pain inventory modified for proxy report of pain interference in children with cerebral palsy with and without cognitive impairment. *Pain Rep.* (2018) 3:e666. doi: 10.1097/PR9.0000000000000666
29. Wood E, Rosenbaum P. The gross motor function classification system for cerebral palsy: A study of reliability and stability over time. *Dev Med Child Neurol.* (2000) 42:292–6. doi: 10.1017/S0012162200000529
30. Barney CC, Feyma T, Beisang A, and Symons FJ. Pain experience and expression in rett syndrome: subjective and objective measurement approaches. *Journal of Developmental and Physical Disabilities.* (2015) 27:417–29. doi: 10.1007/s10882-015-9427-3
31. Bodfish JW, Harper VN, Deacon JM, Deacon JR, Symons FJ. *Pain in Children and Adults With Developmental Disabilities* (2006). Baltimore, MD: Brookes Publishing.
32. Phan A, Edwards CL, Robinson EL. The assessment of pain and discomfort in individuals with mental retardation. *Res Dev Disabil.* (2005) 26:433–9. doi: 10.1016/j.ridd.2004.10.001
33. Shinde SK, Danov S, Chen C-C, Clary J, Harper V, Bodfish JW, et al. Convergent validity evidence for the pain and discomfort scale (PADS) for pain assessment among adults with intellectual disability. *Clin J Pain.* (2014) 30:536–43. doi: 10.1097/AJP.0000000000000020

34. Breau LM, McGrath PJ, Camfield CS, and Finley GA. Psychometric properties of the non-communicating children's pain checklist-revised. *Pain.* (2002) 99:349–57. doi: 10.1016/S0304-3959(02)00179-3
35. Core Team R. *R: A Language and Environment For Statistical Computing*. R Foundation for Statistical Computing, Vienna (2013).
36. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* (2015) 67:1–48. doi: 10.18637/jss.v067.i01
37. Kuznetsova A, Brockhoff PB, Christensen RHB. lmertest package: tests in linear mixed effects models. *J Stat Softw*. (2017) 82:1–26. doi: 10.18637/jss.v082.i13
38. Lüdecke D. ggeffects: Tidy data frames of marginal effects from regression models. *J Open Source Softw.* (2018) 3:772. doi: 10.21105/joss.00772
39. Säfken B, Rügamer D, Kneib T, Greven S. Conditional Model Selection in Mixed-Effects Models with cAIC4. *J Stat Softw.* (2021) 99:1–30. doi: 10.18637/jss.v099.i08
40. Cheng J, Edwards LJ, Maldonado-Molina MM, Komro KA, Muller KE. Real longitudinal data analysis for real people: building a good enough mixed model. *Stat Med.* (2010) 29:504–20. doi: 10.1002/sim.3775
41. Compagnone E, Maniglio J, Camposeo S, Vespino T, Losito L, De Rinaldis M, et al. Functional classifications for cerebral palsy: Correlations between the gross motor function classification system (GMFCS), the manual ability classification system (MACS) and the communication function classification system (CFCS). *Res Dev Disabil.* (2014) 35:2651–7. doi: 10.1016/j.ridd.2014.07.005
42. Himmelmann K, Lindh K, Hidecker MJC. Communication ability in cerebral palsy: a study from the CP register of western Sweden. *Eur J Paediatr Neurol.* (2013) 17:568–74. doi: 10.1016/j.ejpn.2013.04.005
43. Hechler T, Vervoort T, Hamann M, Tietze, A.-L., Vocks S, Goubert L, et al. Parental catastrophizing about their child's chronic pain: Are mothers and fathers different? *Eur J Pain.* (2011) 15:515.e1–e9. doi: 10.1016/j.ejpain.2010.09.015
44. Brady B, Veljanova I, Chipchase L. Are multidisciplinary interventions multicultural? a topical review of the pain literature as it relates to culturally diverse patient groups. *Pain.* (2016) 157:321–8. doi: 10.1097/j.pain.0000000000000412
45. Beeckman M, Simons LE, Hughes S, Loeys T, Goubert L. Investigating how parental instructions and protective responses mediate the relationship between parental psychological flexibility and pain-related behavior in adolescents with chronic pain: A daily diary study. *Front Psychol.* (2019) 10:2350. doi: 10.3389/fpsyg.2019.02350
46. Dahlquist LM, Pendley JS. When distraction fails: parental anxiety and children's responses to distraction during cancer procedures. *J Pediatr Psychol.* (2005) 30:623–8. doi: 10.1093/jpepsy/jsi048
47. Simons LE, Claar RL, Logan DL. Chronic pain in adolescence: parental responses, adolescent coping, and their impact on adolescent's pain behaviors. *J Pediatr Psychol.* (2008) 33:894–904. doi: 10.1093/jpepsy/jsn029
48. Simons LE, Smith A, Kaczynski K, Basch M. Living in fear of your child's pain: the parent fear of pain questionnaire. *Pain.* (2015) 156:694. doi: 10.1097/j.pain.0000000000000100
49. Crombez G, Paepe ALD, Veirman E, Eccleston C, Verleysen G, Ryckeghem DMLV. Let's talk about pain catastrophizing measures: an item content analysis. *PeerJ.* (2020) 2020:e8643. doi: 10.7717/peerj.8643
50. Akbari F, Dehghani M, Khatibi A, Vervoort T. Incorporating family function into chronic pain disability: the role of catastrophizing. *Pain Res Manag.* (2016) 2016:e6838596. doi: 10.1155/2016/6838596

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pain Research

METHODS
published: 23 March 2022
doi: 10.3389/fpain.2022.764128

Combining Electrodermal Activity With the Peak-Pain Time to Quantify Three Temporal Regions of Pain Experience

Viprali Bhatkar[1]*, *Rosalind Picard*[2] *and Camilla Staahl*[3]

[1] Digital Health Independent Consultant, Arlington, MA, United States, [2] MIT Media Lab, Cambridge, MA, United States, [3] Novo Nordisk A/S, R&D Business Development, Copenhagen, Denmark

Background: Self-reported pain levels, while easily measured, are often not reliable for quantifying pain. More objective methods are needed that supplement self-report without adding undue burden or cost to a study. Methods that integrate multiple measures, such as combining self-report with physiology in a structured and specific-to-pain protocol may improve measures.

Method: We propose and study a novel measure that combines the *timing of the peak pain* measured by an electronic visual-analog-scale (eVAS) with continuously-measured changes in electrodermal activity (EDA), a physiological measure quantifying sympathetic nervous system activity that is easily recorded with a skin-surface sensor. The new pain measure isolates and specifically quantifies three temporal regions of dynamic pain experience: I. Anticipation preceding the onset of a pain stimulus, II. Response rising to the level of peak pain, and III. Recovery from the peak pain level. We evaluate the measure across two pain models (cold pressor, capsaicin), and four types of treatments (none, A=pregabalin, B=oxycodone, C=placebo). Each of 24 patients made four visits within 8 weeks, for 96 visits total: A training visit (TV), followed by three visits double-blind presenting A, B, or C (randomized order). Within each visit, a participant experienced the cold pressor, followed by an hour of rest during which one of the four treatments was provided, followed by a repeat of the cold pressor, followed by capsaicin.

Results: The novel method successfully discriminates the pain reduction effects of the four treatments across both pain models, confirming maximal pain for no-treatment, mild pain reduction for placebo, and the most pain reduction with analgesics. The new measure maintains significant discrimination across the test conditions both within a single-day's visit (for relative pain relief within a visit) and across repeated visits spanning weeks, reducing different-day-physiology affects, and providing better discriminability than using self-reported eVAS.

Conclusion: The new method combines the subjectively-identified time of peak pain with capturing continuous physiological data to quantify the sympathetic nervous system response during a dynamic pain experience. The method accurately discriminates, for both pain models, the reduction of pain with clinically effective analgesics.

Keywords: sympathetic nervous system, electrodermal, EDA, SCL, VAS, cold pressor, capsaicin

OPEN ACCESS

Edited by:
Kenneth Craig,
University of British Columbia, Canada

Reviewed by:
Hugo F. Posada-Quintero,
University of Connecticut,
United States
Eleonora Vecchio,
University of Bari Aldo Moro, Italy

***Correspondence:**
Viprali Bhatkar
viprali.b@gmail.com

Specialty section:
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

Received: *25 August 2021*
Accepted: *15 February 2022*
Published: *23 March 2022*

Citation:
Bhatkar V, Picard R and Staahl C (2022) Combining Electrodermal Activity With the Peak-Pain Time to Quantify Three Temporal Regions of Pain Experience. Front. Pain Res. 3:764128. doi: 10.3389/fpain.2022.764128

INTRODUCTION

Pain involves a subjective experience influenced by factors such as fear, emotion, anxiety, cognitions, autonomic responses and malaise (1). Self-reported pain intensity does not correlate well with the severity of the pathological condition (2). Thus, quantification of analgesic effects in clinical trials, even with established analgesics, is frequently inconclusive (3).

Today's standard for pain measurement, the visual-analog scale (VAS) or electronic VAS (eVAS), allows participants to self-report their subjective experience of pain either statically–by reporting a single number, or dynamically–by turning a dial or moving a slider, usually along a scale from 1 to 100. While such scales have become the gold standard, being quick and easy to use, they have long been recognized to have problems with accuracy and reliability, with many factors beyond pain that influence the scores people give (1, 4). While many efforts are made to optimize self-report measures, e.g., customizing how it is presented for a particular population such as older adults (5), a holy grail of pain measurement is to obtain a more objective measure that sensitively reflects changes in pain experience and is easy to use. It also needs to work reliably and repeatably, across different participants experiencing different levels of pain in a lab study, and also on different days and visits, with repositioned equipment making valid measures, across different types of pain and analgesic use.

Recent surveys reviewed a growing number of automated methods to quantify pain objectively using facial expressions, vocalizations, physiology, brain-activity sensing, and more, and indicated the need for personalization of measures (6), as well as many wearable sensing approaches that can help quantify pain more objectively (7). While all of these measures show promise, each fully-objective method has its limitations, typically ignoring user-dependent subjective information, and focusing only on the objective data for one type of pain model and only during one day's visit or assessment period. The same emotions in the same person can exhibit patterns of physiology that change from day to day (8), so it is important to make sure that any pain-sensing method can account for this day-based variation.

Methods to elicit pain in a controlled manner have been refined via a large number of human pain models (9). In this work we use two well-established methods to induce pain: (1) the cold pressor, placing a limb into icy-cold water and holding it there, known for deep intense pain activating the descending pain system and its sensitivity to opioids (10), and (2) intradermal injection of capsaicin, which generates stable, long-lasting, and reproducible primary and secondary hyperalgesia lasting 2 to 3 h (11–13).

While many attempts have been made to develop an objective measure of pain, we focus in this work on a new measure that can be used easily and efficiently deployed in a variety of environments including the emergency room, post-operative recovery space, etc. This requirement rules out EEG, MRI, MEG, and fNIRS, despite that there has been exciting progress with these brain-based methods, e.g., (14, 15). We choose a measure that can be assessed as easily as vitals are assessed today with a readily applied wearable sensor, and measure the sympathetic nervous system response using a new characterization of electrodermal activity (EDA), which can be obtained quickly and easily by placing a sensor on the surface of the wrist or lower leg. The sensor can optionally be worn for continuous monitoring 24/7. Unlike the heart, which receives both sympathetic and parasympathetic nervous system innervation, the skin receives only sympathetic innervation (16). EDA thus provides a sensitive measure of sympathetic nervous system activity that can be captured effortlessly, and that changes continuously during a pain experience.

Sometimes EDA is considered non-specific, because it can be influenced by changing humidity and sweating, hydration, and strong emotions. Our method addresses the specificity problem by synchronizing the quantification of the EDA temporally to two precisely defined moments: (a) the moment of onset when applying a painful stimulus, and (b) the moment, identified subjectively, of "peak pain" experience. By using relative values in the regions anchored by these time points, EDA-based measures are likely to be highly specific to pain because they occur during an elicited experience of pain, acknowledged by a self-reported peak pain. Also, changes due to humidity, sweating, hydration, and emotions-unrelated-to-pain are minimal within the time-frame measured. We evaluate the proposed measure across people, across different days, across two pain models, and across three treatments, showing it addresses these traditional concerns.

While objective physiological data often have the strong and helpful property of being able to be continuously measured, sometimes they are limited because they change only at a *subset* of the moments of interest during a pain experience; for example, facial expressions might be most likely to occur at the onset of a cold pressor task, but the expression might fade or disappear completely, even as pain continues to increase, an observation identified decades ago (17). We seek in this work to develop a measure that continuously represents the trajectory of pain's anticipation, response, and decay.

While the subjectively reported levels provided on a VAS or eVAS can vary because of many factors unrelated to severity of pain, it is still routinely used. In our work, we use it in a way that extracts the timing of its peak, but then we discard the actual eVAS values. More specifically, when self-reporting pain, the exact value selected is highly subjective: it might be low simply because the participant wants to appear stoic. However, when a dial is turned continuously after a painful stimulus, it usually will increase up to a point, before it falls. Thus, each participant shows a moment of peak pain—the highest value relative to their other values. In our work, we find that the time to arrive at this peak is stable across pain sessions, even on different days with different pain treatments. The temporal position of the peak eVAS value is used to delineate two regions: The region rising up to this peak, and the region recovering from this peak. Our new measure then quantifies the EDA in these regions.

We also choose to include in our measure one more region: the assessment of the physiology during a period of anticipation immediately before the pain onset. This choice was inspired by hearing pediatric nurses discuss how some children flinch as if in pain or utter "ouch" *before* the needle touches them and by work showing that pre-pain anxiety can predict self-reported

pain (18). Quantifying this pre-pain anxiety is not typically done in pain research, but we think it is important for better understanding patient pain experiences and we recommend its measurement, at least as a contextualizing factor before the actual pain stimulus occurs.

To summarize, the proposed new method precisely characterizes and quantifies physiology over three temporal regions:

I (Anticipation): From the announcement of an imminent painful stimulus to the pain stimulus onset
II (Response) From the stimulus onset to the moment of subjectively reported peak pain
III (Recovery) From the peak pain moment to recovery from pain, or for a fixed time after the peak

These three quantities characterize our three-region pain measure.

It is well-established that pain should be highest during a painful-stimulus condition when no treatment is provided, reduced slightly under a placebo treatment, and reduced the most by effective analgesics. On placebo effectiveness, see for example Colloca and Barsky (19) and also demonstrations that higher-priced placebos work better than lower-priced ones (20). Using this knowledge, we test the novel three-region measure in a rigorous study with a 3-armed, placebo-controlled, randomized crossover trial design including 24 healthy adults. We systematically compare each measure before and after the effects of placebo, oxycodone and pregabalin. We also examine temporal situations known to affect pain measures, including the heightened anxiety expected during a "first visit," which can be expected to translate into a report of higher pain on the first time than when the identical procedure is repeated later. Finally, we show that the new measure outperforms the eVAS in all of these tests, demonstrating excellent pain discriminability.

METHODS

The methods used in this study are designed to evaluate a new measure of pain in the context of a clinical trial setting. We use treatments of previously established efficacy against pain (pregabalin, oxycodone) in a design of a randomized control trial. The trial applies a double-blind placebo-controlled multi-treatment, multi-day design. Outcomes were compared for all treatments both within and across participants, across days and weeks, and across different placements of the sensors, in order to comprehensively evaluate if the new pain measure is robust to all of these important variations. All study procedures were pre-approved by an ethics review board and the study was registered by ICON Development Solutions, under registration number EudraCT 2012-000484-25.

Participants

We recruited 24 healthy male adults, with normal body mass indexes (18–30 kg/m^2) and normal laboratory health tests. Each committed to attend four visits experiencing pain stimuli on four different days within a two-month period. Participants were non-smokers or light smokers (up to 5 cigarettes or equivalent per day). We focused this study on males since resources were limited and we wanted to reduce gender-based interactions and effects, as well as avoid menstrual-cycle changes and their impact on pain and physiology, which is a complex topic of ongoing research (21–23). Properly controlling the complexity associated with the female physiology would require a larger and longer study, even if it results in the same measure working for women as what we study here for men. Informed consent was obtained before commencing the study.

Pain-Elicitation: Cold Pressor and Capsaicin

The Cold Pressor Test

After preliminary equilibration of the hand temperature, and after alerting the participant that the process would start in 2 min, the participant was instructed to put one hand (the one without the palmar EDA sensor) into a cold-water bath (2°C) for 2 min whilst continually recording the pain intensity using the eVAS with the other hand. The right hand and left hand were used alternatingly on different visits. At the end of 2 min, the participant was instructed to remove his hand from the water bath. Pain was scored continuously using the eVAS starting at the time of the immersion and continued throughout the immersion.

The Intradermal Capsaicin Test

Participants were familiarized with pain evoked from 100 μg of capsaicin at the Training Visit. A single intradermal injection of capsaicin (100 μg) was made into the volar surface of the upper forearm (Manufacturer: ICON Development Solutions Manchester. Composition: 1 mg/ml capsaicin in 10% v/v ethanol, 7.5% v/v Tween 80 in 0.9% sodium chloride solution (100 μg/100 μL).) The injection of intradermal capsaicin was announced to the participant 5–8 min before the injection. The right arm and left arm were used alternatingly on different visits. Pain was scored continuously using the eVAS starting just prior to the intradermal injection and continued for 15 min after the injection.

Pain-Measurement: eVAS and EDA

Recording of eVAS

Pain intensity was assessed using an eVAS with the left end (=0) being equivalent to "no pain" and the right end (=100) referring to "worst pain imaginable". Participants were asked to evaluate pain intensity continuously by selecting the point on the eVAS that corresponds to the pain intensity they have at that moment in time. Participants were instructed to take both pain intensity and unpleasantness into account when scoring pain. While these can be considered two different dimensions, many studies show similar behavior of both dimensions during experimental pain, e.g., Duncan et al. (24), and efforts to distinguish them are ongoing (25) and not addressed in our study design.

EDA

EDA was measured electrically as skin conductance, using the Affectiva Q sensor, which measures skin conductance level (SCL) in microSiemens using 1 cm Ag-AgCl dry electrodes. Sampling rates were 8Hz. Each participant wore synchronized Q sensors

on five different locations, left wrist (LW), right wrist (RW), left ankle (LA), right ankle (RA) and right or left palm (P). The palm side was alternated over the four visits and was worn on only one side because the cold pressor test required submersion of one hand into ice water. In the rest of this paper, only the data from the four limbs was used as the palm data was too often noisy from movement artifacts.

Protocol

The Protocol is illustrated in **Figure 1**. Each patient made four visits: An initial training visit (TV), followed by three treatment visits (Treatments A, B, and C) in randomized order. Treatments were applied double-blind to treatment condition, and all data analyses in this paper were conducted initially without the condition being revealed. Later, they were revealed to be: A=pregabalin, B=oxycodone, and C=placebo. All four visits had a similar structure: First the patient put on the five EDA sensors. Next, a baseline heat-pain stimulation on the non-dominant hand was performed (but is not analyzed in this work, as the timing of each part of the series of rapid stimulations was not reliably recorded for comparison to the EDA). Next, they experienced the cold pressor, while filling out eVAS continuously during the immersion. Then the treatment was applied in the form of an oral capsule, except during the first visit, the TV, when no treatment was made. Then, the patient rested for an hour, which allowed treatment A, B, or C to take effect. Next, they experienced again the same cold pressor test, while continuously reporting eVAS levels. Then, 10 minutes elapsed while they filled out the State Trait Anxiety Inventory (STAI) (26) (not analyzed here). Next, the capsaicin was administered while they continuously reported eVAS levels. Finally, the sensors were removed and the participant was dismissed.

Overall, this design enabled systematic examination of multiple important comparisons including: (1) measuring the same stimulus (cold pressor) before and after treatment (no treatment, placebo, oxycodone, or pregabalin) within the same day; and (2) measuring reported pain across different days (cold pressor and capsaicin) x (no treatment, placebo, oxycodone, pregabalin). Since a person's physiological patterns can vary a lot from day to day, it is important to see if the proposed eVAS-peak-anchored electrodermal measure shows consistent differentiation both across days as well as within days.

In an effort to mitigate the effects of anticipatory arousal, which is likely to increase with increasing uncertainty, the sequence of events was first shown to all participants up front during their first visit (TV), and then this same sequence of events was used in that visit and all subsequent visits. Only the treatment (blinded administration of A, B, or C) was randomized across the subsequent visits. All visits followed the same procedure to reduce the influence that uncertainty has on autonomic stress responses. Also, participants are given an indication 2 min before each cold pressor task that it is going to start in 2 min and similarly 5–8 min before the capsaicin injection, so that anticipatory effects and time periods are held as constant as possible across the procedures and days. This helps eliminate "surprise" effects on autonomic responses.

Data Processing

ICON recruited 27 healthy adult male participants. Of these, three men dropped out of the study early and we received data sets for 24 participants x 4 sensors (LW, RW, LA, RA) x 4 visits = 368 sets of physiological responses. One participant's data had the wrong sampling rate for visit 1, no eVAS for visit 2, and no data for visit 3, so we dropped his data, leaving 23 sets.

Each file was visually inspected to confirm that the data record contained quality signals throughout the entire visit. Some files needed to be omitted due to bad data quality (malfunctioning sensor or sensor placed too loosely to record, causing visibly high levels of noise). Also, a total of 9 participants missed some visits or dropped out at some point after completing the training visit. Overall, 295 of the potential 368 files from the four limbs and 92 visits were obtained with high quality (80.2%). These 295 are distributed as: TV = 76 files, A = 73 files, B = 73 files, C = 73 files. All of these are used in the analyses that follow.

EDA Filtering, SCL Normalization, and Down-Sampling

Electrodermal activity can be divided into the "tonic component," the slowly varying part of the signal usually referred to as skin conductance level (SCL), and the "phasic component," the relatively fast changing peaks usually referred to as skin conductance responses (SCR's). The SCL is usually measured over intervals ranging from tens of seconds to hours, while SCR's are usually measured within 1–6 s after a discrete event.

Our analysis over the cold and capsaicin regions used SCL's derived as follows:

To separate the tonic from phasic EDA, a 5th order, zero-phase, lowpass Butterworth filter was applied to the raw skin conductance signal. The filter's cutoff frequency was set to 0.05 Hz as tonic activity is observed in 0–0.05 Hz. The SCL for each 1-min epoch was estimated using a 1-min wide centered moving average filter.

We compared data from multiple bodily locations and from multiple people over multiple days as baseline physiology can vary from day to day. We needed a robust way to make the data values comparable across all these files. Also, to accurately assess the changes in SCL after an analgesic, it is necessary to compare the SCL before and after the treatment on a common scale. We chose to use Z-score normalization before making all of these comparisons. To perform Z-score normalization, the (low-pass filtered) SCL for each file (one sensor, one day's session) was used to compute the mean and standard deviation for that session. Then the file's SCL was normalized by subtracting the mean value for the day's session and dividing it by the standard deviation for that session, such that the normalized SCL for the day has zero mean and unit standard deviation. Thus, the two cold and one capsaicin session for a person's visit were normalized using the same mean and standard deviation for that day.

The normalized SCL was subsequently analyzed over each of the regions I, II, and III for each cold and capsaicin segment.

Computation of Three-Region Measure

Pain is a mix of psychologically perceived phenomena and physically experienced phenomena (13). In this work, we

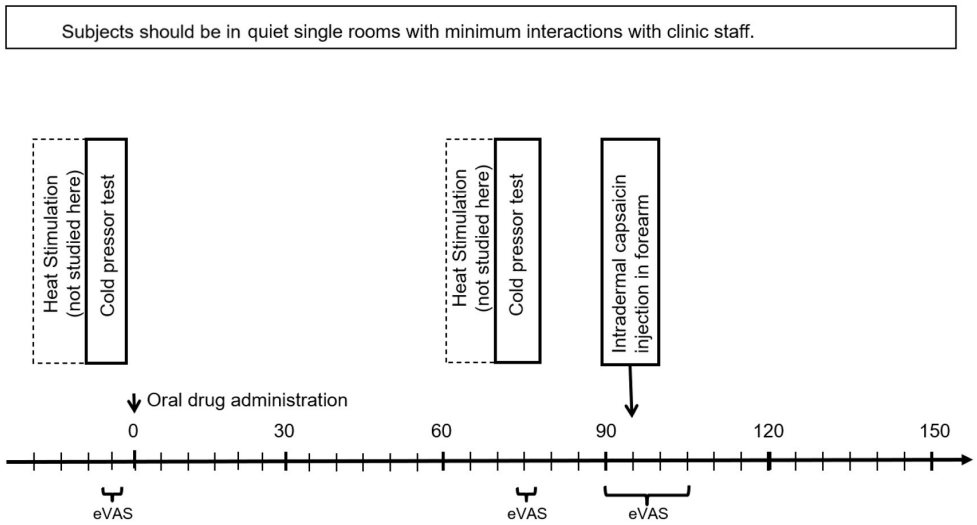


FIGURE 1 | Study protocol, showing three time periods where eVAS was measured. EDA was measured for the entire session (~3 h).

combine these two components in a novel way, using two time points—the objectively measured onset time of the pain stimulus and the subjectively measured time of the peak of the self-reported eVAS data—to structure the analysis of the physiological EDA data into three regions of the pain experience. The result gives a measure that improves on eVAS by adding objective data yet incorporates a valuable aspect of the self-reported pain experience—its peak pain moment as perceived by the participant in pain.

Here is how the three-region method works (See **Figure 2**): Using eVAS data and timing information of when the person was warned of the impending cold or capsaicin stimulus, we define three non-overlapping regions. These three regions separately quantify three regions of the pain response: I. Anticipation, II. Response, and III. Recovery.

Region I = Pain Anticipation

For the cold stimulus, participants were warned approximately 2 min before the cold pressor test. For the capsaicin stimulus, the warning period was from 5 to 8 min. We define Region I, "the anticipatory period", to be the region of time from the onset of the warning to the onset of the pain stimulus. We expect that SCL during this region is affected more by anticipatory anxiety than by physical pain. It is important to include responses during this region because sometimes people appear to actually experience pain before the stimulus touches them: For example, a child might jerk back and scream with "pain" before a needle touches them, and adults sometimes exhibit a facial grimace as if in pain before the onset of actual sensory pain. Thus, we include Region I, the subjective pain anticipatory experience, as part of the pain experience. The eVAS was not reported during region I so we cannot compare physiology with eVAS in that region. However, SCL is hypothesized to rise with anticipation, uncertainty, and anxiety, and our study data confirm that the SCL usually rises during Region 1, even sometimes taking on high values here.

Region II = Pain Response (Rising From Onset to Peak)

Region II is defined as the region of time that begins with the onset of the pain stimulus and ends when the person reports their peak pain level. In this study, the cold stimulus begins when the hand is placed in the ice water, and the capsaicin stimulus begins when the needle is inserted. The participant begins to report eVAS at this onset moment. Region II spans the time from the start of the pain stimulus and start of the eVAS recording to the peak reported eVAS level. The timing of this peak is clearly visible for capsaicin, which has eVAS that tends to follow the shape shown in **Figure 2** (green line = eVAS, blue line = SCL).

For cold pressor, we compute the peak location differently, as the eVAS often climbs monotonically and doesn't peak until the 2-min cold pressor test ends (See **Figure 3**). If we counted the peak as the right-most point, then we would often have just Region II and no Region III. We think it is valuable, even though the hand is still immersed, to examine this later portion of the cold pressor task where the eVAS tends to "level off" separately from the first portion of the immersion, where the eVAS typically climbs fast. Thus, for cold pressor pain we define the peak to occur at the time that the eVAS levels off–specifically, where it ceases to go up more than 0.005 units or 99.99% of the maximum value.

Region III = Pain Recovery (Sustain or Decay)

In this study, for both cold pressor and capsaicin, Region III is defined to begin at the peak identified in the eVAS. For cold pressor, Region III is measured until the cold pressor is ended (2 min from cold pressor onset), while for capsaicin, Region III is measured until 15 min following the onset of the capsaicin stimulus. For cold pain, this region is where the eVAS is usually leveling off–pain is "sustained." For capsaicin pain, this region tends to be where the eVAS values "decay" as the person is in recovery from the initial capsaicin injection.

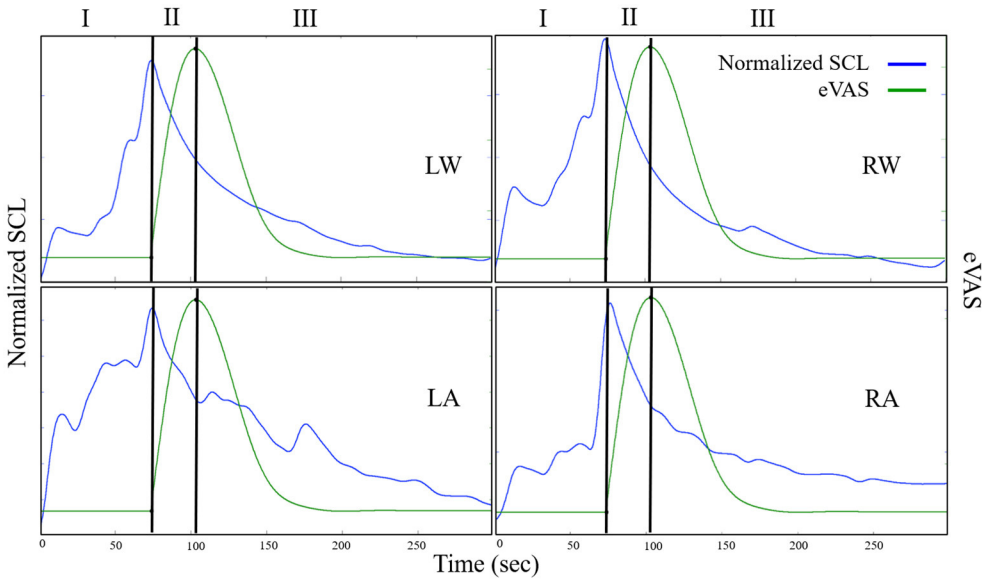


FIGURE 2 | Examples from one person illustrating SCL (blue) and eVAS (green) during the Capsaicin segment for the left wrist (LW), right wrist (RW), left ankle (LA) and right ankle (RA). Region I starts when the person is told that the capsaicin treatment is next, and ends when the needle is inserted. Region II starts with needle insertion and continues until the person's self-reported eVAS reaches a peak. Region III is measured from the peak eVAS until 15 min after the onset of the capsaicin treatment. Region III for Capsaicin usually contains decreasing values of SCL, as in the examples shown here.

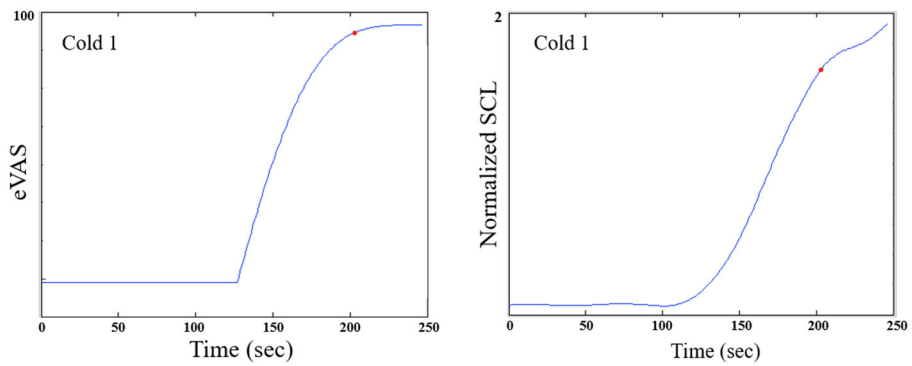


FIGURE 3 | For the cold pressor, the eVAS tends to grow monotonically throughout the task. Thus, a region defining Region III strictly as following the "peak" could be of zero duration. We thus define a function of the peak as the point at which the eVAS does not change > 0.005. The red dot shown above marks the onset of Region III for this participant. Typically, the SCL does not drop during Region III of the Cold pressor test; thus, we measure Region III over a fixed 2-min duration.

Using these pain regions, I, II, and III, defined by the times of announcement of the stimulus, the onset of the pain stimulus, and the time of the patient-reported eVAS peak, we next examine how an objective measure–the normalized, low-pass filtered SCL–changes both within each region and across the regions for each type of visit and each type of treatment. In particular, we wish to evaluate if the new method presented here, anchoring physiology relative to these three regions, is useful for more objectively measuring pain and for measuring the efficacy of active vs. placebo treatments.

Below we examine the performance across three pain-stimuli events: Cold1 (cold pressor applied before the treatment pill was consumed), Cold2 (cold pressor applied more than an hour after the pill was consumed), and Capsaicin (after the Cold2) during each of four visits made by each participant: Training visit (TV), Treatment A, Treatment B, and Treatment C. The oral treatment was administered on visits A, B, and C between the first and second half of the session in a double-blind way by a staff person who came in the room to give them the pill and otherwise did not observe the patients during the trials. Thus, the analgesic is given time to take effect before Cold2 and Capsaicin on visits A and B, while the placebo is given for visit C. The visit ordering was randomized across the patients, with the exception that no treatment is given during the first visit, TV.

RESULTS

Cold Pressor Pain

For each visit, a participant experiences two cold pressor tests: Cold1 (before the pill is given) and Cold2 (after the pill is given). On the training visit, there is no pill given but the patient rests for the hour between the two cold pressor tests. Our hypothesis is that participants will experience significantly less pain during Cold2 than during Cold1 when the treatment contains an analgesic (A or B) and the difference in pain between Cold2 and Cold1 will be small during C (placebo) and insignificant during TV. We also predict that the response to Cold1 will be highest on the first visit (TV) because of the extra anxiety and uncertainty associated with the novel pain experience.

In **Figure 4**, we see the results of the analyses applied to the full set of data. The bars indicate standard errors. The data from sensors on the LA, RA, LW, and RW were compared for $n = 23$ participants. The blue lines represent the mean normalized SCL values of regions I, II, and III during Cold1. As hypothesized, it is seen that, the Cold1 lines (blue) climb in value during all four visits, TV, A, B, and C, i.e., the means increased steadily through all three regions for all the Cold 1 episodes, as there is no treatment present during any day's Cold1 experience. The largest normalized SCL increase is observed during TV's Cold1. While this is not a finding central to our hypotheses, it was expected nonetheless: We have usually seen in studies with SCL that greater uncertainty is associated with higher SCL. Given this was the first visit, and the first experience of an unknown amount of (untreated) pain, it is expected that this visit would have had both the highest uncertainty and the highest increase in SCL. It is one of the reasons we designed the study to have a TV and also a placebo control, so that benefits of the new method are not overestimated by this "first visit" effect.

In these same figures, the red lines represent the mean normalized SCL during the second experience of Cold pressor pain each day, Cold2. We see for TV's untreated and C's placebo treatment that the SCL climbed over time, similar to all the visits in Cold1. For analgesic treatments A and B, however, the objective SCL shows that even if the overall level was a little higher at the moment of starting the pain segment (e.g., from hypothesized higher sweating when taking analgesic B), the relative increases that usually happen from the pain onset to peak, and beyond, were clearly attenuated by both analgesics (Note that the stderr bars are too small to be seen graphically in the plots for five of the measurements shown).

Figure 5 plots the two deltas (difference in mean normalized SCL) between adjacent regions of the pain experiences. As predicted, since there is never an analgesic at the time of Cold1, the deltas for Cold1 (blue bars) are significantly larger than the deltas for Cold2 (red bars). The statistical significances of the measures shown in **Figure 5** are computed and shown in **Table 1**. We performed a Wilcoxon Rank Sum test on the deltas between regions I and II, and regions II and III, during Cold1 against the respective deltas during Cold2 (This test was chosen because the deltas do not have a normal distribution). We compared the data from 23 participants and for all the limb sensor locations together for each treatment. We see all hypotheses confirmed: for TV there is no significant difference in the SCL changes during the first and second cold pressor tests, which is as expected given no treatment is given. As hypothesized for analgesic Treatments A and B, there was a significant difference before and after the treatment was given. This difference is consistent with the hypothesized significant reduction in pain expected following treatment A or B, both of which are well-known effective analgesics.

We hypothesized a small reduction in pain in the placebo condition, C, which we also found using the new measure. Interestingly, in condition C, we see a small but statistically significant difference involving the anticipatory period, but not during regions II & III once the pain has become established. Thus, the placebo shows a transient impact on the onset-to-peak pain but no impact on the sustained pain. This is an interesting aspect of the three-region method: It is deliberately separating out the anticipatory region, which may be the region most impacted by cognitive and affective beliefs, such as belief about the helpfulness of a placebo.

In **Table 1**, $h = 1$ is used to designate when the changes in SCL between regions I and II, and between regions II and III, were significantly lower. We see that h=1 occurs after the drug for analgesic treatments (in Cold2 for both A and B). For the placebo treatment C, the reduction from region I to II for Cold2 can be interpreted as a transient placebo effect that reduced the pain response during the anticipatory period preceding the stimulus onset. The reduction did not occur from region II to region III for the placebo.

We performed another comparison to further test the significance of the new three-region measure by comparing pain responses under a treatment (deltas between regions I and II and between regions II and III) for treatments A, B, and C, with pain responses during the training visit (no treatment during TV). Passing these tests is more challenging since they occur across different days. An objective physiology-based measure that shows reliable results across cold-pain experiences on different days is less likely than one that shows results only within the same visit's Cold1 vs. Cold2 comparisons because people tend to have different physiology from day to day due to hydration, mood, stress, and other natural variables. Further, mood or stress effects can also bias each person's self-reported eVAS range from day to day.

As seen in **Table 2**, looking across the multiple days of visits, the new measure's deltas for treatment A and treatment B were still significantly different than were the deltas during the TV and treatment C (placebo) for Cold2. Comparisons were made with a 1-tailed Wilcoxon test. These results show that the median changes in SCL from region I to II and from region II to III during Cold2 (after the drug) for treatment A and treatment B were significantly smaller, as hypothesized, than the median changes in SCL during Cold2 for TV. Thus, the measure shows that the analgesics result in a significantly reduced pain experience, unlike the placebo and no-treatment conditions, and this effect captured by the new measure is robust across different visits.

These tests illustrate several strengths of the proposed new measure: A valid objective pain measure should show that treatments A and B reduce pain compared to placebo treatment C, and compared to no treatment. The results in **Table 2**

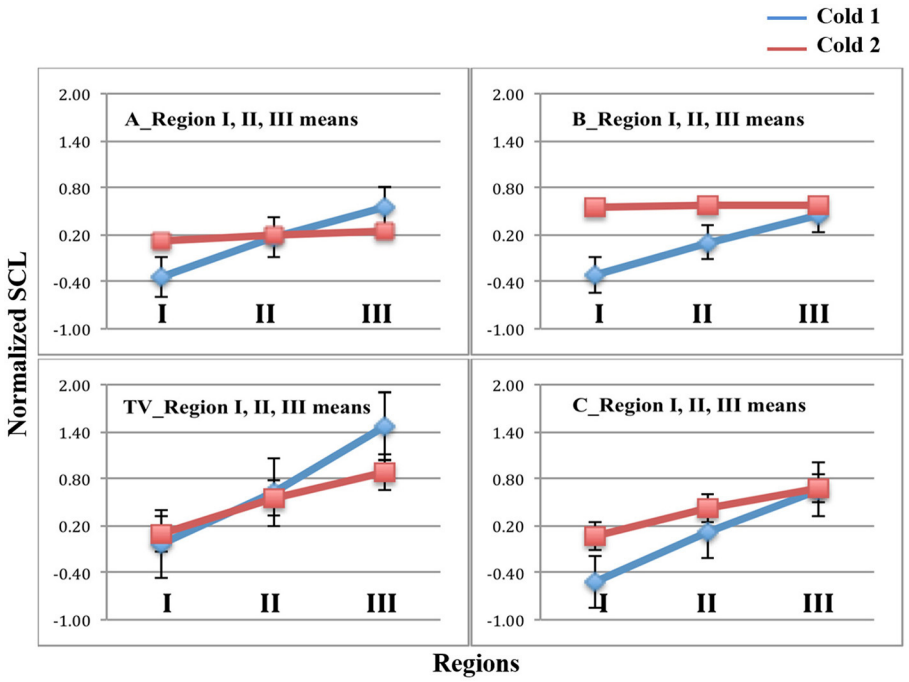


FIGURE 4 | Mean and standard error bars of normalized SCL in Regions I, II, and III. Blue = Cold1. Red = Cold2. Data are from the four limb sensors for n = 23 participants. Analgesics were applied in A and B, and are associated with a lack of increase across the Cold2 pressor.

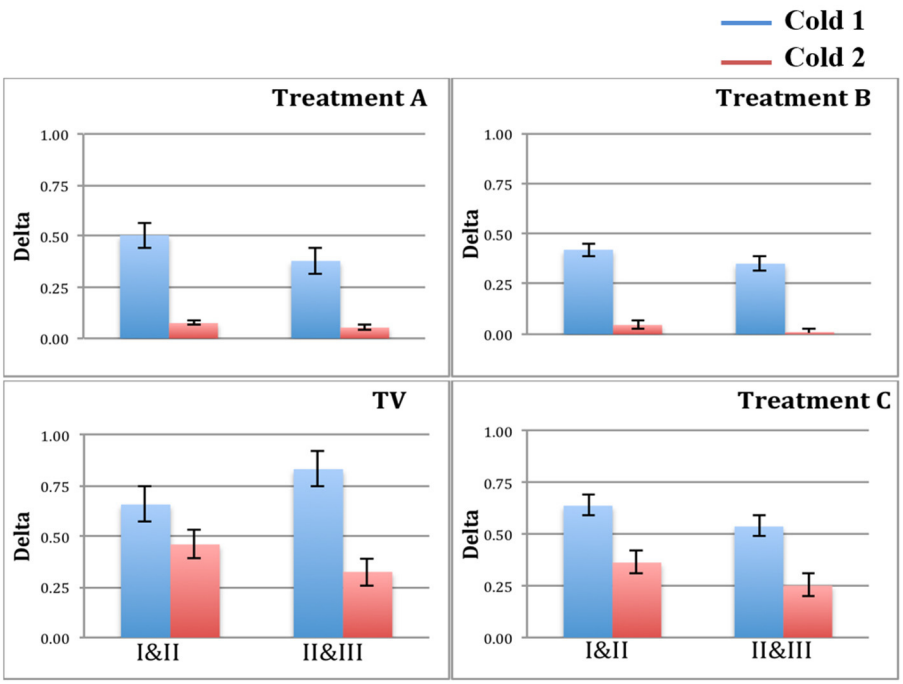


FIGURE 5 | Blue bars = delta values between regions for Cold1. Red bars = delta values between regions for Cold2. Treatments A and B show a significant reduction in the deltas, as hypothesized for these two analgesics.

confirm that the new measure shows these statistically significant reductions for Cold2. We see the significant effect of comparing analgesic conditions, A and B, to non-analgesic condition TV, and the non-significant effect of C's placebo compared to TV.

TABLE 1 | Testing for statistically significant changes in normalized SCL between adjacent regions during Cold1 (before treatment) and Cold2 (after treatment).

Δ	TV I and II	TV II and III	A I and II	A II and III	B I and II	B II and III	C I and II	C II and III
h	0	0	1	1	1	1	1	0
p	0.355	0.151	0.000	0.000	0.000	0.000	0.023	0.065
N	75	75	76	76	73	73	73	73

One-tailed Wilcoxon rank sum test (testing if delta Cold2 < delta Cold1 and assigning h = 1 if this is true). The significant effect of the analgesic is confirmed for treatments A and B, while a significant effect is also seen only at the start of the placebo C.

TABLE 2 | Testing for statistically significant changes in normalized SCL for Cold1 (before treatment) and Cold2 (after treatment) across sessions on different days.

Δ	TV vs. A Cold1_I and II	TV vs. A Cold1_II and III	TV vs. A Cold2_I and II	TV vs. A Cold2_II and III
h	0	0	1	1
p	0.485	0.436	0.000	0.001
N	75 vs. 76	75 vs. 76	75 vs. 76	75 vs. 76
Δ	**TV vs. B Cold1_I and II**	**TV vs. B Cold1_II and III**	**TV vs. B Cold2_I and II**	**TV vs. B Cold2_II and III**
h	0	0	1	1
p	0.213	0.142	0.000	0.000
N	75 vs. 73	75 vs. 73	75 vs. 73	75 vs. 73
Δ	**TV vs. C Cold1_I and II**	**TV vs. C Cold1_II and III**	**TV vs. C Cold2_I and II**	**TV vs. C Cold2_II and III**
h	0	0	0	0
p	0.545	0.788	0.651	0.659
N	75 vs. 73	75 vs. 73	75 vs. 73	75 vs. 73

Comparison is made between the first visit, TV (no treatment for Cold1 or Cold2), and later visits A, B, C. As desired in a measure, all of the no-treatment conditions do not differ over time. The effect of the analgesic is significant only in the diminished responses of SCL to Cold2 in conditions A and B (and not in Cold1 conditions or in placebo C).

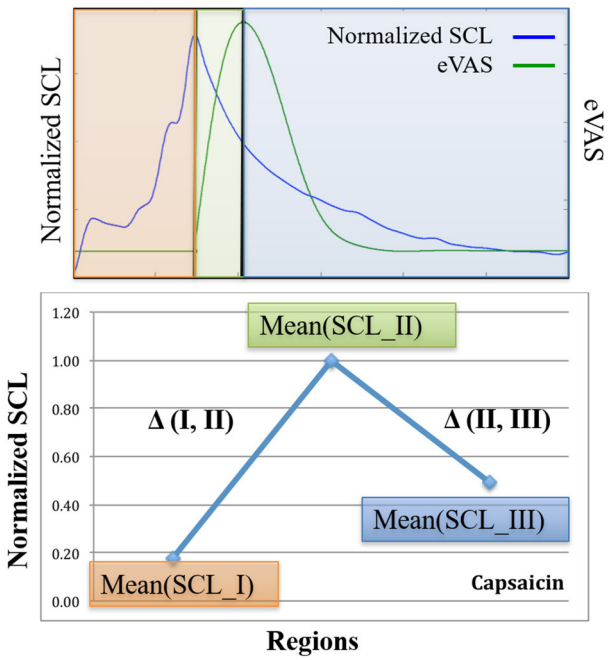


FIGURE 6 | Notation for Capsaicin regions and deltas.

Moreover, the new measure's significant differences cannot be attributed simply to day differences, as the study further confirms the presence of no such difference across the days during Cold1, before the analgesics are applied (where all h = 0).

Capsaicin Pain

For each visit, each participant experiences one inoculation of Capsaicin to elicit pain. We make similar tests for the Capsaicin pain model. The big difference in these tests is that now we have only one inoculation per visit in the second half of each visit, so we cannot compare pre- and post-drug within the same day's visit. Instead, we must evaluate the harder challenge of comparing across visits that have analgesics (A, B) and that don't have analgesics (TV, C), even though these occur on different days. Thus, to find a reliable, repeatable result for Capsaicin pain is a greater test of the new measure's robustness than when the measures are made within the same day's session.

We first characterize the changing pattern of mean SCL across Capsaicin regions I, II, and III, as this is a different kind of pain than cold pressor pain. The notation we use for Capsaicin is described in **Figure 6**, where we again denote the three regions relative to the time of onset of the needle (pain stimulus) and to the (subjective) eVAS-reported peak pain. Again, we show the eVAS in green, and we use its peak to separate regions II and III. The SCL is shown in blue, having an earlier peak at the end of the anticipatory region, the moment when the needle is applied. The first time we saw this "anticipatory" peak preceding the actual reported peak-pain we were surprised (This occurred in a prior pilot study with flu-shot data, where it occurred the moment before the needle was inserted). We find in this clinical trial data that such a peak sometimes occurs as in the example shown here, and sometimes occurs closer to the self-reported peak pain. This phenomenon is another reason to explicitly measure Region I.

In **Figure 7**, we show the mean normalized SCL for all four types of visits, during each of the three regions of the Capsaicin experience. First, we see a general arc across all the visits TV (blue), A (red), B (green) and C (purple): The anticipatory period

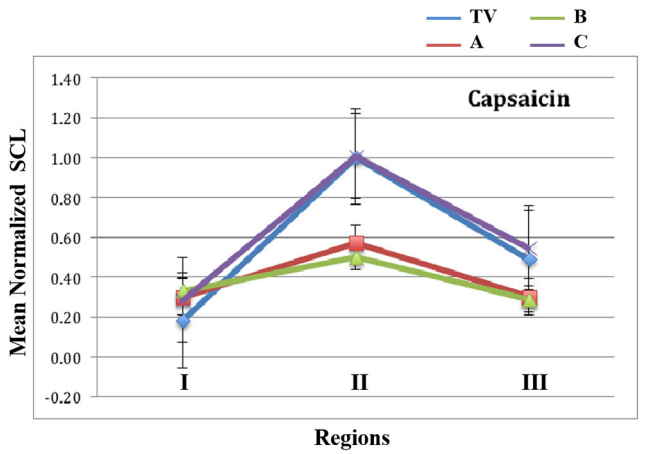


FIGURE 7 | Mean values of normalized SCL within regions I, II, and III for capsaicin. TV = blue, C = Purple, and the analgesics A = red, B = green. Data combine four limb locations from n=23 participants; bars are stderr.

is relatively low in all visits. In the middle, we see that the peak pain eVAS period is also in the region usually having the peak SCL. This property of this measure is seen to be robust for all four types of visits. Finally, the slow recovery of eVAS is similar to that of the SCL during Region III.

The capsaicin pain experience is divided into three regions: the anticipation of pain just before the injection, the needle pain with the injection and its feeling of pain increasing to a peak value, followed by the feeling of the burning wearing off slowly as the peak pain subsides.

As hypothesized, the proposed new measure shows that the arc of the three-region response is less severe for the two analgesic conditions A and B than for the non-analgesic TV and C conditions.

We examine the statistical significance of the measure by comparing the mean values of EDA in regions I and II and in regions II and III of treatments A, B, and C with those of the training visit TV. We apply the Wilcoxon rank sum test to examine the mean values within each region, across the conditions (**Table 3**).

Results show that the mean values of normalized SCL for region II are significantly different between A and TV, and between B and TV, and *not* between C and TV. These results are all in the hypothesized direction: The analgesics reduce the pain response more than placebo, which reduces it more than no treatment.

Importantly, the changes in our new measure are not due to an "overall reduction in SCL" from the analgesic because we confirm (**Table 3**) that the SCL is not different in Region I, before the onset of the pain stimulus, even though all treatments had been given more than 90 min before this time.

Note that these statistically significant effects, for both the capsaicin and the cold pressor pain models, were found before the team doing the data analysis was unblinded to conditions A, B, and C.

TABLE 3 | Capsaicin pain: comparisons of normalized mean SCL in treatments A, B, and C vs. the training visit, TV, within each of regions I, II, III.

	TV vs. A		
	I	**II**	**III**
h	0	1	1
p	0.409	0.000	0.041
N	75 vs. 76	75 vs. 76	75 vs. 76
	TV vs. B		
	I	II	III
h	0	1	1
p	0.109	0.000	0.021
N	75 vs. 73	75 vs. 73	75 vs. 73
	TV vs. C		
	I	II	III
h	0	0	0
p	0.110	0.323	0.478
N	75 vs. 73	75 vs. 73	75 vs. 73

We see the hypothesized reduction of pain confirmed in regions II and III of the analgesic treatments A and B (h=1 for both). We also see the "no difference" hypothesis confirmed for the placebo treatment C, and for all anticipatory periods (before the needle is inserted).

DISCUSSION

This paper presents a novel measure of characterizing pain response based on objectively identifying the time of onset of a pain stimulus, subjectively identifying the peak-pain moment (from the numerical peak of a self-reported eVAS), and then quantifying physiological changes in the three regions delineated by these two time points. The resulting quantitative measures are shown to provide statistically significant discrimination validating the effectiveness of well-known analgesics compared to placebo and no-treatment.

Does the new method work better than self-reported eVAS data alone, and if so, when might it replace it? Before showing this quantitative comparison, it is worth noting some of the features of traditional psychophysical methods of pain assessment, which request a report of subjective pain experience using either one-dimensional pain scales (like eVAS) or multidimensional pain scales; for a more complete picture of self-reported evoked pain response, various assessments must be used (6, 27–30). Different aspects of pain response such as psychological distress or anticipation, pain intensity, and pain recovery interact in complex ways to determine the perception and experience of pain (31–33). The proposed new three-region model summarizes these complex interactions quantitatively with three physiological values that capture meaningful differences in pain level across treatments both within a day and across days; however, the topic of how the measures of the three regions map to the many subjective aspects of pain, and their assessments by multidimensional pain scales, is not currently captured by the method in this paper. These topics remain a challenge for future studies.

The new method adds some complexity to eVAS: eVAS is only one modality, while combining it with EDA integrates a second. Thus, we directly test: Is the new combined EDA + eVAS method performing objectively better than using only eVAS across this data set? The answer is yes, as seen in **Figure 8** where we show the mean eVAS values across conditions TV, A, B, and C for Cold1 vs. Cold2, and **Tables 4**, **5** where statistical significance comparisons are made. We also tested the max eVAS values and the area under the curve of the eVAS, and the results were similar, with the only case of statistically significant discrimination occurring with eVAS and treatment B in the case of Cold pain, and with no significant discrimination with the Capsaicin model using eVAS alone. Another eVAS measure we tested was the time from the start of the stimulus to the max eVAS, which was found to not differ significantly across the visits for a given pain model. While this means that it fails as an eVAS measure at discriminating treatments that reduce pain, it does add strength to its use in our proposed new measure for defining Region II's endpoint, as it is stable across visits and across treatments. Thus, the value a person gives with the eVAS, used alone, fails to discriminate any pain reduction of using pregabalin for the more than 70 visits where eVAS measures compared Cold1 to Cold2, the latter after the treatment was given, and also fails to discriminate any pain reduction of either treatment with the Capsaicin model.

A limit of using only eVAS is also seen in the marginally significant difference found between TV and analgesic B (oxycodone) across visits when all measures are based on using only eVAS (**Table 5**). However, the three-region EDA + eVAS measure clearly distinguished both visits A and B from visits C and TV. Thus, the novel method outperforms traditional eVAS in a randomized control trial evaluating the cold-pressor model of deep pain. The new model is specific to pain (using eVAS to anchor the peak moment of pain) while being more discriminative than eVAS, even with a relatively small number of participants.

Note that it is possible that with a much larger number of participants, the difference between TV and A may eventually become significant when using only eVAS, as might at the same time the difference between TV and C. However, adding more patients adds substantial trial costs, and it requires inflicting pain on a lot more people. If the difference (Cold2 vs. Cold1) using the analgesic with a larger number of participants becomes significant, yet no greater than placebo's significance, then the drug will not be deemed effective. In contrast, the proposed new pain measure is significant in its discriminatory ability when using a small number of participants; thus, it may reduce both clinical trial costs and the ethical costs of inflicting pain on larger numbers of people.

EDA is traditionally recognized as responding to pain, but not specifically to only pain: It usually increases when the sympathetic nervous system is activated, with the fight or flight response, as well as with uncertainty and anticipation (16). Thus, an increase in EDA is usually expected with both anticipation of and experience of painful experiences. Using direct brain stimulation, researchers have shown that EDA is activated ipsilaterally by stimulation of the amygdala, anterior and posterior hippocampus, and anterior cingulate (34), key

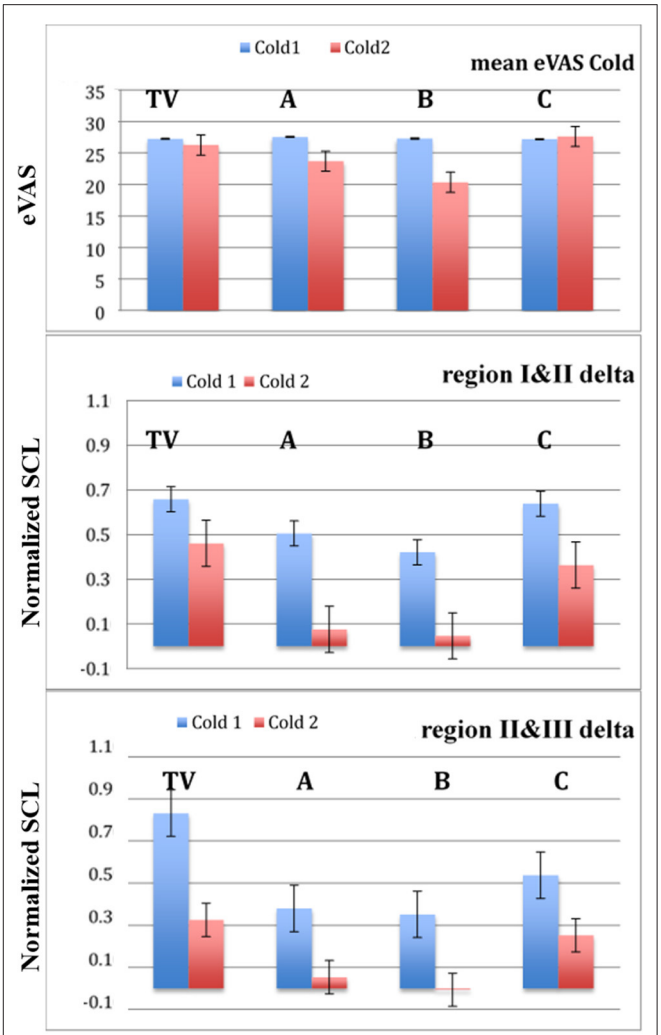


FIGURE 8 | The mean values of eVAS alone do not successfully discriminate both effective analgesics. The deltas comparing normalized SCL values across adjacent regions are more discriminatory. Cold 1 = blue, Cold2 = red, n = 23 participants. Bars are stderr.

regions involved in processing pain, emotion, and anxiety. Thus, the EDA measure in general will be sensitive to pain, changing when pain happens; however, it is not specific to only pain; for example, a significant increase in EDA may occur with brain activity during and soon after a grand mal seizure (35); also, it has been observed to be elevated at the time of death in the minutes following a grand mal seizure (36).

Our work here addresses the problem of specificity in several ways. First, like with early work showing that skin conductance responses reflect infant responses to painful heel sticks (37), we measure the level of pain objectively in a situation known to cause experience of pain, as would be expected in a clinical study, hospital, or recovery room, where contextual factors that might influence the pain measure are both observable and controllable. Second, and novel to our work, we specifically anchor the regions to-be-quantified by using the time point where the person

TABLE 4 | Using eVAS, there is no difference in pain response for analgesic A and we see only a marginal difference for analgesic B.

Test	TV-A-Cold1	TV-A-Cold2	TV-B-Cold1	TV-B-Cold2	TV-C-Cold1	TV-C-Cold2
h	0	0	0	1	0	0
p	0.95	0.56	1.00	0.05	0.95	0.56
N	75 vs. 76	75 vs. 76	75 vs. 73	75 vs. 73	75 vs. 73	75 vs. 73

TABLE 5 | Using only mean eVAS, the reported pain comparison for Cold 1 vs. Cold 2 differs significantly only for analgesic B.

Δ	TV	A	B	C
h	0	0	1	0
p	0.792	0.392	0.031	0.897
N	23	23	23	23

explicitly marks their (subjective) peak pain experience. Third, the way that we process the EDA data within the three regions removes effects likely to be influenced by the environment or other day-to-day varying influences: This was shown in our study design requiring visits on different days, likely to span different conditions of hydration, heat, and humidity. Finally, the way we designed the study with a training visit (TV) helped reduce influences due to study-specific effects on emotions that can be caused, for example, by the first visit's arousal where a patient experiences uncertainty and possible fear or anxiety related to the experimental conditions such as "What are they going to do to me next?" or "How badly will it hurt?" The resulting method thus works specifically for pain, as demonstrated not only with one pain model, but with two very different pain models.

We presented a novel method for improving upon a traditionally subjective method of pain measurement by defining three regions of the pain experience, anchoring these specifically to a patient's personalized 'peak' pain moment, quantifying objective autonomic data for each of the regions, and testing the discriminability of the method over 92 patient visits, including four conditions—two analgesics, a placebo, and a "no treatment" condition—within a randomized control trial. The method uses one piece of information from eVAS—the timing of its peak self-reported pain—but otherwise does not use any of the actual values from the subjectively-reported scale.

One of the interesting findings in this study was how the three-region measure gives insight into physiological changes occurring with the placebo condition, C. In all of the comparisons using the three-region measure, the placebo response was found to lie between that of the no-treatment training visit (TV) and the analgesic treatments (A and B), confirming well-known expectations about placebo effects. This finding is based quantitatively on the objective data from the physiology. Interestingly, when we separately examined the three regions of the pain response, the placebo condition was seen to have its largest affect during the anticipatory period, with smaller effects during regions II and III, once the pain became established. This finding suggests that for those who continue to use only eVAS to measure pain, they may find a different significance level simply by asking patients to report their pain at a different time. We hypothesize that the placebo effect has a different temporal trajectory than the analgesic effect. We suggest that future work examine its dynamics, which could have significant bearing on clinical comparisons, allowing the statistical significance of clinical findings that rely upon eVAS to be manipulated by adapting the timing of when pain is assessed. Methodologically, this timing is an important piece of information, and we suggest it should be reported in future pain study designs to add extra integrity to the design. Note that when today's methods use a "one value" rating of pain for the entire experience (corresponding to our regions II and III), then it will obscure this information (e.g., **Figure 8**'s average eVAS ratings).

Our method was shown to appropriately address the concern that an opioid (oxycodone, treatment B) causes higher sweating than another analgesic (pregabalin, treatment A) and increased sweating might interfere with a method based on SCL (38). We thought it was especially important to test this effect since we measure SCL on the wrists and lower legs, and these sites are sometimes (without evidence that we have seen) claimed as being more thermoregulatory than emotional, even though there is plenty of evidence, across many types of studies, that non-thermoregulatory events, such as those due to changes in neurological (e.g., seizures and sleep stages) and cognitive-affective states result in changes in SCL at these limb locations (16, 35, 36, 39–42). We thus examined the mean values of SCL across the four limbs during Cold 2 to see if there is more overall response with treatment B. As we saw previously in **Figure 4**, Treatment B's mean normalized SCL has a higher value for Cold2 in region I compared to A, C, and TV. This region is immediately after the treatment is given orally (and allowed to take effect during rest) so we do see increased sweating in the baseline at the start of Cold2. However, our measure considers the change in SCL from region I to II, and from II to III, both of which remain low.

In **Figure 9**, we continue this examination by plotting the mean normalized SCL values from the four limb sensors during Cold2 region III, across all participants. This region corresponds to the highest SCL in the non-analgesic visits (TV followed by C = placebo) and is reduced significantly for analgesic A. While analgesic B shows values that are reduced from those in the TV and C conditions, we do indeed see higher SCL on average in B than in A, which is consistent with the reports of increased patient sweating with this drug. However, the mitigating effect of the analgesic on the three-region physiological measure is still statistically significant despite the increased sweating. In

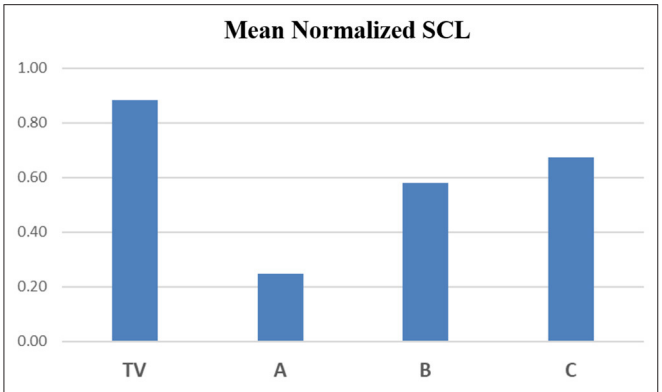


FIGURE 9 | Mean values of normalized SCL from the four limb sensors during Cold2's Region III.

short, the proposed three-region measure of pain appears to change in a way consistent with reduction in pain, not simply with the amount of sweating. It is robust at discriminating pain relief even given the effects of increased sweating from an opioid (oxycodone).

Returning to **Figure 4**, not only is the mean SCL higher for Cold2 treatment B in region I, but also it appears to be relatively higher than Cold1 in all of the non-first visits (A, B, and C) in region I. While the reason it is elevated for the opioid condition B is described above, it is interesting to consider what might affect this region I disparity in the other cases. While we cannot be sure of the cause without running a future causal experiment design, we consider some likely contributing factors here: (i) The region I for Cold1 is lower on repeated visits, which we expected because the Cold1 experience is unchanged from the first visit, and thus uncertainty about it is reduced. (SCL tends to increase with increasing uncertainty). (ii) While the repeated visits repeat the entire protocol, there is a novelty at the start of Cold2 during the visits A, B, and C. In these visits, participants have just consumed an unknown drug, A, B, or C, and they don't know if it is going to help or not. They also know that they are starting another round of painful experiences, which during the TV, was the most painful part of the visit. Thus, it is possible that now they are starting to have a little more anxiety, based on that earlier experience. Anticipation and increased uncertainty can raise SCL. Thus, these possibilities could raise baseline SCL for Cold2 during A, B, and C, making it higher than baseline for Cold1 on that visit. We note that the method still works across these variable conditions.

Anchoring the measurements to the three regions defined in this work addresses the specificity problem and using the deltas between these regions provides significant discrimination even with a small number of people ($n = 23$). Most other measures, such as the absolute ratings given subjectively on an eVAS, require much larger groups of people to achieve significant discriminability. Thus, the new method appears to provide a scientifically significant advantage AND to provide practical cost-saving improvements over the commonly used visual-analog scale. While both are specific to a pain event, the combination used in our method can achieve the level of discrimination desired between study conditions, using a trial with a smaller group of participants.

While the new measure has shown significant performance across many challenging comparisons, this work has limitations. The participants were all healthy adult males, who had to be capable of self-reporting their peak pain moment; thus, this method would not work with infants or others who could not communicate their pain, although adaptations of it may still be useful in such cases (43). It is also unclear how it would work in women; a larger study that examines their physiological changes across months exhibiting the variety of hormonal changes is needed. This work did not have the resources to address the larger study required to control for hormonal influences. The three-region method also requires knowing the timing for the onset of a pain stimulus, which is not likely to be present in all situations, although it can be controlled in many clinical trial studies. While we use a capsaicin injection to model chronic pain, it is also important to examine long-term real-world chronic pain, where we might expect to see significant baseline shifts as well as asymmetries in EDA, and measure also effects on daily behaviors, including pain's impact on sleep and activity patterns and how these relate to the EDA measures (41, 44, 45).

In the future, there are many possible ways to extend and possibly improve upon this work. We decomposed the tonic and phasic EDA data in a very basic, traditional way, and today there is a growing literature describing more advanced ways to analyze components of EDA data, e.g. see the systematic review by Posada-Quintero and Chon (46) especially contributions examining narrower frequency bands that may improve upon our use of a simple low-pass filter to give even better results (47). These additional ways to process the data would be interesting to explore. Many lab-based studies using heat-based pain induction have found that processing the EDA to extract the more rapid SCR's provides more accurate estimation of the pain level than does using the SCL. Particularly intriguing are findings from Posada-Quintero et al. showing that using a narrower frequency band inspired by analyzing sympathetic nervous system activity (similar to in heart-rate variability studies) leads to improved results for estimating the pain stimulation intensity (48, 49). Their findings also showed that EDA was better at estimating the stimulus intensity than the subjects' self-report scores. Our work is not directly comparable because we are not trying to use EDA to estimate self-report level, nor are we trying to estimate the numeric intensity level of a pain stimulus; instead, we are trying to examine, following a typical clinical trials protocol for evaluating a new treatment, if it is showing a significant difference in pain reduction compared to placebo or to no treatment. Our work doesn't directly use self-reported "level of pain" other than as an intermediate step to locate the self-reported moment of peak pain, which is then used to bound a region for measuring the skin conductance.

Despite the different goals of this work, an important future direction is to closely examine the contributing features of an EDA signal, including content from different frequency bands, and how they relate to (1) distress or discomfort, which would be expected to be higher for a first visit (expected in the TV condition and untreated or anticipated-as-untreated conditions), and to (2) characteristics of different pain models and their

stimuli. Many studies include repeated stimuli such as electrical or heat pulses, each of which can elicit an anticipatory orienting response in the EDA (these were unable to be examined in our study, because their stimuli were not timed properly for allowing synchronization to EDA). Stress and orienting responses, and small movements they elicit, may confound the pain response. At the same time, several researchers have shown that the high-speed fluctuating changes in EDA have been some of the most valuable features for classifying pain-related distress, (50), levels of self-reported pain, e.g. (51), and objective levels of applied pain intensity (48, 52). The latter have used particularly novel and well-performing frequency-specific methods of extracting SCR's, which would be important to examine in future work. We did not compute SCR's in this work, despite that we initially expected that they would be more informative than SCL, especially with our many-visit study design, since a mean SCL can vary highly within a person across days. Upon inspecting our data, even before smoothing, we observed very few SCR's, and many cases with zero SCR's, even during strongly-reported pain. See, as exemplary, the relatively smooth examples in **Figures 2**, **3**, where there is a dominant change in SCL after the onset of the pain stimulus, but with very few fluctuations around the large rise. In short, counting SCR's in the traditional way would not have given significant pain discrimination in this clinical trial. At the same time, our method does not use SCL in the traditional way, where typically the SCL after the pain stimulus is compared across treatments. That approach does not show a significant difference for analgesics vs. non-treatment or placebo in our study (in part because the opioid increased the SCL). Instead, our method anchors SCL's region of computation specifically to the pain event, normalizes it across a day's session, and computes changes in levels across three regions in a way that apparently reduces its influences from other factors that may have caused "average SCL" to not perform discriminatively in past studies. The new method's results outperform eVAS for discriminating the effects of analgesic vs. placebo under gold-standard blinded test conditions in a professionally-conducted clinical trial.

One might ask "Why compute three regions when only one measure of pain is typically sought?" Indeed, if a medical professional wants to know quickly whether or not a patient is hurting badly, asking for a subjective report is faster and if the patient can provide it, it can suffice for triage. However, a better characterization is needed in clinical trials to examine if one treatment reduces pain more than another. For clinical trials, the proposed new measure provides a better result than using eVAS. The new measure provides objective physiology data, anchored specifically in an onset-of-pain event and a subjectively-timed "peak experience of pain" event, establishing quantitative changes in the anticipatory, peak pain, and recovery regions of the pain experience. All three regions may be targets for future improved treatments. Our study shows that different treatments may affect these regions differently, and that quantifying these three regions in the way described provides greater discrimination of treatment effects than using self-reported pain.

Future work might examine, for different pain models, which way to use the three regions to give the best discrimination. As seen for capsaicin (**Figure 7**), the biggest differences between the no-treatment or placebo conditions and the two analgesics occur in Region II. If we were to simplify the three-region model to a two-region model (combining Region II and III, and comparing their combined value to that of Region I) then the size of the effect will be reduced, even if in some cases the difference is still significant. For the cold pressor model, it is not Region II but it is Region III that shows the biggest difference; this can be visualized in **Figure 4**, by shifting the red plots down to match their normalized SCL to that of Cold1 in Region I. The differences are due to the different pain models: capsaicin pain peaks immediately, while cold pressor pain takes minutes before it climbs. By attending to where the regions are most likely to differ for a given pain model, it becomes possible to examine more precisely where the benefits of a treatment occur.

Overall, our work contributes to the important goal of improving the measurement of pain, not trying to make it completely objective or deny its subjective reality, but making it *more* objective, and making its quantification more specific to three regions of a dynamic pain experience. The method is low-cost, practical, and easily combined (if desired) with studies that use more costly measures such as fNIRS, fMRI, and new kinds of brain imaging. The new method is well-suited for studies that evaluate different treatments for pain, such as clinical trials. Not only does the new method provide better discriminability than eVAS with fewer participants, it reduces the psychological and ethical costs of inflicting pain on a larger than necessary group of people. While the results we have shown suggest that the new method is more sensitive than traditional eVAS for clinical trials, our work has not focused on what is the underlying line of action. The evidence of a more sensitive measurement, compared to eVAS, showing the effects of the two analgesic drugs used, especially pregabalin, may be related to its possible action directly on the sympathetic nervous system rather than specifically targeting the perception of pain. More work is needed to understand how effective the proposed measure continues to be when tested with additional kinds of pain models and treatments. The three-region method also may potentially improve the methodology for studies designed to elicit and measure responses to pain by giving better insights into placebo interactions and the impact of cognitive and affective contexts that can influence the experience of pain, whether these occur before the onset of the actual pain stimulus, after its onset, or during the recovery period after the peak pain. Overall, this study, within the format of a clinical trial, has shown that the proposed method works better than eVAS across multi-day visits by healthy men, across two pain models, and across conditions of no-treatment, placebo, and two well-known effective analgesics.

DATA AVAILABILITY STATEMENT

The study participants were in Europe and while they consented to have their anonymized data published as part of this research, they did not consent to have their raw data shared for other research purposes; thus, under European law, their datasets cannot be made available.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by ICON plc IRB, Manchester, UK. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

VB contributed the idea and first implementation of anchoring the pain measurement timing to the peak self-reported pain, conducted data analyses, and helped write the paper. RP contributed to the design of the study, the methods of data analysis, and wrote most of the paper. CS contributed to the design of the study, writing the trial protocol, obtaining approval by ethics committees and medicines agencies in Manchester, helping write the paper, and was part of training the CRO performing the trial and conducting clinical oversight with the trial. All authors contributed to the article and approved the submitted version.

FUNDING

This work received funding from Grunenthal GmbH. The study was co-designed by author RP (working at MIT and consulting for Affectiva at the time) and by author CS (working for Grunenthal at the time). The study and all data collection were performed by ICON, independent of Grunenthal and Affectiva. The funder did not influence data collection, analysis or interpretation of the data, or influence the writing of this article or the authors' decision to submit it for publication. CS was no longer affiliated with Grunenthal at the time that she assisted in the preparation of this manuscript for submission.

ACKNOWLEDGMENTS

The authors wish to thank Dan Bender for his expertise helping set up the study, Evan Kodra for his input on analyses related to the study, and the reviewers for comments that greatly improved the presentation of this work.

REFERENCES

1. Melzack R. The McGill pain questionnaire: major properties and scoring methods. *Pain.* (1975) 1. doi: 10.1016/0304-3959(75)90044-5
2. Mao J. Translational Pain Research: Achievements and Challenges. *J Pain.* (2009) 10:1001–1011. doi: 10.1016/j.jpain.2009.06.002
3. Bjordal JM, Ljunggren AE, Klovning A, Slørdal L. Non-steroidal anti-inflammatory drugs, including cyclo-oxygenase-2 inhibitors, in osteoarthritic knee pain: Meta-analysis of randomised placebo controlled trials. *Br Med J.* (2004). doi: 10.1136/bmj.38273.626655.63
4. Haefeli M, Elfering A. Pain assessment. *Eur spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc.* (2006) 15 Suppl 1:S17–24. doi: 10.1007/s00586-005-1044-x
5. Kang Y, Demiris G. Self-report pain assessment tools for cognitively intact older adults: Integrative review. *Int J Older People Nurs.* (2018) 13:e12170. doi: 10.1111/opn.12170
6. Werner P, Lopez-martinez D, Walter S, Al-hamadi A, Gruss S, Picard RW. *Automatic Recognition Methods Supporting Pain Assessment : A Survey.* IEEE Transactions on Affective Computing. (2019) doi: 10.1109/TAFFC.2019.2946774
7. Chen J, Abbod M, Shieh JS, Pain and stress detection using wearable sensors and devices—a review. *Sensors (Switzerland).* (2021) 21:1–18. doi: 10.3390/s21041030
8. Picard RW, Member S, Vyzas E, Healey J. Toward machine emotional intelligence : analysis of affective physiological state. *IEEE Trans Pattern Anal Mach Intell.* (2001) doi: 10.1109/34.954607
9. Olesen AE, Andresen T, Staahl C, Drewes AM. Human experimental pain models for assessing the therapeutic efficacy of analgesic drugs. *Pharmacol Rev.* (2012) 64. doi: 10.1124/pr.111.005447
10. Edwards R, Ness T, Fillingim R. Endogenous opioids, blood pressure, and diffuse noxious inhibitory controls: a preliminary study. *Percept Mot Skills.* (2004) 99:679–87. doi: 10.2466/pms.99.2.679-687
11. Andresen T, Staahl C, Oksche A, Mansikka H, Arendt-Nielsen L, Drewes AM. Effect of transdermal opioids in experimentally induced superficial, deep and hyperalgesic pain. *Br J Pharmacol.* (2011) 164. doi: 10.1111/j.1476-5381.2010.01180.x
12. Dirks J, Petersen K, Dahl J. The heat/capsaicin sensitization model: a methodologic study. *J pain.* (2003) 4:122–8. doi: 10.1054/jpai.2003.10
13. Modir J, Wallace M. Human experimental pain models 3: heat/capsaicin sensitization and intradermal capsaicin models. *Methods Mol Biol.* (2010) 617:169–74. doi: 10.1007/978-1-60327-323-7_14
14. Su Q, Song Y, Zhao R, Liang M. A review on the ongoing quest for a pain signature in the human brain. (2020) 274–287. doi: 10.26599/BSA.2019.9050024
15. Wager TD, Atlas LY, Lindquist M, Roy M, Woo, C-W, et al. An fMRI-based neurologic signature of physical pain. *N Engl J Med.* (2013) 368:1388–97. doi: 10.1056/NEJMoa1204471
16. Boucsein W. *Electrodermal activity, 2nd ed*, *Techniques in psychophysiology*. New York: Springer-Verlag. (2012). doi: 10.1007/978-1-4614-1126-0
17. Craig KD, Patrick CJ. facial expression during induced pain. *J Pers Soc Psychol.* (1985) 48:1080–91. doi: 10.1037/0022-3514.48.4.1089
18. Kuscu OO, Akyuz S. Is it the injection device or the anxiety experienced that causes pain during dental local anaesthesia? *Int J Paediatr Dent.* (2008) 18:139–45. doi: 10.1111/j.1365-263X.2007.00875.x
19. Colloca L, Barsky AJ. Placebo nocebo effects. *N Engl J Med.* (2020) 382:554–61. doi: 10.1056/NEJMra1907805
20. Waber RL, Shiv B, Carmon Z, Ariely D. Commercial features of placebo and therapeutic efficacy. *JAMA.* (2008) 299:1016–7. doi: 10.1001/jama.299.9.1016
21. Rezaii T, Hirschberg A, Carlström K, Ernberg M. The influence of menstrual phases on pain modulation in healthy women. *J pain.* (2012) 13:646–55. doi: 10.1016/j.jpain.2012.04.002
22. Rhudy J, Bartley E. The effect of the menstrual cycle on affective modulation of pain and nociception in healthy women. *Pain.* (2010) 149:365–72. doi: 10.1016/j.pain.2010.02.041
23. Tommaso M. Pain perception during menstrual cycle. *Curr Pain Headache Rep.* (2011) 15:400–6. doi: 10.1007/s11916-011-0207-1
24. Duncan GH, Bushnell CM, Lavigne GJ. Comparison of verbal and visual analogue scales for measuring the intensity and unpleasantness of experimental pain. *Pain.* (1989) 37:295–303. doi: 10.1016/0304-3959(89)90194-2
25. Nahman-Averbuch H, Hughes C, Hoeppli, M-E., White K, Peugh J, Leon E, et al. Communication of pain intensity and unpleasantness through magnitude ratings: Influence of scale type, but not sex of the participant. *medRxiv.* (2021). 06.24.21259433. doi: 10.1101/2021.06.24.21259433
26. Spielberger C, Gorsuch R, Lushene R, Vagg P, Jacobs G. *Manual for the State-Trait Anxiety Inventory.* Consult Psychol Press Inc. (1983) doi: 10.1037/t06496-000
27. Arendt-Nielsen L, Yarnitsky D. Experimental and clinical applications of quantitative sensory testing applied to skin, muscles and viscera. *J pain.* (2009) 10:556–72. doi: 10.1016/j.jpain.2009.02.002

28. Edwards R, Sarlani E, Wesselmann U, Fillingim R. Quantitative assessment of experimental pain perception: multiple domains of clinical relevance. *Pain.* (2005) 114:315–9. doi: 10.1016/j.pain.2005.01.007
29. Frampton CL, Hughes-Webb P. The measurement of pain. *Clin Oncol.* (2011) 23. doi: 10.1016/j.clon.2011.04.008
30. Giordano J, Abramson K, Boswell M. Pain assessment: subjectivity, objectivity, and the use of neurotechnology. *Pain Physician.* (2010) 13:305–15. doi: 10.36076/ppj.2010/13/305
31. Breivik H, Borchgrevink PC, Allen SM, Rosseland LA, Romundstad L, Breivik Hals EK, et al. Assessment of pain. *BJA Br J Anaesth.* (2008) 101:17–24. doi: 10.1093/bja/aen103
32. Fillingim R, Loeser J, Baron R, Edwards R. Assessment of chronic pain: domains, methods, and mechanisms. *J pain.* (2016) 17:T10–20. doi: 10.1016/j.jpain.2015.08.010
33. Handwerker H, Kobal G. Psychophysiology of experimentally induced pain. *Physiol Rev.* (1993) 73:639–71. doi: 10.1152/physrev.1993.73.3.639
34. Mangina CA, Beuzeron-Mangina JH. Direct electrical stimulation of specific human brain structures and bilateral electrodermal activity. *Int J Psychophysiol.* (1996) 22. doi: 10.1016/0167-8760(96)00022-0
35. Poh MZ, Loddenkemper T, Reinsberger C, Swenson NC, Goyal S, Madsen JR, et al. Autonomic changes with seizures correlate with postictal EEG suppression. *Neurology.* (2012) 78. doi: 10.1212/WNL.0b013e318258f7f1
36. Picard RW, Migliorini M, Caborni C, Onorati F, Regalia G, Friedman D, et al. Wrist sensor reveals sympathetic hyperactivity and hypoventilation before probable SUDEP. *Neurology.* (2017) 89:633–5. doi: 10.1212/WNL.0000000000004208
37. Storm H. Skin conductance and the stress response from heel stick in preterm infants. *Arch Dis Child Fetal Neonatal Ed.* (2000) 83:F143–7. doi: 10.1136/fn.83.2.F143
38. Doberenz S, Roth W, Wollburg E, Maslowski N, Kim S. Methodological considerations in ambulatory skin conductance monitoring. *Int J Psychophysiol.* (2011) 80:87–95. doi: 10.1016/j.ijpsycho.2011.02.002
39. Johnson KT, Picard RW. Advancing neuroscience through wearable devices. *Neuron.* (2020) 108:8–12. doi: 10.1016/j.neuron.2020.09.030
40. Kasos K, Csirmaz L, Vikor F, Zimonyi S, Varga K, Szekely A. Electrodermal correlates of hypnosis: current developments. *OBM Integr Complement Med.* (2020) 5:1–20. doi: 10.21926/obm.icm.2002017
41. Sano A, Picard RW, Stickgold R. Quantitative analysis of wrist electrodermal activity during sleep. *Int J Psychophysiol.* (2014) 94. doi: 10.1016/j.ijpsycho.2014.09.011
42. van Dooren M, de Vries JJGGJ, Janssen JH. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiol Behav.* (2012) 106. doi: 10.1016/j.physbeh.2012.01.020
43. O'Leary HM, Marschik PB, Khwaja OS, Ho E, Barnes KV, Clarkson TW, et al. Detecting autonomic response to pain in Rett syndrome. *Dev Neurorehabil.* (2017) 20:108–14. doi: 10.3109/17518423.2015.1087437
44. Schumm J, Bächlin M, Setz C, Arnrich B, Roggen D, Tröster G. Effect of movements on the electrodermal response after a startle event. *Proc. 2nd Int. Conf. Pervasive Comput. Technol. Healthc. PervasiveHealth*. (2008) 315–318. doi: 10.1109/PCTHEALTH.2008.4571101
45. Picard RW, Fedor S, Ayzenberg Y. Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emot Rev.* (2016) 8. doi: 10.1177/1754073914565517
46. Posada-Quintero HF, Chon KH. Innovations in electrodermal activity data collection and signal processing: a systematic review. *Sensors (Basel).* (2020) 20. doi: 10.3390/s20020479
47. Posada-Quintero HF, Florian JP, Orjuela-Cañón AD, Aljama-Corrales T, Charleston-Villalobos S, Chon KH. Power Spectral Density Analysis of Electrodermal Activity for Sympathetic Function Assessment. *Ann Biomed Eng.* (2016) 44:3124–35. doi: 10.1007/s10439-016-1606-6
48. Posada-Quintero HF, Kong Y, Chon KH. Objective pain stimulation intensity and pain sensation assessment using machine learning classification and regression based on electrodermal activity. *Am J Physiol Regul Integr Comp Physiol.* (2021) 321:R186–96. doi: 10.1152/ajpregu.00094.2021
49. Posada-Quintero HF, Dimitrov T, Moutran A, Park S, Chon KH. Analysis of reproducibility of noninvasive measures of sympathetic autonomic control based on electrodermal activity and heart rate variability. *IEEE Access.* (2019) 7:22523–31. doi: 10.1109/ACCESS.2019.2899485
50. Gjerstad AC, Wagner K, Henrichsen T, Storm H. Skin conductance versus the modified COMFORT sedation score as a measure of discomfort in artificially ventilated children. *Pediatrics.* (2008) 122:e848–53. doi: 10.1542/peds.2007-2545
51. Aqajari SAH, Cao R, Kasaeyan Naeini E, Calderon M-D, Zheng K, Dutt N, et al. Pain assessment tool with electrodermal activity for postoperative patients: method validation study. *JMIR mHealth uHealth.* (2021) 9:e25258. doi: 10.2196/25258
52. Kong Y, Posada-Quintero HF, Chon KH. Sensitive physiological indices of pain based on differential characteristics of electrodermal activity. *IEEE Trans Biomed Eng.* (2021) 68:3122–30. doi: 10.1109/TBME.2021.3065218

Conflict of Interest: RP is employed by MIT as a full professor and consults for Empatica Inc, where she is a co-founder and shareholder, serves as Chairman of the Board and consults part-time as Chief Scientist. RP receives royalties on patents for her AI-related and wearables inventions owned by MIT. RP also serves as an expert witness representing Apple, Inc. in matters related to wearable technologies. At MIT, her lab's research receives funding from the NIH via the Massachusetts General Hospital Depression Research Clinic and from a consortium of over fifty companies listed at https://www.media.mit.edu/posts/member-companies. She receives speaker fees through Stern Strategy. These relationships present no conflict of interest at the time of this paper submission. VB is self-employed. CS is employed by Novo Nordisk A/S. CS was employed by Grünenthal GmbH during the design and conduct of the trial. This study received funding from Grünenthal GmbH. The funder had the following involvement with the study: assisting in designing the trial, overseeing the trial conduct, and engaging in discussions around the data analysis.

frontiers | Frontiers in Pain Research

TYPE Brief Research Report
PUBLISHED 24 October 2022
DOI 10.3389/fpain.2022.969867

OPEN ACCESS

EDITED BY
Mathieu Piché,
Université du Québec à Trois-Rivières, Canada

REVIEWED BY
Mathieu Roy,
McGill University, Canada
Nobuhiro Watanabe,
Tokyo Metropolitan Institute of Gerontology, Japan

*CORRESPONDENCE
Giovanna Stella
giovanna.stella@phd.unict.it

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

RECEIVED 15 June 2022
ACCEPTED 26 September 2022
PUBLISHED 24 October 2022

CITATION
Bucolo M, Rance M, Nees F, Ruttorf M, Stella G, Monarca N, Andoh J and Flor H (2022) Cortical networks underlying successful control of nociceptive processing using real-time fMRI. Front. Pain Res. 3:969867. doi: 10.3389/fpain.2022.969867

Cortical networks underlying successful control of nociceptive processing using real-time fMRI

Maide Bucolo[1], Mariela Rance[2], Frauke Nees[2,3], Michaela Ruttorf[4], Giovanna Stella[1]*, Nicolò Monarca[1], Jamila Andoh[2,5†] and Herta Flor[1,2†]

[1]Department of Electrical Electronic and Computer Engineering, University of Catania, Catania, Italy, [2]Institute of Cognitive and Clinical Neuroscience, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany, [3]Institute of Medical Psychology and Sociology, University Medical Center Schleswig-Holstein, Kiel, Germany, [4]Computer Assisted Clinical Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany, [5]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

Real-time fMRI (rt-fMRI) enables self-regulation of neural activity in localized brain regions through neurofeedback. Previous studies showed successful up- and down-regulation of neural activity in the anterior cingulate cortex (ACC) and the insula (Ins) during nociceptive stimulation. Such self-regulation capacity is, however, variable across subjects, possibly related to the ability of cognitive top-down control of pain. Moreover, how specific brain areas interact to enable successful regulation of nociceptive processing and neurofeedback-based brain modulation is not well understood. A connectivity analysis framework in the frequency domain was used to examine the up- or down-regulation in the ACC and Ins and pain intensity and unpleasantness ratings were assessed. We found that successful up- and down-regulation was mediated by the ACC and by its functional connectivity with the Ins and secondary somatosensory cortex. There was no significant relationship between successful up- or downregulation and pain ratings. These findings demonstrate functional interactions between brain areas involved in nociceptive processing during regulation of ACC and Ins activity, and the relevance of the frequency domain connectivity analysis for real-time fMRI. Moreover, despite successful neural regulation, there was no change in pain ratings, suggesting that pain is a complex perception, which may be more difficult to modify than other sensory or emotional processes.

KEYWORDS

brain connectivity, regulation of neural activity, nociceptive processing, coherence analysis, fMRI

Introduction

Real time functional magnetic resonance imaging (rt-fMRI) permits the feedback of neuronal activity, which can then be controlled and regulated. rt-fMRI has been well established over the past 15 years (1, 2) and has often been associated with behavioral changes (3), including pain perception (4), although this could not be consistently replicated (5). Brain responses to nociceptive processing have been shown to involve

areas such as primary (SI) and secondary (SII) somatosensory cortices, insula (Ins), the anterior (ACC) or the mid-cingulate (MCC) cortices (6, 7).

The prefrontal cortex (PFC) has also been shown to be involved in pain processing but may be more important in chronic than acute pain (8, 9), and therefore may not be an ideal target for neuromodulation. In addition, the PFC structure is quite complex, and includes different regions, namely, dorsal, medial and ventral prefrontal cortices, involved in various aspects of pain processing, e.g., intensity of pain, spatial aspects of pain processing, emotion regulation, but also involved in various cognitive processes such as attention or decision making (10, 11).

The rostral anterior cingulate cortex (rACC) in particular, has been involved in pain regulation (12, 13), and is therefore a target of choice for rt-fMRI studies (4). In a previous study, we showed that participants could successfully downregulate neural activity related to nociceptive processing in the rACC and the posterior insula (pIns) and upregulate pInsL but not rACC (14). Upregulation or downregulation of either region was unrelated to pain intensity or unpleasantness ratings. The ability to successfully regulate brain activity was also shown to be variable across participants (14), which might be related to lack of cognitive top-down control of pain and deserves further investigation.

We also showed that lower covariation between the two regions correlated positively with the training effect and thus learning, suggesting that the state of the network involved in the processing of pain should be considered in the modulation of pain-evoked activation and related behavioral effects (15). Therefore, in this study, we aimed at examining functional connectivity in pInsL and in rACC and their effect on learning.

In addition, it is unclear how brain areas interact to enable successful regulation of nociceptive processing (16) and neurofeedback-based brain modulation (17, 18). For example, Hinterberger et al., analyzed successful regulation of slow cortical potentials and found that a number of brain regions were involved in successful regulation with a focus on sensorimotor and frontal control regions (19).

We aimed to assess the temporal dependence of activation patterns between brain regions, specifically, the functional connectivity of regulation- and pain-associated brain regions during up- or down-regulation of neural activity related to nociceptive processing. The methodological framework used here to evaluate functional connectivity combines signal processing from data-driven mathematical methods and complex network analysis (20). This integrated approach has previously been applied to various brain signals from electroencephalographic (EEG), magnetoencephalographic (MEG) (21, 22), and functional magnetic resonance imaging (fMRI) data (23). Despite the proliferation of mathematical methods and toolboxes (24–26), there is no general consensus on the most robust and efficient way to assess functional connectivity (27). The use of different computational parameters such as the frequency range, time lag, or the choice of a significance threshold can affect the results at an individual and group level analysis (28, 29). Group analyses allow statistical measures on the validity of the result, but are still affected by sample size and individual variability in complex brain activity and can sometimes hide relevant key brain mechanisms. We investigated some of these issues in our previous works (30, 31) and established a connectivity analysis framework in the frequency domain that we used in the present study. We assessed the effect of self-regulation of the activity from two target ROIs, the rACC and pIns and examined functional connectivity to other areas such as the somatosensory cortex (SII), the anterior and posterior insula (aIns, pIns) and MCC. Based on its role in top-down control, we expected the ACC to play a key role in successful regulation of nociceptive processing and to show functional connections to SII, pIns, MCC.

We added the posterior insula, because it has been involved in the sensory pain aspects (32–34) and was associated with a reliable activation pattern across all subjects (14).

Materials and methods

Participants

Ten healthy right-handed participants were enrolled in the study [mean age, standard deviation $M = 29.0$, $SD = 6.48$, range (20, 41)], four females ($M = 27.0$, $SD = 3.92$), and six males ($M = 30$, $SD = 7.81$). Exclusion criteria were cardiovascular or neurological disorders, brain injury, acute pain, current analgesic medication, pregnancy, lifetime and current substance abuse or dependence, any mental disorder, and metallic implants. The study adhered to the Declaration of Helsinki and was approved by the Ethics Committee of the Medical Faculty Mannheim, Heidelberg University, Germany. All subjects gave written informed consent after a detailed description of the complete study. The sample of this study is identical to that described in a previous study (14). Here we reanalysed the data with respect to patterns of connectivity in Learners and non-Learners of neurofeedback control.

fMRI neurofeedback procedure

The neurofeedback protocol consisted of a baseline run and 24 training trials spread over the course of 4 consecutive days. On the first day, the participants were introduced to the experimental setup and protocol, and the baseline run was recorded. Each session (training day) consisted of six successive training trials; each trial of 7 min was composed of

six regulation phases, each lasting 45 s and seven non-regulation phases, each lasting 22.5 s, evenly distributed across each session. The sequence of regulation and non-regulation phases is depicted, the overall duration of a trial is 7 min and consists of 258 samples based on the acquisition time TR = 1.5 s (sampling frequency 0.66 Hz). On the right of **Figure 1**, as an example, the time-series extracted from a ROI fMRI is shown. During the regulation phases, 2 ms of painful electrical stimulation at a frequency of 2 Hz were carried out using a digitimer DS7A stimulator and applied over the fourth digit of the right hand using concentric bipolar electrodes (see **Figure 1A**).

Individual detection and pain thresholds were determined by the method of limits, averaging over the last two of three ascending and descending stimulation sequences (15). Pain tolerance was averaged over the last two of three ascending stimulation sequences. Stimulation strength was set at 70% between pain threshold and pain tolerance and adjusted to be rated between 6 and 7 on an 11 point verbal rating scale (ranging from 0 = no pain to 10 = strongest imaginable pain), allowing for a possible increase or decrease of perceived pain strength. The individually adjusted mean stimulation strength was 2.27 mA (SD = 1.76), the pre baseline intensity of this stimulus was rated as 6.40 (SD = 0.61) and the unpleasantness was assessed on a verbal rating scale (raining from 0 = not unpleasant to 10 = extremely unpleasant) amounting to 6.70 (SD = 1.32). The postbaseline stimulus intensity was rated 6.10 (SD = 1.68) and the pain unpleasantness 7.25 (SD = 1.51).

The visual feedback consisted of a moving blue or yellow ball in front of a black background (**Figure 1C**).

During the regulation phases of the training trials, a stationary white arrow appeared next to the ball on the left side of the screen indicating the vertical direction in which the ball should be moved. Movements of the ball corresponded to changes in the computed BOLD signal from the regions of interest (ROI), i.e., rACC or pInsL and a control ROI (UNR), with activity unrelated to the nociceptive stimulation or pain processing (located in the parietal lobe, bordering the occipital lobe and the height of pInsL), see (14) for detailed information, **Figure 1B**. The target ROIs, i.e., rACC or pInsL were discernible by the colour of the moving ball (blue or yellow) for rACC and pInsL and the colour was randomized across participants. The baseline run was similar to the training trial, the subjects were presented a stationary white ball on the screen, but no visual feedback was given. The participants were instructed that the vertical change of the blue or yellow ball was an indicator of their own brain activity in selected brain regions and that they would be able to observe the changes with a delay of a few seconds. The subjects were allowed to use any kind of strategy that would not involve body movement (e.g., muscle tension or relaxation). During the non-regulation phases, i.e., in the absence of visual feedback, the participants were told to perform simple mental arithmetic for the purpose of stopping regulation attempts and ensuring comparability across subjects.

MRI acquisition

MRI data were acquired on a 3 T MAGNETOM Trio TIM whole body scanner using a standard 12-channel head coil (Siemens Medical Solutions, Erlangen, Germany). fMRI data were acquired using gradient-echo and echo-planar imaging (EPI) sequence (TR/TE = 1,500/22 ms, matrix size 96 × 96, flip angle 90o, and bandwidth BW = 1,270 Hz/px). Twenty-four AC/PC aligned slices were acquired with voxel size 2.2 mm × 2.2 mm × 3.5 mm and 0.5 mm gap. A three-dimensional fast low angle shot high-resolution T1-weighted anatomical scan was also acquired for each participant (TR/TE = 23/5.02 ms, matrix size 448 × 448, flip angle 25o, BW 190 Hz/px, voxel size 0.5 mm × 0.5 mm × 1.0 mm) as anatomical reference.

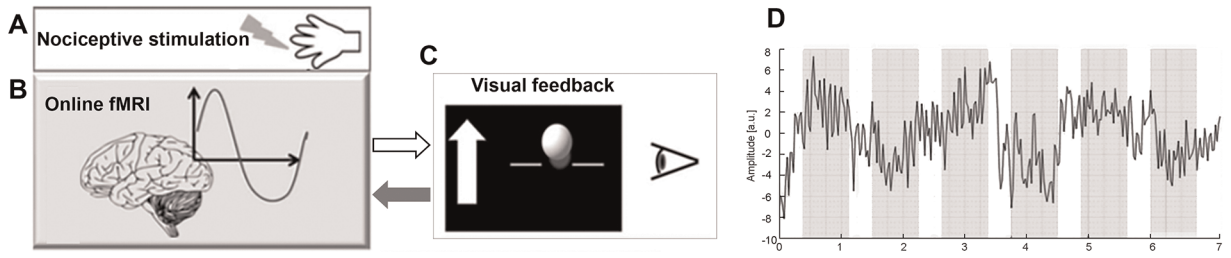


FIGURE 1
Schema of the fMRI neurofeedback setup. (**A**) Painful electrical stimuli were applied over the fourth digit of the right hand. (**B**) Online pre-processing and statistical analysis of the BOLD response were carried out from the target ROIs (rACC, pInsL) and a control ROI (UNR). (**C**) Differences in BOLD response between target ROIs and UNR ROI were computed and represented by a moving ball in front of a black background on a display screen. On the left side of the ball, a white arrow was displayed with the up or down directions depending on if the participants had to up- or down-regulate neural activity from the target ROIs. (**D**) Trend of a ROI time series in a trial of 7 min where the regulation (grey) and non-regulation (white) phases are highlighted.

Foam pegs (Siemens Medical Solutions, Erlangen, Germany) were used to immobilize the subject's head during MR scanning.

MRI pre-processing and statistical analysis

Online fMRI data pre-processing and statistical analyses

Brain responses to nociceptive stimulation were recorded and analysed in real-time during the fMRI acquisition using Turbo BrainVoyager Version 1.1 (Brain Innovation, Maastricht, TheNetherlands) as described in (2). The mean BOLD signal change from two target regions of interest (ROIs), i.e., rACC and pInsL was compared with a control region ("UNR"), see Figure 2 and Table 1 for coordinates. The feedback signal was calculated as the difference of the percent BOLD signal change between one of the target ROIs and the UNR ROI (Figure 1B) and visually fed back to the subject in the form of a moving ball on a screen (Figure 1C). The feedback computation and visualization were performed with in-house written scripts based on Presentation® Version 13.0 Build 01.23.09 (Neurobehavioral Systems Inc., Albany, CA, USA) on another computer connected with Turbo BrainVoyager *via* LAN. The location of the rACC, pInsL and UNR regions was determined in an offline analysis of the baseline run. The criteria for the target ROIs were (a) a position over the most significant cluster active during the stimulation phase and not active during the non-regulation phase and (b) being at the respective areas in the rACC and pInsL regions (15).

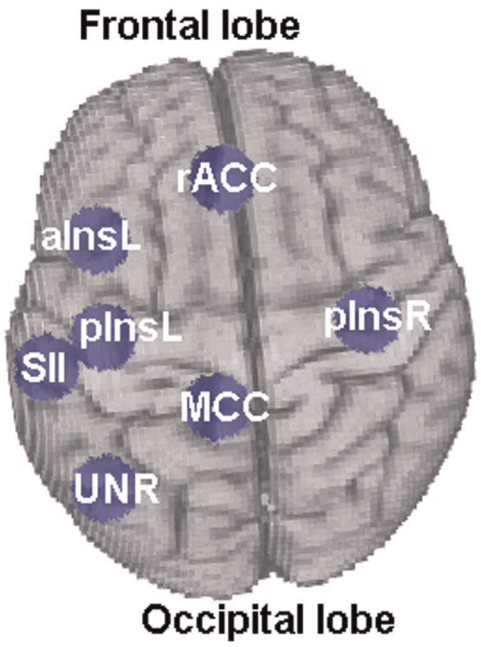


FIGURE 2
Location of the six ROIs (rACC, pInsL, pInsR, MCC, aInsL, SII) used for functional connectivity analyses. The ROI "UNR" corresponds to the control ROI and is shown for illustration purposes.

TABLE 1 MNI coordinates of the regions of interest used in the manuscript.

	MNI coordinates		
ROI	**x**	**y**	**z**
rACC	−1	36	6
pInsL	−42	−20	9
pInsR	40	−13	6
MCC	−3	−25	38
aInsL	−45	6	−6
SII	−58	−29	23
UNR	−42	−61	29

rACC, rostral anterior cingulate cortex; pInsL, left posterior insula; pInsR, right posterior insula; SII, secondary somatosensory cortex; MCC, medial cingulate cortex; UNR, parieto-occipital area (control ROI).

TABLE 2 The four conditions for the generation of the feedback signal and their association with the ball displacement on the visual feedback.

Condition	Feedback	Ball displacement
1	ACCD: rACC—UNR < 0	Down
2	ACCU: rACC—UNR > 0	Up
3	INSD: pIns—UNR < 0	Down
4	INSU: pIns—UNR > 0	Up

(1) ACCD represents down-regulation of the BOLD activity in the rACC (compared with the UNR ROI). (2) ACCU represents up-regulation of the BOLD activity in the rACC (compared with the UNR ROI). (3) INSD represents down-regulation of the BOLD activity in the pINSL (compared with the UNR ROI). (4) INSU represents up-regulation of the BOLD activity in the INS (compared with the UNR ROI).

The positioning of the UNR region was also monitored online during the training trials to not exhibit significant activation or deactivation. Four feedback conditions were used (see Table 2), assuming that for each target ROI the activity should be larger or smaller than the activity of the UNR region, displayed with up or down vertical directions in the ball displacement on the screen. Table 2 summarizes the four feedback conditions: ACCD, ACCU, INSD, INSU, the feedback signal computation (i.e., control ROI UNR), and the ball displacement on the screen.

Offline fMRI data pre-processing and statistical analyses

The offline data pre-processing of the fMRI scans was performed using BrainVoyager QX 2.3 (Brain Innovation, Maastricht, The Netherlands, Goebel, 2001). Time courses for the brain connectivity analysis were extracted using an offline

GLM analysis of all training fMRI datasets from six ROIs: the two target ROIs (rACC, pInsL), and four other ROIs selected for their involvement with pain processing: the medial cingulate gyrus (MCC), the right posterior insula (pInsR), the left anterior insula (aInsL), the left secondary somatosensory cortices (SIIL), see **Figure 2** for their anatomical location.

Classification of Learners and non-Learners were based on the following criteria (14), see also **Table 2**:

- For conditions 1 and 2 (i.e., ACCD and ACCU), if the average difference of the activation of rACC and UNR was negative and positive respectively, and if this was the case for at least four out of six training trials, and if the modulation effect for the specific condition improved from trial 1 to trial 6, a subject was considered a Learner.
- For conditions 3 and 4 (i.e., INSD and INSU), if the average difference of the activation of pInsL and UNR was negative and positive respectively, and if this was the case for at least four out of six training trials, and if the modulation effect for the specific condition improved from trial 1 to trial 6, a subject was considered a Learner.

The other subjects were categorized as non-Learners.

Functional connectivity analysis

The Coherence function was used to evaluate the functional connectivity between the six ROIs for each of the four conditions (see **Figure 3**). The connectivity matrix (CM), which represents the level of inter- relationship between pairs of brain areas, was computed as a measure of the linear independence in the frequency domain between pairs of brain time-courses. Considering the six ROIs, a CM (6×6) was obtained and the values in the diagonal associated with the ROI self-similarity were set to zero. Based on the sampling rate (TR = 1.5 s), the frequency range investigated was [0.0012–0.33] Hz considering the limit due to the Nyquist theorem ($f_{MAX} = 2/TR$). It was scanned with a frequency step of $f_{step} = f_{MIN} = (1/N)\ f_{MAX} = 0.0012$ Hz where $N = 258$ is the number of samples acquired. The CMs for each subject for trials 1, 6 per condition were computed by averaging the overall frequency range. In the group analysis the connectivity matrices for the Learner and non-Learner groups were obtained by normalizing the CMs in the range [0, 1] and then averaging all of them by group. The normalization was necessary to make the results independent of subject variabilities of the CM levels. In the last step a cut-off threshold was established to extract the four strongest connections. Those connections were then plotted using a weighted graph representation for a visual inspection. In this graph the nodes are representative of the ROIs, and connections between ROIs pairs are for the CM weights above the defined threshold, otherwise the ROIs remain unconnected. Connectivity strengths were defined by weights of functional connectivity obtained from the coherence analysis. The colours code the strength of the connection between the ROIs. Colors from blue to red indicate weak to

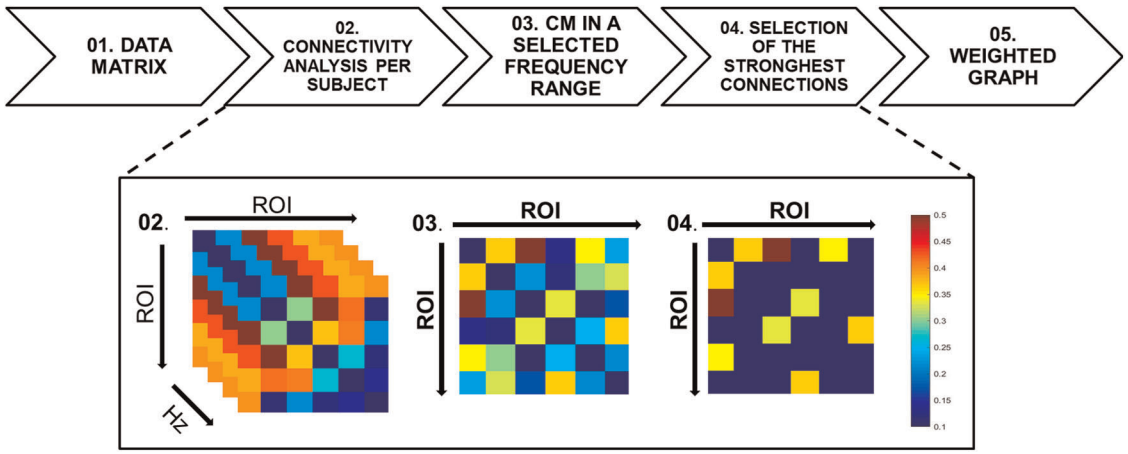


FIGURE 3
Flowchart of the functional connectivity analysis. (Top) The sequence of mathematical method used from the fMRI data matrix (step 1) to the generation of the weighted graphs (step 5). (Bottom) An example of fMRI data transformation from step 1 to 5. Step 1: for each participant, the fMRI data are represented by a matrix with dimensions $N \times M$, in which the N dimension represents the ROIs and the M dimension represents number of samples of the time series ($N = 6$ and $M = 258$). Step 2: the coherence mathematical method is performed and a frequency-varying CM in the overall frequency range [0.0012–0.33] Hz is obtained with a frequency step of 0.0012 Hz. Step 3: After the average of the time-varying CM in the selected frequency range of interest [0.15–0.33] Hz, a single CM for each participant is obtained. For group analyses, the single subject CMs are first normalized in the range [0, 1], then averaged to obtain a group-level CM. Step 4: The CM is thresholded using the percentile thresholding method in order to enhance the strongest connections. Step 5: the CM is then used as an adjacency matrix for the graph extraction. The color bar indicates the connection strengths.

strong connection strength respectively and are expressed in arbitrary units (a.u).

In Figure 3 the analysis pipeline is shown as a flow chart representing the connectivity analysis from the data matrix to the CMs averaging in the frequency range and the threshold selection to evidence the strongest links. The simplest method for estimating functional connectivity in the frequency domain is Coherence analysis (35). The Coherence between two individual time-series (y_i, y_j) over a frequency range f is defined as follows:

$$Coh2i,\, j(f) = \frac{E[|(C_{i,j}(f)|^2]}{E[|C_{i,i}(f)|]*E[|C_{j,j}(f)|]}$$

$$C_{i,j}(f) = \; Yi(f)Yj*(\mathrm{f})$$

The squared coefficient of Coherence can be interpreted as the proportion of the power in one of the two time-series (at a selected frequency), which can be explained by its linear relation with the other time course. Coherence is a positive function bounded by [0, 1] and symmetric in i and j. A measure of Coherency, such as an average over a frequency band, is capable of detecting zero-time lag synchronization and fixed time non-zero-time lag synchronization, which may occur when there is a significant delay between two brain sites. However, it does not provide any information on directionality of the coupling between the two recording sites.

Figure 4 shows for a representative Learner (Subject-4), the trends of Coherence obtained for the ROI rACC paired with the ROIs [pInsL, pInsR, SIIL] versus frequency. We note a high variability in the frequency range. The analyses were then carried out averaging the information in the range [0.15–0.33] Hz. No change in the results was detected when changing this range. Due to the CM symmetry, we selected the four strongest connections over fifteen, the threshold value was set based on the 70th percentile and the CM was represented by a weighted and undirected graph.

In addition, we also assessed possible regulation of the control ROI, i.e., UNR ROI. We extracted time courses from the UNR ROI and calculated %BOLD signal change and compared it with rACC for both ACCD and ACCU conditions and with pInsL for both INSD and INSU conditions using paired t-tests (R package version 1.3.1093).

Finally, we compared pain intensity and unpleasantness ratings from the last training sessions (where the maximum effect would be expected) between Learners and non-Learners for all conditions using one way ANOVA (R package version 1.3.1093).

Results

The functional connectivity analysis was investigated in the frequency domain at the group level and at a single subject level for the target ROIs in rACC and pInsL. The results are shown for the first trial (trial 1) and for the last trial (trial 6), the latter considered to be the trial when individuals had learned to regulate neural activity for ACC or Ins. The data showed

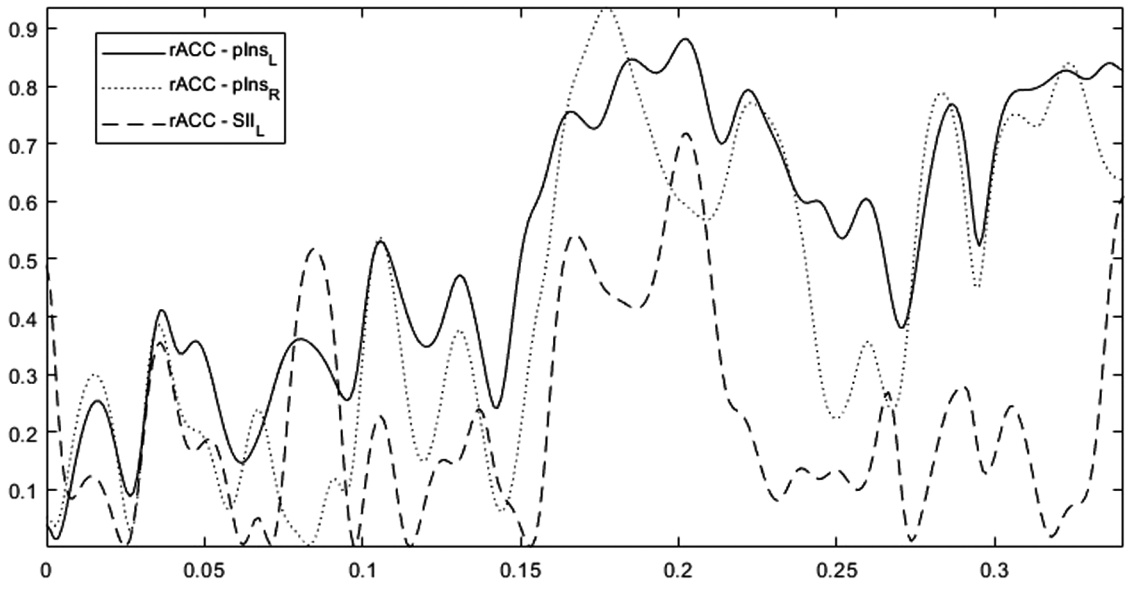


FIGURE 4
The trends of coherence versus frequency for the rACC paired with the ROIs pInsL, pInsR, SIIL for a representative learner (subject-4). The overall frequency range is [0.0012–0.33] Hz with a frequency step of 0.0012 Hz.

high inter-subject variability in the learning outcome, which is not represented in the group results. We therefore show also individual results with the best and worst learning outcome in order to provide additional insights in the differences in connectivity. At the single subject level, we selected two participants, one representative of the Learner group (Subject-4) and the other one representative of the non-Learner group (Subject-8). The brain networks underlying the learning process were identified and compared among each other and with the group networks.

On average each condition includes four or five participants, with the exception of condition (INSD) in which the Learner and the non-Learner groups included seven and two participants, respectively (Table 3).

Differences between Learners and non-Learners were found after normalization of the connectivity matrices in the range [0, 1] and thresholded using the percentile thresholding method to select the strongest connections. After such normalization and thresholding procedures, for ACCD, data from 4/6 Learners and 4/4 non-Learners remained.

TABLE 3 Classification of participants in the learner or non-learner groups for each condition (ACCD, ACCU, INSD, INSU).

Condition	Learner	non-Learner group
ACCD	[3, 4, 5, 6, 7, 10]	[1, 2, 8, 9]
ACCU	[1, 2, 4, 6, 9]	[3, 5, 7, 8, 10]
INSD	[2, 3, 4, 5, 6, 8, 9, 10]	[1, 7]
INSU	[1, 2, 4, 6, 7]	[3, 5, 8, 9, 10]

For ACCD we found a connection between rACC and pInsL, mean connectivity indices between rACC and pInsL were (mean ± SD) 0.34 ± 0.25 for Learners and 0.30 ± 0.16 for non-Learners. For ACCU, data from 4/5 Learners and 5/5 non-Learners remained, and mean connectivity indices between rACC and pInsR were 0.61 ± 0.20 for Learners and 0.35 ± 0.19 for non-Learners (resp., Figures 5A, 6A).

In addition, for both ACCD and ACCU conditions and for both Learners and non-Learners, there was a network of three nodes, composed of SII, pInsL and pInsR that were interconnected at a group level (Figures 5A, 6A).

We also found a connection between left and right pIns for both ACCD and ACCU conditions for both Learners and non-Learners and for both group- and single-subject levels (Figures 5A,B, 6A,B).

In the non-Learner group for the ACCD and ACCU conditions, there was no connection between ACC and pInsL or pInsR, neither at a group level (Figure 6A) nor at a single subject level (Figure 6B). In addition, a connection between SII and aInsL is present for ACCD at a group and single subject level (Figure 6A left and Figure 6B left) and for ACCU at a single subject level (Figure 6B right).

For INSD, after data normalization and thresholding, data from 7/8 Learners and 2/2 non-Learners remained. We found a connection between rACC and pInsL, mean connectivity indices between rACC and pInsL were (mean ± SD) 0.39 ± 0.20 for Learners and 0.23 ± 0.13 for non-Learners (resp., Figures 7A, 8A). In the non-Learners we found a connection between rACC and pInsR, mean connectivity indices between rACC and pInsR were 0.25 ± 0.08 for non-Learners and

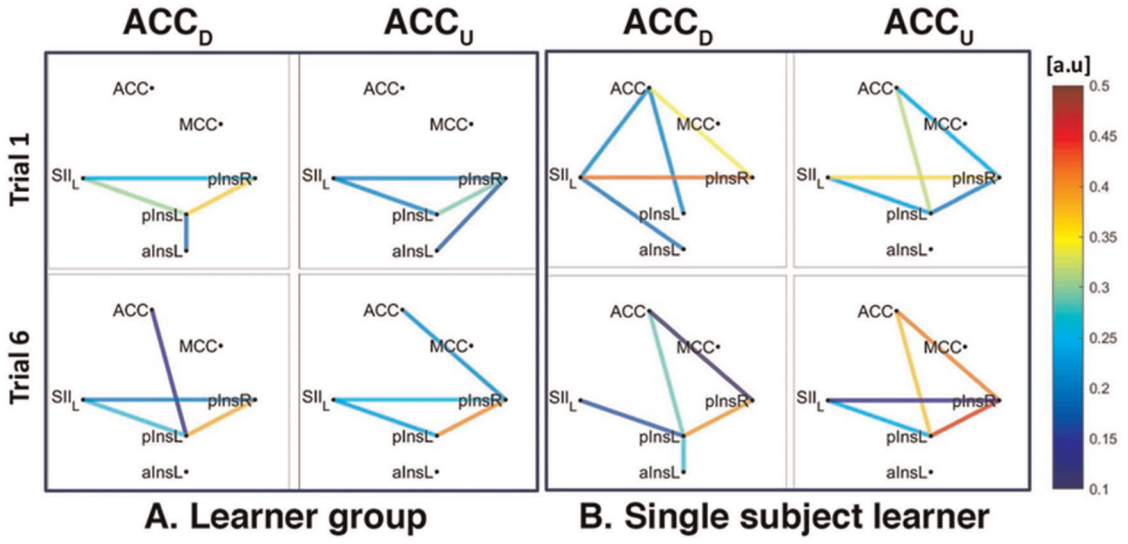


FIGURE 5
Connectivity strengths identified using the Coherence analysis for the ACC condition (ACC_D, ACC_U) in the Learner group (A) and for a representative Learner (B) in the first trial (Trial 1, top) and in the last trial (Trial 6, bottom). The colours from blue to red indicate weak to strong connection strength and are expressed in arbitrary units (a.u).

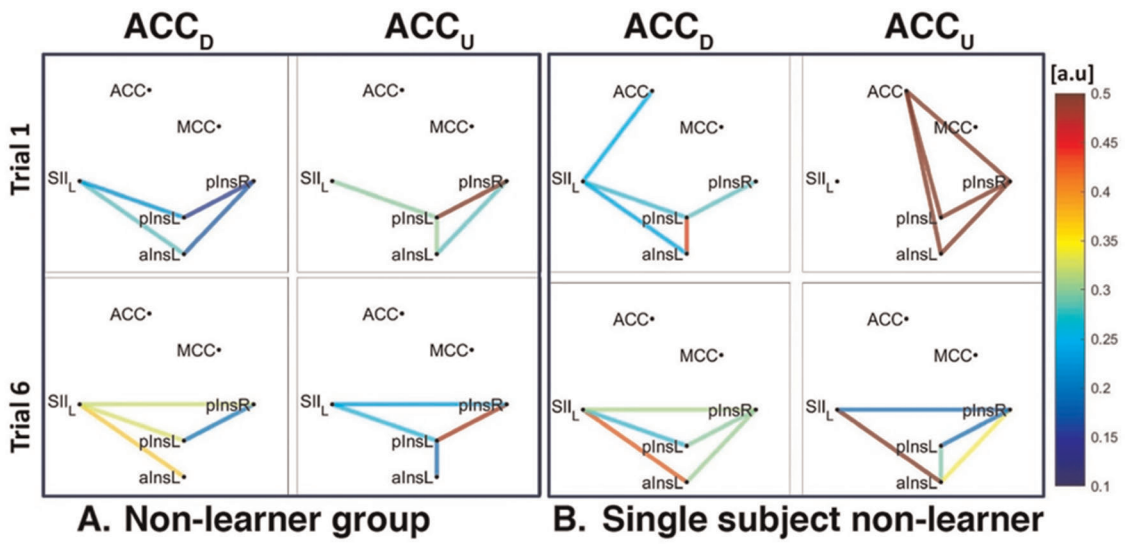


FIGURE 6
Connectivity strengths identified using the Coherence analysis for the ACC condition (ACC_D, ACC_U) in the non-Learner group (A) and for a representative non-Learner (B) in the first trial (Trial 1, top) and in the last trial (Trial 6, bottom). The colors from blue to red indicate weak to strong connection strength respectively and are expressed in arbitrary units (a.u).

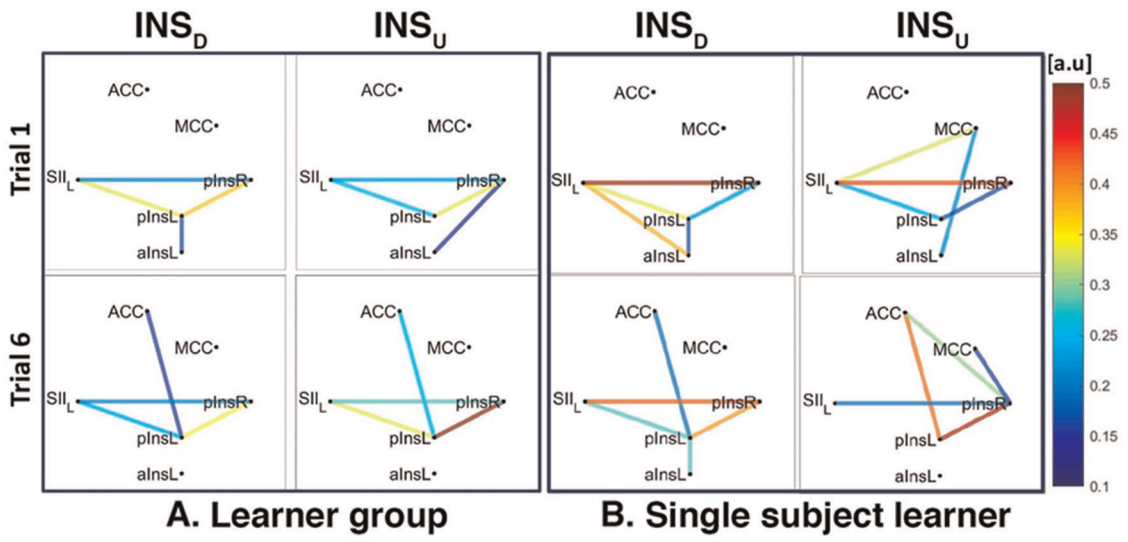


FIGURE 7
Connectivity strengths identified using the Coherence analysis for the Ins condition in the Learner group (A) and for a representative Learner (B) in the first trial (Trial 1, top) and in the last trial (Trial 6, bottom). The colors from blue to red indicate weak to strong connection strength respectively and are expressed in arbitrary units (a.u).

0.14 ± 0.06 for non-Learners (resp., Figures 7A left, Figure 8A left).

For INSU, data from 4/5 Learners and 3/6 non-Learners remained. We found a connection between rACC and pInsL for the Learners, mean connectivity indices between rACC and pInsL were (mean ± SD) 0.29 ± 0.47 for Learners and 0.14 ± 0.11 for non-Learners (resp., Figures 7A, 8A).

The network of three interconnected nodes (SII, pInsL, pIns) was also found for INSD and INSU at a group level for both Learners and non-Learners (Figures 7A, 8A). This

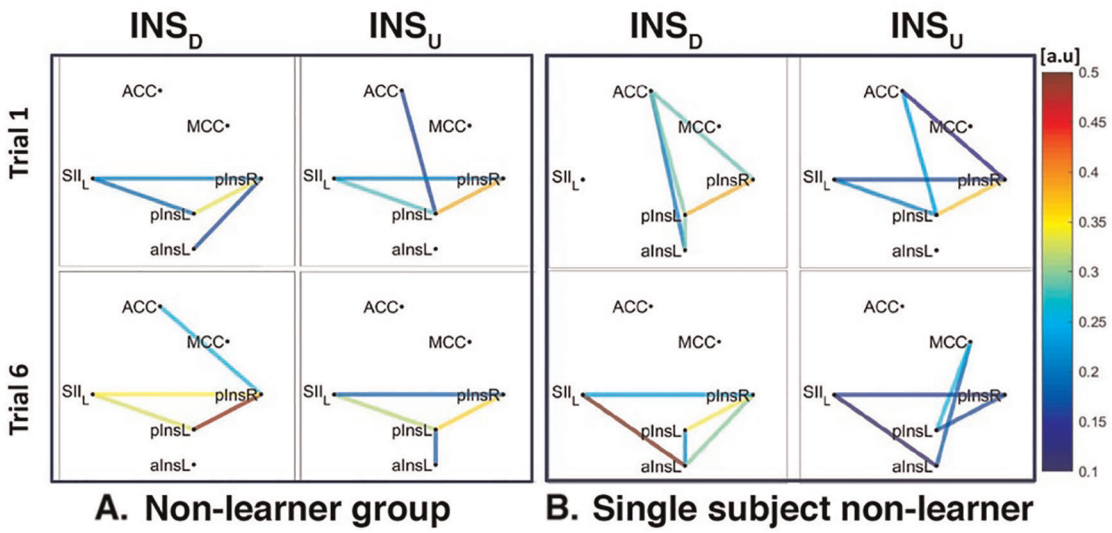


FIGURE 8
Connectivity strengths identified using the Coherence analysis for the Ins condition for the non-Learner group (A) and for a representative non-Learner (B) in the first trial (Trial 1, top) and in the last trial (Trial 6, bottom). The colors from blue to red indicate weak to strong connection strength respectively and are expressed in arbitrary units (a.u).

network was also found for the single subject Learner for INSD (Figure 7A left) but not for INSU and not for the single subject non-Learner (Figure 7B right, Figure 8B).

Regulation of the UNR ROI was significantly smaller compared with regulation of rACC for ACCD condition [$t(1,9) = 5.62$, $p < 0.001$], mean %BOLD signal change ± SD (UNR ROI: 0.03 ± 1.27 and rACC: 0.49 ± 1.79).

Regulation of the UNR ROI was significantly smaller compared with regulation of rACC for ACCU condition [$t(1,9) = 5.00$, $p < 0.001$], mean %BOLD signal change ± SD (UNR ROI: 0.05 ± 0.99 and rACC: 0.12 ± 2.35) for ACCU condition.

For INSD, regulation of the UNR ROI was significantly smaller compared with regulation of pInsL [$t(1,9) = -3.29$, $p < 0.01$], mean %BOLD signal change ± SD (UNR ROI: 0.05 ± 0.07 and pInsL: 0.26 ± 0.55).

For INSU, regulation of the UNR ROI was significantly smaller compared with regulation of pInsL [$t(1,9) = -5.32$, $p < 0.001$], mean %BOLD signal change ± SD (UNR ROI: 0.07 ± 0.10 and pInsL: 0.14 ± 0.17).

Finally, pain ratings did not differ significantly between Learners and non-Learners with respect to pain intensity and unpleasantness for all conditions [$F(1,8) < 0.52$, $p > 0.50$ for pain intensity and $F(1,8) < 2.28$, $p > 0.17$ for pain unpleasantness], Table 4.

Discussion

We used a connectivity analysis framework in the frequency domain to examine up- or down-regulation of neural activity in the ACC and Ins and pain-associated brain areas during nociceptive processing. We found that successful up- and down-regulation of ACC and Ins is mediated by the ACC and by its functional connectivity with the posterior Insula. These findings are in line with the literature showing that individuals can learn to control activation in the ACC, a region known to be important for both pain perception and pain regulation (4, 12, 13, 36). We extended these findings by showing that we can not only down-regulate but also up-regulate neural activity in ACC and Ins. Such self-regulation aptitude could be related to factors such as pain coping (Emmert et al., 2017) and it would be interesting to investigate how pain coping relates to functional connectivity strength between SII, left and right posterior insula.

We also showed that voluntary control over activation in rACC and posterior Insula was consistently related to a network of three interconnected nodes composed of SII, left and right posterior insula. This network has been shown to be involved in pain processing (6, 37, 38), the nature of connections between the network nodes has, however, not been investigated. We found that the three nodes (SII, left and right posterior insula) were functionally connected, for both the Learner and the non-Learner group, although the strength of the connections differed between groups and conditions.

Interestingly, individuals who can successfully control activation in the Ins showed functional connections between ACC to pInsL (for both INSD and INSU). The non-Learner group showed an additional connection from ACC to the pInsR, albeit weaker. The non-Learner group showed connections between aInsL and SII (for ACCD), or between

TABLE 4 Pain intensity and unpleasantness ratings for all conditions (ACCD, ACCU, INSD, INSU) for learners (1) and non-learners (0) for trial 6.

Subjects	Condition	Learners (1)/non-Learners (0)	Pain intensity	Pain unpleasantness
S1	ACCD	0	5	7
S2	ACCD	0	5	5
S3	ACCD	1	6	6
S4	ACCD	1	5	6
S5	ACCD	1	7	7
S6	ACCD	1	5	5
S7	ACCD	1	5	5
S8	ACCD	0	4.5	7.5
S9	ACCD	0	7.5	7.5
S10	ACCD	1	7	6
S1	ACCU	1	7	10
S2	ACCU	1	5	5
S3	ACCU	0	5	6
S4	ACCU	1	4	4
S5	ACCU	0	6	6
S6	ACCU	1	6	7
S7	ACCU	0	6	5
S8	ACCU	0	2.5	5.5
S9	ACCU	1	5	5
S10	ACCU	0	5	6
S1	INSD	0	5	5
S2	INSD	1	4	4
S3	INSD	1	5	6
S4	INSD	1	5	5
S5	INSD	1	6	6
S6	INSD	1	4	6
S7	INSD	0	5	4
S8	INSD	1	4.5	7.5
S9	INSD	1	7	7
S10	INSD	1	5	4
S1	INSU	1	6	8
S2	INSU	1	5	5
S3	INSU	0	5	6
S4	INSU	1	5	6
S5	INSU	0	6	6
S6	INSU	1	4	5
S7	INSU	1	5	4
S8	INSU	0	3.5	6.5
S9	INSU	0	5	5
S10	INSU	0	9	9

The ratings relate to a verbal rating scale with pain intensity ranging from 0 = no pain to 10 = strongest imaginable pain and pain unpleasantness ranging from 0 = not unpleasant to 10 = extremely unpleasant. ACCD, rACC downregulation; ACCU, rACC upregulation; INSD, pInsL downregulation; INSU, pInsL upregulation.

aInsL and pInsL (for ACCU and INSU). The Learner group did not show connections to aInsL. The anterior and posterior portions of the insula have been shown to be involved in different aspects of pain processing (38), with the posterior portion processing touch and pain sensation (37) and the anterior part involved in affective-motivational processes of pain perception (39). The connections to the aIns in the non-Learner group could therefore suggest that these individuals were negatively affected by pain perception, which might have disrupted their neural regulation task keeping attention on the nociceptive stimuli instead of regulation.

Despite successful control of the activation in ACC (down only) and pInsL (up and down) in the Learner group, pain intensity and unpleasantness ratings did not significantly differ between Learners and non-Learners. A lack of relationship between the regulation of brain activity and changes in behavior or cognition has also previously been reported (2, 17). A possible reason could be that the targets to be regulated, for example, ACC and pIns are involved in a variety of cognitive and behavioural functions such as emotional processing or somatosensory integration (17, 40–42) and may therefore not yield obvious specific behavioral or cognitive changes (43) and, in addition, pain as complex sensory and emotional experience may be more difficult to target by focusing of individual brain regions, as it involves many brain circuits (7).

Limitations

This study involved four distinct conditions (ACCD, ACCU, INSD, INSU), over the course of 4 days. Switching between up- and down- regulation and between two different regions might have been challenging. This may be an additional reason why there were no significant differences in pain intensity and pain unpleasantness ratings between Learners and non-Learners.

Previous neurofeedback studies modulated only one target (4, 44) and used several training trials (up to four) on only 1 day. In this study, we increased the amount of training trials to six, giving subjects more time to train since two brain regions had to be upregulated and downregulated and the tasks were counterbalanced over the 4 days thus switching both the order of regions and the order of upregulation and downregulation. The use of separate training trials for the directions enabled to compare the course of controllability, as well as to monitor systematic changes that might occur independently from regulation efforts such as an overall decrease (habituation) or increase (sensitization) in the response of the area. Furthermore, we previously showed that the insula could be up and down regulated but ACC could only be down regulated, suggesting that both targets can be regulated successfully (apart from ACC upregulation).

Moreover, we showed that success in the modulation of one region and direction of the modulation was not significantly correlated with success in another condition, indicating that regulation of one region might not interfere with regulation of another brain target (14).

Conclusion

Real-time fMRI (rt-fMRI) enables self-regulation of neural activity in localised brain regions through neurofeedback. Previous studies showed successful up- and down-regulation of neural activity in the anterior cingulate cortex (ACC) and the insula (Ins) during nociceptive stimulation. In this work, the brain connectivity analysis was used to investigate how specific brain areas interact to enable successful regulation of nociceptive processing. A connectivity analysis framework in the frequency domain was used to identify a network of interconnected ROIs underlying regulation of neural activity during nociceptive processing. Both the analysis at group level and for single subjects showed that ACC is a key node for a successful control over somatosensory and pain-related areas, and pain regulation underlies an up-down control of ACC. Further work is needed to determine causal influences between somatosensory and pain-related areas during neural regulation of ACC and Ins.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the Medical Faculty Mannheim, Heidelberg University, Germany. The patients/participants provided their written informed consent to participate in this study.

Author contributions

The authors contributed equally to this work in conceptualization, methodology, software and writing. All authors contributed to the article and approved the submitted version.

Funding

This work was partially supported by the Ateneo Italo-Tedesco in the framework of the Programma Vigoni 2010. This study was also supported by grants from the Deutsche Forschungsgemeinschaft (SFB1158/B07 to HF and JA and SFB1158/B03 to FN and HF) and by Project 4-FRAILTY, PNR 2015-2020-Area dispecializzazione: tecnologie per gli ambienti di vita.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Weiskopf N. Real-time fMRI and its application to neurofeedback. *Neuroimage*. (2012) 62(2):682–92. doi: 10.1016/j.neuroimage.2011.10.009

2. Weiskopf N, Veit R, Erb M, Mathiak K, Grodd W, Goebel R, et al. Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. *Neuroimage*. (2003) 19(3):577–86. doi: 10.1016/S1053-8119(03)00145-9

3. Birbaumer N, Ruiz S, Sitaram R. Learned regulation of brain metabolism. *Trends Cogn Sci*. (2013) 17(6):295–302. doi: 10.1016/j.tics.2013.04.009

4. deCharms RC, Maeda F, Glover GH, Ludlow D, Pauly JM, Soneji D, et al. Control over brain activation and pain learned by using real-time functional MRI. *Proc Natl Acad Sci U S A*. (2005) 102(51):18626–31. doi: 10.1073/pnas.0505210102

5. Emmert K, Kopel R, Sulzer J, Bruhl AB, Berman BD, Linden DEJ, et al. Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: how is brain regulation mediated? *Neuroimage*. (2016) 124(Pt A):806–12. doi: 10.1016/j.neuroimage.2015.09.042

6. Apkarian AV, Bushnell MC, Treede RD, Zubieta JK. Human brain mechanisms of pain perception and regulation in health and disease. *Eur J Pain*. (2005) 9(4):463–84. doi: 10.1016/j.ejpain.2004.11.001

7. Wager TD, Atlas LY, Lindquist MA, Roy M, Woo CW, Kross E. An fMRI-based neurologic signature of physical pain. *N Engl J Med*. (2013) 368 (15):1388–97. doi: 10.1056/NEJMoa1204471

8. Apkarian AV, Sosa Y, Sonty S, Levy RM, Harden RN, Parrish TB, et al. Chronic back pain is associated with decreased prefrontal and thalamic gray

matter density. *J Neurosci*. (2004) 24(46):10410–5. doi: 10.1523/JNEUROSCI.2541-04.2004

9. Baliki MN, Chialvo DR, Geha PY, Levy RM, Harden RN, Parrish TB, et al. Chronic pain and the emotional brain: specific brain activity associated with spontaneous fluctuations of intensity of chronic back pain. *J Neurosci*. (2006) 26 (47):12165–73. doi: 10.1523/JNEUROSCI.3576-06.2006

10. Hika B, Al Khalili Y. *Neuronatomy, prefrontal association cortex*. Treasure Island (FL): StatPearls (2022).

11. Ong WY, Stohler CS, Herr DR. Role of the prefrontal cortex in pain processing. *Mol Neurobiol*. (2019) 56(2):1137–66. doi: 10.1007/s12035-018-1130-9

12. Peyron R, Laurent B, Garcia-Larrea L. Functional imaging of brain responses to pain. A review and meta-analysis (2000). *Neurophysiol Clin*. (2000) 30 (5):263–88. doi: 10.1016/S0987-7053(00)00227-6

13. Rainville P. Brain mechanisms of pain affect and pain modulation. *Curr Opin Neurobiol*. (2002) 12(2):195–204. doi: 10.1016/S0959-4388(02)00313-6

14. Rance M, Ruttorf M, Nees F, Schad LR, Flor H. Real time fMRI feedback of the anterior cingulate and posterior insular cortex in the processing of pain. *Hum Brain Mapp*. (2014) 35(12):5784–98. doi: 10.1002/hbm.22585

15. Rance M, Ruttorf M, Nees F, Schad LR, Flor H. Neurofeedback of the difference in activation of the anterior cingulate cortex and posterior insular cortex: two functionally connected areas in the processing of pain. *Front Behav Neurosci*. (2014) 8:357. doi: 10.3389/fnbeh.2014.00357

16. Chapin H, Bagarinao E, Mackey S. Real-time fMRI applied to pain management. *Neurosci Lett*. (2012) 520(2):174–81. doi: 10.1016/j.neulet.2012.02.076

17. Johnston S, Linden DE, Healy D, Goebel R, Habes I, Boehm SG. Upregulation of emotion areas through neurofeedback with a focus on positive mood. *Cogn Affect Behav Neurosci*. (2011) 11(1):44–51. doi: 10.3758/s13415-010-0010-1

18. Ochsner KN, Bunge SA, Gross JJ, Gabrieli JD. Rethinking feelings: an FMRI study of the cognitive regulation of emotion. *J Cogn Neurosci*. (2002) 14 (8):1215–29. doi: 10.1162/089892902760807212

19. Hinterberger T, Neumann N, Pham M, Kubler A, Grether A, Hofmayer N, et al. A multimodal brain-based feedback and communication system. *Exp Brain Res*. (2004) 154(4):521–6. doi: 10.1007/s00221-003-1690-3

20. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. (2009) 10(3):186–98. doi: 10.1038/nrn2575

21. Graziaa FD, Sapuppoa F, Shannahoff-Khalsab D, Bucoloa M. Network parameters for studying functional connectivity in brain MEG data (2009).

22. Stam CJ. Use of magnetoencephalography (MEG) to study functional brain networks in neurodegenerative disorders. *J Neurol Sci*. (2010) 289(1-2):128–34. doi: 10.1016/j.jns.2009.08.028

23. van den Heuvel MP, Hulshoff Pol HE. Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur Neuropsychopharmacol*. (2010) 20(8):519–34. doi: 10.1016/j.euroneuro.2010.03.008

24. He B, Dai Y, Astolfi L, Babiloni F, Yuan H, Yang L. Econnectome: a MATLAB toolbox for mapping and imaging of brain functional connectivity. *J Neurosci Methods*. (2011) 195(2):261–9. doi: 10.1016/j.jneumeth.2010.11.015

25. Oostenveld R, Fries P, Maris E, Schoffelen JM. Fieldtrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. (2011) 2011:156869. doi: 10.1155/2011/156869

26. Zhou D, Thompson WK, Siegle G. MATLAB toolbox for functional connectivity. *Neuroimage*. (2009) 47(4):1590–607. doi: 10.1016/j.neuroimage.2009.05.089

27. Woo CW, Roy M, Buhle JT, Wager TD. Distinct brain systems mediate the effects of nociceptive input and self-regulation on pain. *PLoS Biol*. (2015) 13(1): e1002036. doi: 10.1371/journal.pbio.1002036

28. Friston KJ, Frith CD, Fletcher P, Liddle PF, Frackowiak RS. Functional topography: multidimensional scaling and functional connectivity in the brain. *Cereb Cortex*. (1996) 6(2):156–64. doi: 10.1093/cercor/6.2.156

29. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*. (2010) 52(3):1059–69. doi: 10.1016/j.neuroimage.2009.10.003

30. Bucolo M, Rance M, Muscarello A, Spampinato A, Ruttorf M, Flor H. *Brain connectivity networks and fmri brain-computer interface*. In: *Proceedings of the 7th international workshop on biosignal interpretation (BSI2012)* (2012). p. 17–20

31. Bucoloa M, Rance M, Muscarello A, Spampinato A, Ruttorf M, Flor H. Functional neuro-imaging and brain networks. *Int J Bioelectromagn*. (2012) 14:100–7. ISSN 1456-7865

32. Frot M, Magnin M, Mauguiere F, Garcia-Larrea L. Human SII and posterior insula differently encode thermal laser stimuli. *Cereb Cortex*. (2007) 17(3):610–20. doi: 10.1093/cercor/bhk007

33. Hofbauer RK, Rainville P, Duncan GH, Bushnell MC. Cortical representation of the sensory dimension of pain. *J Neurophysiol*. (2001) 86 (1):402–11. doi: 10.1152/jn.2001.86.1.402

34. Rainville P, Duncan GH, Price DD, Carrier B, Bushnell MC. Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science*. (1997) 277(5328):968–71. doi: 10.1126/science.277.5328.968

35. Faes L, Pinna GD, Porta A, Maestri R, Nollo G. Surrogate data analysis for assessing the significance of the coherence function. *IEEE Trans Biomed Eng*. (2004) 51(7):1156–66. doi: 10.1109/TBME.2004.827271

36. Petrovic P, Ingvar M. Imaging cognitive modulation of pain processing. *Pain*. (2002) 95(1-2):1–5. doi: 10.1016/S0304-3959(01)00467-5

37. Segerdahl AR, Mezue M, Okell TW, Farrar JT, Tracey I. The dorsal posterior insula subserves a fundamental role in human pain. *Nat Neurosci*. (2015) 18 (4):499–500. doi: 10.1038/nn.3969

38. Wiech K, Jbabdi S, Lin CS, Andersson J, Tracey I. Differential structural and resting state connectivity between insular subdivisions and other pain-related brain regions. *Pain*. (2014) 155(10):2047–55. doi: 10.1016/j.pain.2014.07.009

39. Gu X, Gao Z, Wang X, Liu X, Knight RT, Hof PR, et al. Anterior insular cortex is necessary for empathetic pain perception. *Brain*. (2012) 135(Pt 9):2726–35. doi: 10.1093/brain/aws199

40. Baumgartner U, Iannetti GD, Zambreanu L, Stoeter P, Treede RD, Tracey I. Multiple somatotopic representations of heat and mechanical pain in the operculo-insular cortex: a high-resolution fMRI study. *J Neurophysiol*. (2010) 104(5):2863–72. doi: 10.1152/jn.00253.2010

41. Bush G, Luu P, Posner MI. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci*. (2000) 4(6):215–22. doi: 10.1016/S1364-6613(00)01483-2

42. Li X, Hartwell KJ, Borckardt J, Prisciandaro JJ, Saladin ME, Morgan PS, et al. Volitional reduction of anterior cingulate cortex activity produces decreased cue craving in smoking cessation: a preliminary real-time fMRI study. *Addict Biol*. (2013) 18(4):739–48. doi: 10.1111/j.1369-1600.2012.00449.x

43. McCaig RG, Dixon M, Keramatian K, Liu I, Christoff K. Improved modulation of rostrolateral prefrontal cortex using real-time fMRI training and meta-cognitive awareness. *Neuroimage*. (2011) 55(3):1298–305. doi: 10.1016/j.neuroimage.2010.12.016

44. Caria A, Veit R, Sitaram R, Lotze M, Weiskopf N, Grodd W, et al. Regulation of anterior insular cortex activity using real-time fMRI. *Neuroimage*. (2007) 35 (3):1238–46. doi: 10.1016/j.neuroimage.2007.01.018

frontiers | Frontiers in Pain Research

TYPE Review
PUBLISHED 20 December 2022
DOI 10.3389/fpain.2022.1067103

OPEN ACCESS

EDITED BY
Loren Martin,
University of Toronto, Canada

REVIEWED BY
Hadas Nahman-Averbuch,
Washington University in St. Louis, United States
Maria Adele Giamberardino,
University of Studies G. d'Annunzio Chieti and Pescara, Italy

*CORRESPONDENCE
Louise Kuhlmann
louise.frandsen@rn.dk

SPECIALTY SECTION
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

RECEIVED 11 October 2022
ACCEPTED 06 December 2022
PUBLISHED 20 December 2022

CITATION
Kuhlmann L, Olesen SS and Drewes AM (2022) Assessment of visceral pain with special reference to chronic pancreatitis.
Front. Pain Res. 3:1067103.
doi: 10.3389/fpain.2022.1067103

Assessment of visceral pain with special reference to chronic pancreatitis

Louise Kuhlmann[1,2]*, Søren Schou Olesen[1,2] and Asbjørn Mohr Drewes[1,2]

[1]Mech-Sense, Department of Gastroenterology and Hepatology, Aalborg University Hospital, Aalborg, Denmark, [2]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

A thorough pain assessment is of utmost importance when managing pain in clinical practice as it is the foundation for defining pain in need of treatment, either interventional or pharmacological. Pain characteristics can also guide interventional strategies and help evaluate the effect of treatment. In research settings, standardized pain assessment is crucial to improve comparability across studies and facilitate meta-analysis. Due to the importance of thorough visceral pain assessment, this manuscript describes the key elements of pain evaluation focusing on chronic pancreatitis. Most studies in pain assessment have focused on somatic pain, and although chronic pain often shares characteristics between etiologies, some differences must be addressed when assessing visceral pain. Especially differences between somatic and visceral pain are apparent, where visceral pain is diffuse and difficult to localize, with referred pain aspects and often autonomic symptoms dominating the clinical picture. These aspects need to be incorporated into the pain assessment instrument. The manuscript will discuss the different ways of assessing pain, including unidimensional measurement scales, multidimensional questionnaires, and quantitative sensory testing. The advantages and challenges linked to the different methods will be evaluated.

KEYWORDS
pain, chronic pain, visceral pain, pain characterization, chronic pancreatitis, pain assessment

Introduction

Pain is a frequent symptom in the adult population, with a prevalence of up to 20% (1). It is a common cause of seeking medical advice in primary, secondary, and tertiary health sectors. Chronic pain has significant consequences for the patient's life quality, as it affects not only physical health but also psychological well-being, daily activities, and economic functioning (2–5). Besides this, chronic pain also has enormous direct and indirect, associated societal costs (6). Pain treatment is essential in optimizing patient quality of life and disease-related cost. It can minimize pain-related admission and diminish the need for disability payments by maintaining the ability to work (5, 6).

Somatic and visceral pain have many similarities; however, the differences are also considerable. The transmission of visceral pain sensation varies from somatic pain as the afferent nerves innervating viscera terminate at several spinal levels leading to

diffuse pain perception. In the projection to the spinal cord alongside sympathetic fibers, cross-talk often occurs with resultant autonomic symptoms such as nausea, sweating, early satiety, and diarrhea (7). In the spinal cord, the fibers converge with somatic fibers (8). This may lead to pain referred to somatic and other visceral structures.

Visceral diseases are typically associated with severe and disabling pain. According to the International Association for the Study of Pain (IASP) classification of chronic pain, visceral pain can be either primary (previously labeled functional) or secondary (to organic diseases) (9). Although pain is a hallmark of primary visceral pain, in this chapter, we will focus on organic pain, where the diseases are better characterized and understood, with special reference to chronic pancreatitis (CP). However, the principles mentioned in this article can typically also be used in other types of visceral pain, including irritable bowel syndrome, bladder pain syndrome, and endometriosis.

CP is a progressive fibroinflammatory disease where the dominating symptom is visceral pain (10). Pain affects up to 60%–70% of patients, affecting mental health and quality of life (11). The pathophysiology of pain in CP is multifactorial and often caused by a complex interplay between factors such as pancreatic duct obstruction, inflammation, and pancreatic neuropathy (12, 13).

Patients often describe their pain as a continuous, severe, epigastric pain radiating to the back (12), but pain localization varies between individuals (14). The pain is typically fluctuating over time, some patients have pain-free intervals, and other patients have chronic pain with exacerbations (15).

In clinical studies, pain assessment in CP varies considerably. The Pancreatitis-Quantitative Sensory Testing (P-QST) consortium (16) is currently working on a meta-analysis assessing the effect of endoscopic and surgical pain treatment in CP, and preliminary results show that although pain score improvements are similar in the two groups, there are problems with comparing the treatments, as the pain assessment differs considerably between studies. As such, pain assessment varies from comprehensive pain questionnaires to simply asking the patients how they feel. As pain relief is often the primary endpoint in interventional studies of CP, the greatly varying methods for pain assessment across studies are problematic, and studies addressing different treatments can hardly be compared.

Pain assessment tools in chronic pancreatitis

Pain treatment is a difficult and complicated task as chronic pain patients are very heterogenous due to many different origins of pain, diversity in affected pain mechanisms, many pain-associated risk factors, differences in coping strategies, and differing responses to pharmacological treatments that, again, often are associated with many side effects (17). Due to the heterogenicity of the patient group, treatment should be individualized to fit the patient's pain phenotype, depending on, for instance, pain characteristics and affected pain mechanisms. In this context, pain assessment is essential. A scoring of pain severity is used to evaluate the need for analgesic treatment; pain management strategies can be developed from thorough pain characterization; finally, pain assessment is central in evaluating treatment effects (18).

The subjective nature of pain sensation makes objective estimation of pain intensity impossible (18). Therefore, the gold standard for pain assessment is patients' pain self-reports. The method of the patient report can vary from verbal, unidimensional measurements to written comprehensive multidimensional pain assessment. Pain has many components, including pain intensity, localization, pattern, factors provoking pain, factors exacerbating pain, pain-related symptoms, current treatment (pharmacological), previous treatments (pharmacological as well as interventional), quality of life, mental health, and risk factors for pain. The many aspects of pain underline the need for multidimensional pain assessment.

Unidimensional pain scales such as the Visual Analogue Scale and the Numerical Rating Scale are commonly used in clinic and research practice to assess pain intensity. There are, however, several challenges in using unidimensional scales for pain assessment. The scales are simple measurements, but the interpretation reflects the individual's conceptualization of pain, resulting in significant differences between reports. It has been suggested that the unidimensional scales should be converted to ratio scales to provide information on changes over time rather than a single measurement (19). As the unidimensional scales leave several aspects of pain assessment in the dark, it is likely more suited for assessing acute rather than chronic pain when used as a stand-alone measure (18).

Different recommendations on pain assessment, including the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) and Validation and Application of a patient-relevant core set of outcome domains to assess multimodal PAIN therapy (VAPAIN) recommendations, specify that several core domains in pain should be considered in clinical studies (20, 21). When complying with these recommendations, pain assessment will provide information on pain intensity, characteristics, and how pain affects different aspects of patients' lives, including sleep, economic function, and psychological health. In visceral pain, changes in pain characteristics can be caused by new disease-related complications, where targeted treatments might exist and are therefore mandatory to assess (22). A more in-depth assessment can provide important clinical knowledge that can be used to evaluate the need for further examinations. Multidimensional scales are, therefore, useful in

visceral pain. It can focus on disease-specific characteristics and evaluate further aspects of pain if developed for a specific condition. This gives a complete image of how pain affects the patients' lives. However, a detailed multidimensional pain characterization is time-consuming, limiting the use of comprehensive pain assessment tools in research and clinical practice.

Until recently, there has been a lack of formally validated pain assessment tools developed specifically for CP. The Izbicki pain scale has commonly been used, as it is developed specifically for CP, but it still lacks the formal validation process. It was presented in a study in 1995 and has been used extensively afterward (23). The questionnaire is quite simple, evaluating pain on intensity, frequency of pain attacks, use of analgesic treatment, and inability to work. Each subpart accounts for 25% of the score, but due to the workability assessment, it is limited in its response to treatments over shorter periods. Besides the Izbicki pain scale, other non-chronic-pancreatitis-specific questionnaires, such as the brief pain inventory, have been validated and used for pain assessment. However, these questionnaires lack evaluation of pancreatic pain-specific domains such as postprandial pain and gastrointestinal manifestations (24).

Recently a comprehensive pain assessment questionnaire, the Comprehensive Pain Assessment Tool (COMPAT), has been developed specifically for CP, complying with the IMMPACT and VAPAIN recommendations (14). It is useful for a comprehensive evaluation of pancreatic pain, but due to the extensive length of 17 pages, some patients might not be able to answer the questionnaire sufficiently. Consequently, a short form of the COMPAT questionnaire, the COMPAT-SF, has been developed (25) and validated as a separate questionnaire. The COMPAT-SF scores correlate to scores from the brief pain inventory and the Izbicki pain scale. It also correlates to patient quality of life and hospitalizations due to pain in the previous year. Reliability has been evaluated both on internal consistency and in a test-retest examination. It has been proven acceptable, especially when considering chronic pain's fluctuating nature (25). Predictive validity and the power as a decision-making tool are still lacking but will be examined in future years.

An international guideline for using different pain questionnaires and recommendations for their use in painful CP has recently been published. For further details, the reader is referred to (26).

Neurophysiological assessment of pain in chronic pancreatitis

Questionnaires can however fail to capture the complexity of visceral pain in CP, and research has focused on identifying additional methods for assessing pain and guiding treatment strategies (27). Quantitative Sensory Testing (QST) can be used to assess pain, where it serves to characterize sensory processing in both peripheral and central pain pathways (**Figure 1**). It can serve as a means to phenotype the patient's nociceptive profile. In QST, standardized stimulations of somatic and visceral tissue are used to explore different neural pathways and networks. This results in a response quantified with psychophysical and/or objective methods (28). Visceral stimulations of patients are often not well accepted in a clinical setting, and due to convergence between visceral afferents from the pancreas and somatic afferents from the T10 dermatome, QST of the skin can be used to assess whether pain processing from the pancreas to the central nervous system is sensitized (29). In addition, when adding more specific examinations, such as assessment of endogenous descending inhibition from centers in the brainstem and temporal summation, we can analyze whether pain processing in the central pathways is abnormal. QST can be used as a biomarker to categorize pain phenotypes based on affected pain mechanisms (30).

Quantitative sensory testing involves several tests to enlighten the sensory function of pain perception, from peripheral stimulation to evaluation of the processing in the brain. The tests consist of different standardized stimulations and evaluate patients' subjective pain intensity response. The stimulation can, for instance, be thermal, mechanical, electrical, chemical, or ischemic (31). The resulting pain intensity registration can be supplied with electroencephalography (EEG) to examine an objective measure. Resting state EEG can be used to examine the brain's default mode, whereas EEG during a painful stimulation in QST gives us information on evoked brain potentials as a result of the pain perception (32). Autonomic reactivity to nociceptive input can also be measured objectively, where heart rate variability is a promising measure (33). It changes due to increased sympathetic-baroreflex activity and a decrease in vagal-parasympathetic activity.

QST, EEG, and functional magnetic resonance imaging have been used in several clinical studies and can be used to identify different dysfunctional pain mechanisms. However, pain itself is a subjective sensation and should be assessed as such (34–36). The German Research Network on Neuropathic Pain has published guidelines on comprehensive QST batteries to examine pain thoroughly (37). Although this gives a detailed description of pain processing, it is unsuitable for clinical practice and cannot be used as a bedside tool to examine visceral pain. Invasive visceral stimulations are also used to examine pain processing in visceral pain (34, 38, 39). These examinations may involve significant discomfort for the patients, fragile and expensive equipment and are therefore not optimal for bedside use in clinical routine work. However, as visceral pain processing can be partly reflected in convergent somatic dermatomes, this can be used as a proxy

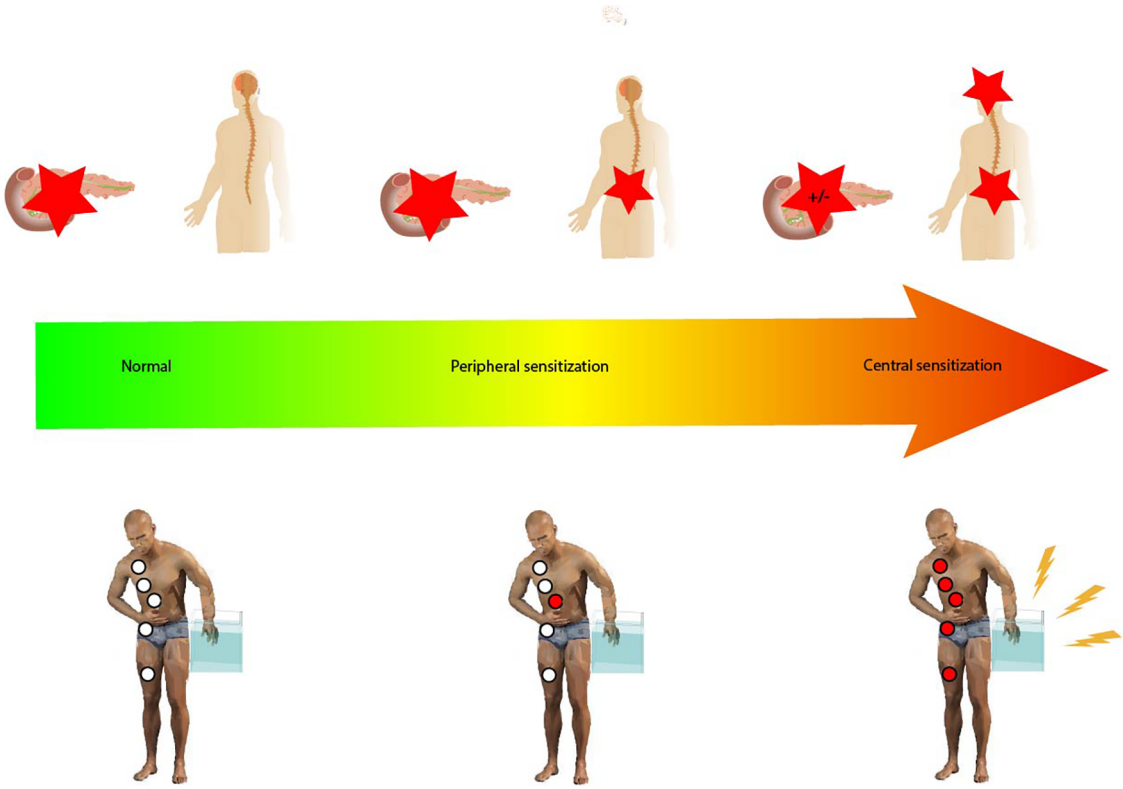


FIGURE 1
Pain responses ranging from normal through peripheral sensitization to central sensitization. Peripheral sensitization is characterized by increased excitability of the second-order neurons sharing spinal innervation with the pancreatic gland but lacking changes in the central pain processing pathways. Central sensitization is associated with changes in central pain processing pathways. The red stars indicate the generation of a nociceptive response to stimulation, either peripherally or centrally. Hypersensitivity as tested by the P-QST study protocol (pain thresholds as well as temporal summation) is indicated by the red circles; white circles indicate normal P-QST responses. In central sensitization, dynamic QST measures such as conditioned pain modulation can also be affected as shown in the figure.

of central aspects of visceral pain (40). The P-QST consortium was formed to develop and promote knowledge of a bedside QST examination that could be used to evaluate pain processing (16). It consists of three simple tests. One examines sensitization in several anatomical regions using pain detection and tolerance thresholds to standardized pressure (27, 41). If the nervous structures in anatomical regions associated with the pancreas are evaluated as sensitized, and other regions have normal thresholds, it could indicate segmental hyperalgesia corresponding to the pancreatic segment of the spinal cord. If pain thresholds are affected at several sites, it indicates generalized hyperalgesia where the damage is not only located in the peripheral nerve and associated spinal segments but also changes in the central processing of painful stimulations (42).

Another test in P-QST examines the temporal summation score. First, a single pinprick stimulation is performed on a somatic structure, such as the dominant forearm, the pancreatic dermatome (T10), or both, and the patient's corresponding pain intensity is reported. Afterward, a repetitive series of pinpricks are performed one second apart, and the resulting pain intensity is reported. The ratio between the two pain intensities is calculated as the temporal summation score (43). The final P-QST examination evaluates conditioned pain modulation, the change in pain perception after a conditioning stimulus. Conditioned pain modulation has previously been shown to correlate to clinical pain intensity and is an important aspect to include in the neurophysiological examination (41). In the P-QST consortium, the cold pressor test is recommended where a conditioning stimulus, lowering the patient's hand in water with a temperature of 2 degrees for 2 min, is performed. The test stimulus, pressure stimulation on the thigh until the pain tolerance level is reached, is performed before and immediately after the conditioning stimulus and the two values are compared (28).

The P-QST consortium has proposed an algorithm for identifying central and segmental sensitization, where at least 2 out of the following 4 measures (conditioned pain modulation, cold pressor endurance time, sum of pain

detection thresholds, and temporal summation on the forearm) indicate central sensitization, whereas 1 out of the following 2 measures (1: Ratio between pain thresholds in T10 and control sites and 2: Enhanced temporal summation at the upper abdominal) indicate segmental sensitization (27).

Over the last decades of pain research, it has been debated whether QST results are gender-specific, as evidence has pointed to a difference in pain perception and pain frequency between males and females (44, 45). A systematic review by Racine et al. from 2012 did however not find results proving lower pain tolerance in women, as there was significant differences in results when comparing different test stimuli and stimulation sites (46). Research does point to variations due to the menstrual cycle, and QST measures should optimally be standardized in regards to menstrual phase (47). This would however be difficult to plan in clinical practice.

In the P-QST consortium, the reference values for pressure pain detection threshold has been differentiated between sexes, as studies show variation in this exact test stimulus (27, 46).

The reliability of QST has been discussed. Static measurements, such as pressure pain stimulation, have acceptable reproducibility, whereas dynamic parameters, especially conditioned pain modulation, show variability over time (28). However, the variability differs between stimulation methods, and there are several ways to improve reliability. This includes, among other, comprehensive training of both examinators and patients and choosing the most reproducible painful stimulation; if this is considered, the assessment method is still a potent prognostic factor in clinical studies (48, 49). The test stimulus in the conditioned pain modulation regimen was chosen due to its reproducibility as a static examination. It is the most commonly used test stimulus and is well tolerated. Although reliability varies according to test site, both inter- and intrasession reliability are generally good (43, 48). The conditioning stimulus can be of different types, as well as different intensities. It can be discussed whether the stimulus can be too painful for certain individuals, and thereby possibly excluding them from completing the stimulus. On the contrary, the stimulus also can be too mild to evaluate pain modulation. Generally, there is low reproducibility for dynamic QST (the conditioning stimulus) due to the complex mechanisms of pain modulation, and the results must be evaluated with this in mind (43). It is however accessible and easy to control, and although some patients might not endure the full conditioning stimulation, this is also a usable result in the final evaluation, as described above.

In recent studies, QST has been used to predict the outcomes of treatments. Olesen et al. have shown that hyperalgesia to electrical stimulations in the T10 dermatome is predictive of the efficacy of pregabalin treatment in CP (50). QST has also been used as a predictor in other types of patients, including diabetic neuropathy, where conditioned pain modulation predicts the efficacy of duloxetine (51), and a mixed group of chronic pain patients, where cold pain intensity and EEG activity induced by cold pain, predicted the pain reduction of opioid treatment (52). Further studies are, however, needed before using this as a decision-making tool in clinical practice, although different pain treatment algorithms using neurophysiological evaluations have been proposed (29, 53).

Conclusion

Pain assessment in visceral pain is complex. As pain affects life in many ways, several aspects besides pain intensity must be evaluated. These include mental health, autonomic symptoms, and quality of life. Unidimensional scales are mostly suited for evaluating changes in pain intensity in acute pain but are too simple for assessing the complexity of chronic visceral pain. Questionnaire validity is increased when developed for a specific disease, as it can provide information beyond general characteristics.

Besides pain questionnaires, QST is gaining ground. It is used to quantify the loss or gain of sensory function and can be performed as a quick bedside examination with only a few instruments available. It can help evaluate the progression of chronic pain from a segmental to a central origin. In the future, it might also help tailor analgesic treatment focusing on affected pain mechanisms (29).

It's important to strive for a uniform assessment of pain in clinical studies, as this will increase the comparability of results. For further information on pain assessment in CP, please see the international guidelines on the subject (26).

Author contributions

LK has drafted the initial version of the manuscript, which has been commented on and edited by SO and AD. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Reid KJ, Harker J, Bala MM, Truyers C, Kellen E, Bekkering GE, et al. Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. *Curr Med Res Opin*. (2011) 27 (2):449–62. doi: 10.1185/03007995.2010.545813

2. Magni G, Marchetti M, Moreschi C, Merskey H, Luchini SR. Chronic musculoskeletal pain and depressive symptoms in the national health and nutrition examination I. Epidemiologic follow-up study. *Pain*. (1992) 53:163–8. doi: 10.1016/0304-3959(93)90076-2

3. Viane I, Crombez G, Eccleston C, Devulder J, De Corte W. Acceptance of the unpleasant reality of chronic pain: effects upon attention to pain and engagement with daily activities. *Pain*. (2004) 112(3):282–8. doi: 10.1016/j.pain.2004.09.008

4. Teasell RW, Bombardier C. Employment-related factors in chronic pain and chronic pain disability. *Clin J Pain*. (2001) 17(4 Suppl):839–45. doi: 10.1097/00002508-200112001-00010

5. Bonathan C, Hearn L, Williams AC. Socioeconomic status and the course and consequences of chronic pain. *Pain Manag*. (2013) 3(3):159–62. doi: 10.2217/pmt.13.18

6. Turk DC, Wilson HD, Cahana A. Treatment of chronic non-cancer pain. *Lancet*. (2011) 377(9784):2226–35. doi: 10.1016/S0140-6736(11)60402-9

7. Drewes AM, Olesen AE, Farmer AD, Szigethy E, Rebours V, Olesen SS. Gastrointestinal pain. *Nat Rev Dis Prim*. (2020) 6(1):1–16. doi: 10.1038/s41572-019-0135-7

8. Gebhart GF, Bielefeldt K. Physiology of visceral pain. *Compr Physiol*. (2016) 6 (4):1609–33. doi: 10.1002/cphy.c150049

9. Aziz Q, Giamberardino MA, Barke A, Korwisi B, Baranowski AP, Wesselmann U, et al. The IASP classification of chronic pain for ICD-11: chronic secondary visceral pain. *Pain*. (2019) 160(1):69–76. doi: 10.1097/j.pain.0000000000001362

10. Majumder S, Chari ST. Chronic pancreatitis. *Lancet*. (2016) 387 (10031):1957–66. doi: 10.1016/S0140-6736(16)00097-0

11. Olesen SS, Poulsen JL, Drewes AM, Frøkjær JB, Laukkarinen J, Parhiala M, et al. The scandinavian baltic pancreatic club (SBPC) database: design, rationale and characterisation of the study cohort. *Scand J Gastroenterol*. (2017) 52 (8):909–15. doi: 10.1080/00365521.2017.1322138

12. Poulsen JL, Olesen SS, Malver LP, Frøkjær JB, Drewes AM. Pain and chronic pancreatitis: a complex interplay of multiple mechanisms. *World J Gastroenterol*. (2013) 19(42):7282–91. doi: 10.3748/wjg.v19.i42.7282

13. Drewes AM, Krarup AL, Detlefsen S, Malmstrøm M-L, Dimcevski G, Funch-Jensen P. Pain in chronic pancreatitis: the role of neuropathic pain mechanisms. *Gut*. (2008) 57(11):1616–27. doi: 10.1136/gut.2007.146621

14. Teo K, Johnson MH, Drewes AM, Windsor JA. A comprehensive pain assessment tool (COMPAT) for chronic pancreatitis: development, face validation and pilot evaluation. *Pancreatology*. (2017) 17(5):706–19. doi: 10.1016/j.pan.2017.07.004

15. Mullady DK, Yadav D, Amann ST, O'Connell MR, Barmada MM, Elta GH, et al. Type of pain, pain-associated complications, quality of life, disability and resource utilisation in chronic pancreatitis: a prospective cohort study. *Gut*. (2011) 60(1):77–84. doi: 10.1136/gut.2010.213835

16. Phillips AE, Faghih M, Singh VK, Olesen SS, Kuhlmann L, Novovic S, et al. Rationale for and development of the pancreatic quantitative sensory testing consortium to study pain in chronic pancreatitis. *Pancreas*. (2021) 50 (9):1298–304. doi: 10.1097/MPA.0000000000001912

17. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: impact on daily life, and treatment. *Eur J Pain*. (2006) 10:287–333. doi: 10.1016/j.ejpain.2005.06.009

18. Breivik H, Borchgrevink PC, Allen SM, Rosseland LA, Romundstad L, Breivik Hals EK, et al. Assessment of pain. *Br J Anaesth*. (2008) 101:17–24. doi: 10.1093/bja/aen103

19. Hartrick CT, Kovan JP, Shapiro S. The numeric rating scale for clinical pain measurement: a ratio measure? *Pain Pract*. (2003) 3(4):310–6. doi: 10.1111/j.1530-7085.2003.03034.x

20. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: iMMPACT recommendations. *Pain*. (2003) 106(3):337–45. doi: 10.1016/j.pain.2003.08.001

21. Kaiser U, Kopkow C, Deckert S, Neustadt K, Jacobi L, Cameron P, et al. Developing a core outcome domain set to assessing effectiveness of interdisciplinary multimodal pain therapy: the VAPAIN consensus statement on core outcome domains. *Pain*. (2018) 159(4):673–83. doi: 10.1097/j.pain.0000000000001129

22. D'Haese JG, Ceyhan GO, Demir IE, Tieftrunk E, Friess H. Treatment options in painful chronic pancreatitis: a systematic review. *HPB*. (2014) 16 (6):512–21. doi: 10.1111/hpb.12173

23. Bloechle C, Izbicki JR, Knoefel WT, Kuechler T, Broelsch CE. Quality of life in chronic pancreatitis, results after duodenum-preserving resection of the head of the pancreas. *Pancreas*. (1995) 11(1):77–85. doi: 10.1097/00006676-199507000-00008

24. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the brief pain inventory for chronic nonmalignant pain. *J Pain*. (2004) 5(2):133–7. doi: 10.1016/j.jpain.2003.12.005

25. Kuhlmann L, Teo K, Olesen SS, Phillips AE, Faghih M, Tuck N, et al. Development of the comprehensive pain assessment tool short form for chronic pancreatitis: validity and reliability testing. *Clin Gastroenterol Hepatol*. (2022) 20 (4):e770–83. doi: 10.1016/j.cgh.2021.05.055

26. Drewes AM, Bellin MD, Besselink MG, Bouwense SA, Olesen SS, van Santvoort H, et al. Assessment of pain associated with chronic pancreatitis: an international consensus guideline. *Pancreatology*. (2021) 21(7):1256–84. doi: 10.1016/j.pan.2021.07.004

27. Phillips AE, Faghih M, Kuhlmann L, Larsen IM, Drewes AM, Singh VK, et al. A clinically feasible method for the assessment and characterization of pain in patients with chronic pancreatitis. *Pancreatology*. (2020) 20(1):25–34. doi: 10.1016/j.pan.2019.11.007

28. Olesen SS, van Goor H, Bouwense SAW, Wilder-Smith OHG, Drewes AM. Reliability of static and dynamic quantitative sensory testing in patients with painful chronic pancreatitis. *Reg Anesth Pain Med*. (2012) 37(5):530–6. doi: 10.1097/AAP.0b013e3182632c40

29. Kuhlmann L, Olesen SS, Olesen AE, Arendt-Nielsen L, Drewes AM. Mechanism-based pain management in chronic pancreatitis – is it time for a paradigm shift? *Expert Rev Clin Pharmacol*. (2019) 12(3):1–10. doi: 10.1080/17512433.2019.1571409

30. Smith SM, Dworkin RH, Turk DC, Baron R, Polydefkis M, Tracey I, et al. The potential role of sensory testing, skin biopsy, and functional brain imaging as biomarkers in chronic pain clinical trials: iMMPACT considerations. *J Pain*. (2017) 18(7):757–77. doi: 10.1016/j.jpain.2017.02.429

31. Cruz-Almeida Y, Fillingim R. Can quantitative sensory testing move us closer to mechanism-based pain management. *Pain Med*. (2012) 100(2):130–4. doi: 10.1111/pme.12230

32. Bouwense SAW, de Vries M, Schreuder LTW, Olesen SS, Frøkjær JB, Drewes AM, et al. Systematic mechanism-orientated approach to chronic pancreatitis pain. *World J Gastroenterol*. (2015) 21(1):47–59. doi: 10.3748/wjg.v21.i1.47

33. Koenig J, Jarczok MN, Ellis RJ, Hillecke TK, Thayer JF. Heart rate variability and experimentally induced pain in healthy adults: a systematic review. *Eur J Pain*. (2014) 18(3):301–14. doi: 10.1002/j.1532-2149.2013.00379.x

34. Arendt-Nielsen L, Yarnitsky D. Experimental and clinical applications of quantitative sensory testing applied to skin, muscles and viscera. *J Pain*. (2009) 10(6):556–72. doi: 10.1016/j.jpain.2009.02.002

35. Olesen SS, Hansen TM, Graversen C, Steimle K, Wilder-Smith OHG, Drewes AM. Slowed EEG rhythmicity in patients with chronic pancreatitis: evidence of abnormal cerebral pain processing? *Eur J Gastroenterol Hepatol*. (2011) 23(5):418–24. doi: 10.1097/MEG.0b013e3283457b09

36. Olesen SS, Hansen TM, Graversen C, Valeriani M, Drewes AM. Cerebral excitability is abnormal in patients with painful chronic pancreatitis. *Eur J Pain*. (2013) 17(1):46–54. doi: 10.1002/j.1532-2149.2012.00155.x

37. Maier C, Baron R, Tölle TR, Binder A, Birbaumer N, Birklein F, et al. Quantitative sensory testing in the German research network on neuropathic pain (DFNS): somatosensory abnormalities in 1236 patients with different neuropathic pain syndromes. *Pain*. (2010) 150(3):439–50. doi: 10.1016/j.pain.2010.05.002

38. Dimcevski G, Staahl C, Andersen SD, Thorsgaard N, Funch-Jensen P, Arendt-Nielsen L, et al. Assessment of experimental pain from skin, muscle, and esophagus in patients with chronic pancreatitis. *Pancreas*. (2007) 35 (1):22–9. doi: 10.1097/mpa.0b013e31805c1762

39. Dimcevski G, Schipper KP, Tage-Jensen U, Funch-Jensen P, Krarup AL, Toft E, et al. Hypoalgesia to experimental visceral and somatic stimulation in painful chronic pancreatitis. *Eur J Gastroenterol Hepatol*. (2006) 18(7):755–64. doi: 10.1097/01.meg.0000223903.70492.c5

40. Arendt-Nielsen L, Laursen RJ, Drewes AM. Referred pain as an indicator for neural plasticity. *Prog Brain Res*. (2000) 129:343–56. doi: 10.1016/S0079-6123(00)29026-2

41. Kuhlmann L, Olesen SS, Grønlund D, Olesen AE, Phillips AE, Faghih M, et al. Patient and disease characteristics associate with sensory testing results in chronic pancreatitis. *Clin J Pain*. (2019) 35(9):786–93. doi: 10.1097/AJP.0000000000000740

42. Arendt-Nielsen L, Morlion B, Perrot S, Dahan A, Dickenson A, Kress HGG, et al. Assessment and manifestation of central sensitisation across different chronic pain conditions. *Eur J Pain*. (2018) 22(2):216–41. doi: 10.1002/ejp.1140

43. Marcuzzi A, Wrigley PJ, Dean CM, Adams R, Hush JM. The long-term reliability of static and dynamic quantitative sensory testing in healthy individuals. *Pain*. (2017) 158(7):1217–23. doi: 10.1097/j.pain.0000000000000901

44. Kindler LL, Valencia C, Fillingim RB, George SZ. Sex differences in experimental and clinical pain sensitivity for patients with shoulder pain. *Eur J Pain*. (2011) 15(2):118–23. doi: 10.1016/j.ejpain.2010.06.001.

45. Bulls HW, Freeman EL, Anderson AJ, Robbins MT, Ness TJ, Goodin BR. Sex differences in experimental measures of pain sensitivity and endogenous pain inhibition. *J Pain Res*. (2015) 8:311–20. doi: 10.2147/jpr.s84607

46. Racine M, Tousignant-Laflamme Y, Kloda LA, Dion D, Dupuis G, Choinière M. A systematic literature review of 10 years of research on sex/gender and experimental pain perception-part 1: are there really differences between women and men? *Pain*. (2012) 153:602–18. doi: 10.1016/j.pain.2011.11.025

47. Riley III JL, Robinson ME, Wise EA, Price DD. A meta-analytic review of pain perception across the menstrual cycle. *Pain*. (1999) 81:225–35. doi: 10.1016/S0304-3959(98)00258-9

48. Kennedy DL, Kemp HI, Ridout D, Yarnitsky D, Rice ASC. Reliability of conditioned pain modulation: a systematic review. *Pain*. (2016) 157(11):2410–9. doi: 10.1097/j.pain.0000000000000689

49. O'Neill S, O'Neill L. Improving QST reliability - more raters, tests, or occasions? A multivariate generalizability study. *J Pain*. (2015) 16(5):454–62. doi: 10.1016/j.jpain.2015.01.476

50. Olesen SS, Graversen C, Bouwense SAW, van Goor H, Wilder-Smith OHG, Drewes AM. Quantitative sensory testing predicts pregabalin efficacy in painful chronic pancreatitis. *PLoS One*. (2013) 8(3):e57963. doi: 10.1371/journal.pone.0057963

51. Yarnitsky D, Granot M, Nahman-Averbuch H, Khamaisi M, Granovsky Y. Conditioned pain modulation predicts duloxetine efficacy in painful diabetic neuropathy. *Pain*. (2012) 153(6):1193–8. doi: 10.1016/j.pain.2012.02.021

52. Grosen K, Olesen AE, Gram M, Jonsson T, Kamp-Jensen M, Andresen T, et al. Predictors of opioid efficacy in patients with chronic pain: a prospective multicenter observational cohort study. *PLoS One*. (2017) 12(2):1–13. doi: 10.1371/journal.pone.0171723

53. Drewes AM, Bouwense SAW, Campbell CM, Ceyhan GO, Delhaye M, Demir IE, et al. Guidelines for the understanding and management of pain in chronic pancreatitis. *Pancreatology*. (2017) 17(5):720–31. doi: 10.1016/j.pan.2017.07.006

frontiers | Frontiers in Pain Research

TYPE Original Research
PUBLISHED 15 February 2023
DOI 10.3389/fpain.2023.1057659

OPEN ACCESS

EDITED BY
Kai Karos,
Open University of the Netherlands, Netherlands

REVIEWED BY
Wiebke Gandhi,
University of Reading, United Kingdom
Angela Starkweather,
University of Connecticut, United States

*CORRESPONDENCE
Karen Deborah Davis
karen.davis@uhnresearch.ca

SPECIALTY SECTION
This article was submitted to Pain Research Methods, a section of the journal Frontiers in Pain Research

RECEIVED 29 September 2022
ACCEPTED 17 January 2023
PUBLISHED 15 February 2023

CITATION
Sanmugananthan VV, Cheng JC, Hemington KS, Rogachov A, Osborne NR, Bosma RL, Kim JA, Inman RD and Davis KD (2023) Can we characterize A-P/IAP behavioural phenotypes in people with chronic pain?
Front. Pain Res. 4:1057659.
doi: 10.3389/fpain.2023.1057659

Can we characterize A-P/IAP behavioural phenotypes in people with chronic pain?

Vaidhehi Veena Sanmugananthan[1,2], Joshua C. Cheng[1,2], Kasey S. Hemington[1,2], Anton Rogachov[1,2], Natalie Rae Osborne[1,2], Rachael L. Bosma[1,2], Junseok Andrew Kim[1,2], Robert D. Inman[1,3] and Karen Deborah Davis[1,2,4]*

[1]Division of Brain, Imaging and Behaviour, Krembil Brain Institute, Krembil Research Institute, University Health Network, Toronto, ON, Canada, [2]Institute of Medical Science, University of Toronto, Toronto, ON, Canada, [3]Department of Medicine, University of Toronto, Toronto, ON, Canada, [4]Department of Surgery, University of Toronto, Toronto, ON, Canada

Two behavioural phenotypes in healthy people have been delineated based on their intrinsic attention to pain (IAP) and whether their reaction times (RT) during a cognitively-demanding task are slower (*P*-type) or faster (A-type) during experimental pain. These behavioural phenotypes were not previously studied in chronic pain populations to avoid using experimental pain in a chronic pain context. Since pain rumination (PR) may serve as a supplement to IAP without needing noxious stimuli, we attempted to delineate A-P/IAP behavioural phenotypes in people with chronic pain and determined if PR can supplement IAP. Behavioural data acquired in 43 healthy controls (HCs) and 43 age-/sex-matched people with chronic pain associated with ankylosing spondylitis (AS) was retrospectively analyzed. A-P behavioural phenotypes were based on RT differences between pain and no-pain trials of a numeric interference task. IAP was quantified based on scores representing reported attention towards or mind-wandering away from experimental pain. PR was quantified using the pain catastrophizing scale, rumination subscale. The variability in RT was higher during no-pain trials in the AS group than HCs but was not significantly different in pain trials. There were no group differences in task RTs in no-pain and pain trials, IAP or PR scores. IAP and PR scores were marginally significantly positively correlated in the AS group. RT differences and variability were not significantly correlated with IAP or PR scores. Thus, we propose that experimental pain in the A-P/IAP protocols can confound testing in chronic pain populations, but that PR could be a supplement to IAP to quantify attention to pain.

KEYWORDS

behavioural phenotype, chronic pain, attention, reaction time, rumination

1. Introduction

Pain is an attention-grabbing sensory experience, and acute pain plays a role in protecting our bodies from harm (1, 2). Pain and attention are interconnected. For example, allocation of attention can modulate components of the pain experience (e.g., pain intensity) and activation of pain-related brain regions (3–11). Much of the research that has examined pain and attention interactions has been conducted in healthy individuals, but studies in chronic pain populations are challenging and confounded by the need to apply an experimental pain stimulus in the context of chronic pain.

Our lab has demonstrated that healthy individuals vary in their ability to balance attention towards pain vs. other attentional demands, and that these differences generally can be observed

by two behavioural phenotypes. One behavioural phenotype is based on how performance of an attention-demanding task is impacted by concurrent experimental acute pain stimuli. We have designated individuals as either "Attention to task dominants" (A-types) for those that exhibit faster task reaction times (RTs) or "pain dominants" (P-types) for those that exhibit slower task RTs during concurrent pain stimulation compared to a no-pain condition (10, 12–14). The other behavioural phenotype is based on an assessment we developed to determine intrinsic attention to pain (IAP), which reflects the tendency of an individual to attend to an acute pain stimulus vs. mindwander away from the pain stimulus (1, 15). Healthy individuals designated into these two behavioural phenotypes have been shown to have characteristic structural and functional attributes in areas of the dynamic pain connectome in the brain (1, 10, 12–15).

Behavioural phenotypes that identify how an individual attends to pain could provide some individual predictive value of the potential effectiveness of attentional or psychotherapeutic interventions that aim to alter attentional engagement towards chronic pain (e.g., cognitive behavioural therapy) (16, 17). However, A-P types and IAP have not been characterized in chronic pain populations in part because of potential confounds associated with applying experimental pain stimuli in a chronic pain context and the unknown relevancy of using experimental pain as a model for chronic pain (18–20).

In this exploratory study, our main aim was to examine A-P and IAP behavioral phenotypes in people with chronic pain using our standard A-P/IAP protocols that use acute experimental pain. However, because of the inherent confound of applying experimental pain in a chronic pain context, our secondary exploratory aim was to explore whether pain rumination (PR) could be used to quantify attention to pain and provide an experimental pain-free supplement to IAP.

PR is repetitive and continuous negative thinking about pain and the possible causes and consequences associated with it's experience (21). PR is thought to be related to IAP (1, 21) because it also captures an individual's tendency to attend to pain. However, the relationship between these two metrics has not been examined previously.

2. Materials and methods

2.1. Participants

This study comprised a retrospective analysis of behavioural data collected from 43 right—handed people with chronic pain associated with ankylosing spondylitis (AS) (average age = 28.7, SD = +/−6.4 years old; 30 males, 13 females) and 43 age- and sex- matched pain-free healthy controls (HC) (average age = 28.2, SD = +/−6.1 years old; 30 males, 13 females) with the overall ages ranged from 18 to 40 (+/−2) years old.

All study participants provide informed consent to experimental methods that were approved by the University Health Network Research Ethics Board. We recruited individuals with AS from the Toronto Western Hospital's Spondylitis Clinic and that were diagnosed with AS using the modified New York criteria (22, 23). Both HCs and AS participants were excluded if they met any of the following conditions: (1) current or were previously diagnosed with a psychiatric, neurological, or metabolic disorder, (2) previous major surgeries, (3) any serious infection within 4 weeks of data collection requiring hospitalization and/or antibiotics.

2.2. Tasks

2.2.1. Numeric interference task

Participants were familiarized with a numeric interference (NI) task (see 12, 13, 24–27) and underwent a training session before testing began. The NI task required the participants to view a computer screen that displayed 3 separate boxes, each of which contained a different number of digits that ranged in value from 1 to 9. Within each box there were identical numbers but there were different numbers across the boxes. Each participant was instructed to use a numerical keyboard to indicate as quickly and as accurately as possible, the highest number of digits across the boxes. The cognitive-demanding aspect of the task was that participants had to report the highest number of digits (non-dominant information) rather than the highest number value (dominant information) (12, 13, 24–27). The study included 6 blocks with 24 trials each (trial length = 2.5 s, inter-block interval = 60 s), and blocks alternated between a no-pain condition and a pain condition during which experimental pain was applied concurrently during the task (12, 13). A computer-controlled transcutaneous electrical nerve stimulation (TENS) device (300-PV Empi Inc.) was used to deliver stimuli to the left median nerve and was calibrated prior to testing to elicit pain intensity of approximately 40–60/100 (0 = no pain, 100 = most intense pain imaginable) for each participant. The NI task was run on EPrime v1.1 (Psychological Software tools). See the **Supplementary Materials** for more details about the TENS stimulus calibration procedure. The first two blocks of the NI task (one no-pain block and one pain block) were removed to avoid learning effects for each participant.

2.2.2. Measuring performance on the NI task and data-cleaning

Task performance was quantified from each participants' mean RT and RT variability (RTv) (12) across the no-pain and pain blocks, respectively (see our previous study 12). The RTv in each participant was calculated from the variance of the RTs in all of the trials of the no-pain blocks and the pain blocks separately.

Trials with RTs that were <=200 milliseconds (ms) or >=2500 ms were removed. The upper cut-off was determined based on the maximum trial time. The lower cut- off was based on the postulated time needed for physiological processes (e.g., stimulus detection, decision making, motor response) to occur (approximately 100–200 ms) during a reaction-time (28, 29). Participant exclusion criteria was set at having more than 30% of their total trials missing from each block-type and/or all blocks together after data-cleaning was completed. No participants were excluded from analyses after the data-cleaning procedure was implemented.

2.2.3. A-P categorization of individuals

The differences in RT between the no-pain blocks and the pain blocks of the NI task were used to characterize A and P types as we have done in our previous studies: The RTmean of the no-pain blocks was subtracted from the RTmean of the pain blocks (ΔRTmean = RTmean pain—RTmean no pain) for each participant separately (12, 13). Thus, the A-types exhibit negative ΔRTmean values which reflect a general increase in task performance speed from the no-pain to pain condition of the NI task, whereas the P-types exhibit positive ΔRTmean values, which reflect a general decrease in task performance speed from the no-pain to pain condition of the NI task.

2.3. Quantifying attention to pain

We used two approaches to quantify attention to pain; the IAP measure we have developed in our lab that uses an experimental stimulus, and an assessment of pain rumination which is a measure that does not require applying stimuli:

Participants underwent an experience sampling of experimental pain stimuli previously developed by Kucyi et al. (15) to quantify an individual's IAP. To do this, participants were asked to stare at a blank screen with a white fixation cross during which a 20 s transcutaneous electrical stimulus was delivered to the skin overlying the left median nerve (300-PV, Empi Inc.) at an intensity to evoked pain rated at 40–60/100 (0 = no pain, 100 = most intense pain imaginable) that was calibrated prior to the task for each participant separately. See the **Supplementary Materials** for more information regarding the stimulus calibration procedure. After 20 s, the pain stimulus stopped, and a probe popped up on the screen that asked participants to indicate whether their attention had been "only on pain", "mostly on pain", "mostly on something else", or "only on something else". After the participants responded to this prompt or after 8 s had passed, an inter-stimulus interval with the blank screen and white fixation cross popped up without pain for 22 s. In total, participants underwent 20 trials of this task. Based on the proportions of trials that reported attention towards pain vs. attention towards something else, a single IAP score was calculated for each participant that ranged from −2 (always attending to something else) to +2 (always attending to pain) as follows (15):

$$\mathrm{IAP} = [(2n_{only\ pain} + n_{mostly\ pain}) - (2n_{only\ else} + n_{mostly\ else})]/(n_{total})$$

where n = number of trials

We quantified PR using the 4 item pain rumination subscale of the pain catastrophizing scale (PCS-R). The entire PCS consists of 13 items, each of which is rated on a five-point Likert Scale (0—not at all, 4- all the time). A score between 0 (lowest PR score) and 16 (highest PR score) was generated for each participant based on their responses.

2.4. Analyses

Analyses were conducted using R-Studio, Graphpad-Prism 7, and Microsoft Excel. We used parametric and non-parametric tests as appropriate: Independent sample t-tests were used to examine differences in RTmean and RTv, respectively, between the HCs and AS group. This was done for each NI task condition (no-pain and pain), separately. The difference in IAP scores between HCs and the AS group were examined using an independent sample t-test. The difference in PCS-R scores between the HCs and the AS group were examined using a Mann-Whitney-u test.

We used Spearman's correlations to determine the correlation between IAP and PCS-R scores in the HCs and the AS group, and for the correlations between (i) ΔRTmean values and IAP scores, (ii) ΔRTmean values and PCS-R scores, (iii) ΔRTv values and IAP scores, and (iv) ΔRTv values and PCS-R scores for the HCs and the AS group.

3. Results

3.1. Attention to task-dominant (A-type) and pain-dominant (P-type) characterization and performance on the NI task

Within each cohort of HCs and AS, we delineated 32 A-type individuals and 11 P-type individuals (**Figure 1**). There were no significant differences in RT mean between the HCs and AS group in either the no-pain condition (HC: M = 1320.47 ms, SD = 163.39 ms; AS: M = 1390.40 ms, SD = 186.64 ms) ($t = 1.85$, $p = 0.068$, Cohen's d = 0.40) or in the pain condition (HC: M = 1276.51 ms, SD = 167.39 ms; AS: M = 1343.93 ms, SD = 173.96 ms) ($t = 1.83$, $p = 0.071$, Cohen's d = 0.39) of the NI task (**Figure 2**). Furthermore, there was no significant difference in RTv between the HCs and the AS group in the pain condition of the NI Task (HC: $M = 8.6 \times 10^4$ ms^2, $SD = 3.0 \times 10^4$ ms^2; AS: $M = 9.5 \times 10^4$ ms^2, $SD = 4.2 \times 10^4$ ms^2) ($t = 1.24$, $p = 0.22$, Cohen's d = 0.27). However, as shown in **Figure 3**, there was a significant difference in task RTv between the HCs and the AS group in the no-pain condition ($t = 2.15$, $p = 0.035$, Cohen's d = 0.46) such that the AS group exhibited an overall higher mean RTv than the HCs (HC: $M = 9.1 \times 10^4$ ms^2, $SD = 3.5 \times 10^4$ ms^2; AS: $M = 1.1 \times 10^5$ ms^2, $SD = 4.8 \times 10^4$ ms^2).

3.2. Attention to pain: intrinsic attention to pain and pain rumination

We examined two metrics of attention to pain: IAP and PCS-R. We did not find any significant group differences in IAP scores (HC: M = 0.033, SD = 0.76; AS: M = −0.11, SD = 0.85) ($t = 0.80$, $p = 0.43$, Cohen's d = 0.17) or in the PCS-R scores (HC: M = 4.93, SD = 3.75; AS: M = 4.47, SD = 4.04) ($p = 0.43$, Cohen's d = 0.12) between the HCs and the AS group (**Figure 4**). Furthermore, the IAP scores were not significantly correlated with the PCS-R scores in the HCs (rho = −0.0056, $p = 0.97$). However, the correlation between IAP scores and the PCS-R scores in the AS group showed a statistically significant trend (rho = 0.30, $p = 0.054$) (**Figure 5**).

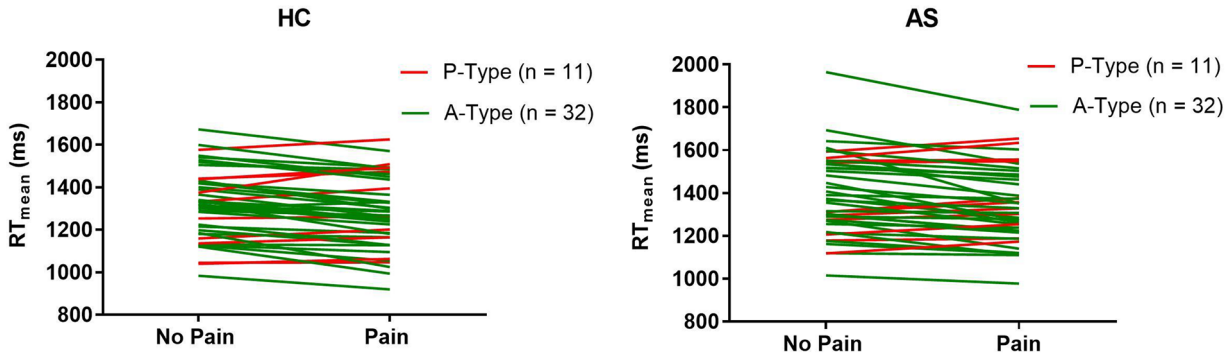


FIGURE 1
Change in individual mean task reaction times (RT) from the no-pain to the pain blocks for each individual in the healthy controls (HCs) and ankylosing spondylitis (AS) groups. A-types are represented by the green lines. *P*-types are represented by the red lines.

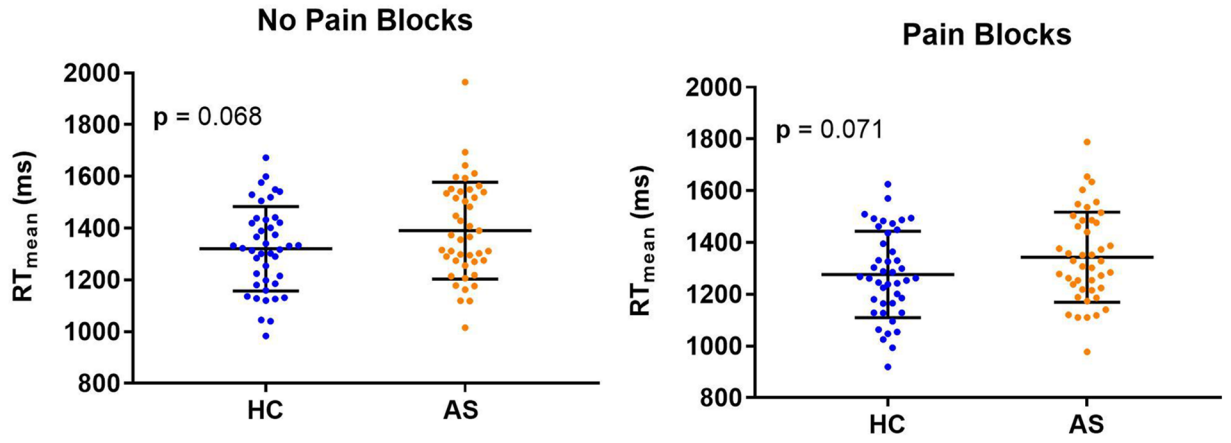


FIGURE 2
Individual and mean group task reaction times (RTs) for the healthy controls (HCs) and the ankylosing spondylitis (AS) group. Individual HCs are represented by the blue dots and individuals in the AS group are represented by the orange dots.

3.3. Relationship between task performance and the metrics of attention to pain

We next examine the relationship between task performance and metrics of attention to pain at the individual and group level. We did not find any significant correlation between the IAP scores and ΔRTmean in the HCs (rho = −0.22, $p = 0.15$) or in the AS group (rho = −0.18, $p = 0.26$) (**Figure 6**). We also examined the consistency of task performance (**Figure 7**) and again found there was no significant correlation between the IAP scores and ΔRTv in the HCs (rho = −0.17, $p = 0.28$) and the AS group (rho = −0.12, $p = 0.44$). Also, there was no significant correlation between the PCS-R scores and ΔRTmean (**Figure 8**) in the HCs (rho = −0.14, $p = 0.38$) and the AS group (rho = −0.16, $p = 0.32$). Finally, we also did not find any significant correlations between the PCS-R scores and ΔRTv (**Figure 9**) in the HCs (rho = 0.093, $p = 0.55$) or in the AS group (rho = −0.035, $p = 0.82$).

3.4. Supplemental analysis: pain scores in ankylosing spondylitis group, sex differences in task performance and metrics of attention to pain

As a supplementary exploratory analysis, pain intensity scores at the time of testing and on average over 4 weeks were obtained from the AS group using the painDETECT questionnaire, and compared across the A-and P-types (30, 31). There was no statistically significant differences in current (i.e., state) pain scores across A- and P-types. However, there was a marginally significant difference in average (i.e., trait) pain scores over 4 weeks between the A- and P-types such that P-types had higher overall average pain scores than the A-types (see **Supplemental Materials**). As well, we examined sex differences in RTmean, RTv, IAP scores and PCS-R scores in the HCs and AS group. There were no statistically significant sex differences identified other than a marginally significant difference in RTv between males and females in the AS

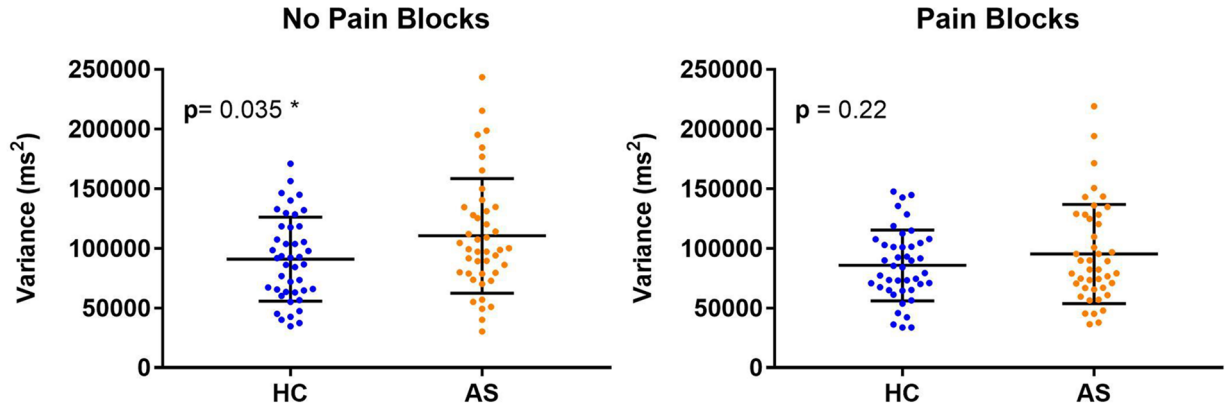


FIGURE 3
Individual and mean group task reaction time variance (RTv). Higher task RTv was found in the healthy controls (HCs) compared to the ankylosing spondylitis (AS) group in the no-pain blocks (left panel) but not in the pain blocks (right panel). The HCs are represented by the blue dots. The AS group is represented by the orange dots.

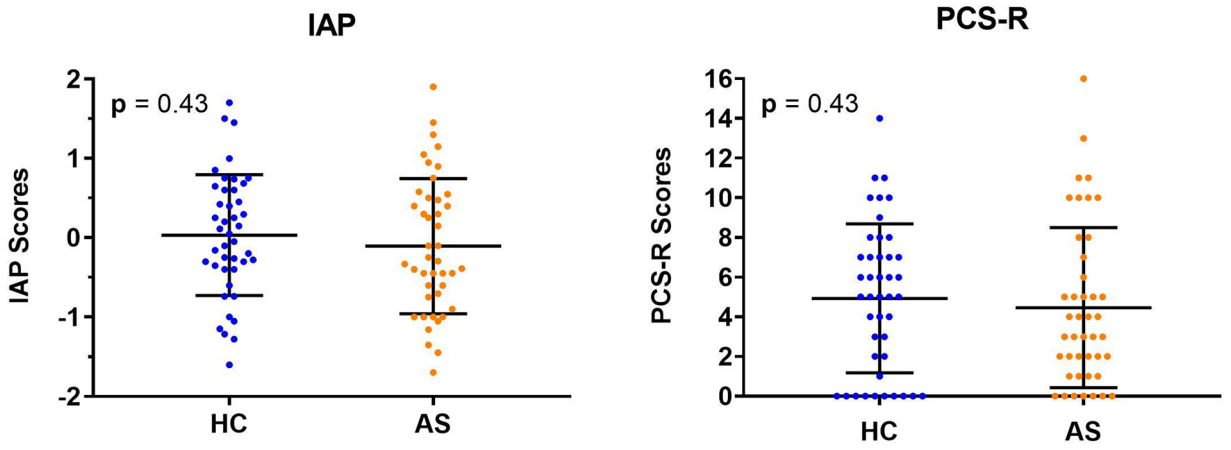


FIGURE 4
Individual and group mean intrinsic attention to pain (IAP) and pain rumination scores. Neither the IAP scores (left panel) or rumination scores (right panel) from the pain catastrophizing scale, rumination subscale (PCS-R) were significantly different between healthy controls (HCs) and those in the ankylosing spondylitis (AS) group. The HCs are represented by the blue dots. The AS group is represented by the orange dots.

group for the pain condition of the NI task (see **Supplemental Materials**). These findings were not included as part of our main analyses due to low and unequal group sizes between the A-P types and the sexes (arising from the male-predominance of AS).

4. Discussion

Categorizing people with the A-P and IAP behavioural phenotypes (10, 12–15) provides insight to understand acute pain and attention interactions but in chronic pain populations this assessment could be confounded by ongoing and fluctuating chronic pain (32–34). Additionally, mechanisms underlying pain-attention interactions may differ for acute and chronic pains. As a first step towards determining the suitability of A-P and IAP testing in chronic pain populations, we characterized A-P and IAP behavioural phenotypes in people with chronic pain associated with AS. Our main findings were that (1) the current A-P and IAP protocols are not suitable for people with chronic pain, and (2) PR can be used as a supplement to IAP to capture attention to pain in chronic pain populations without the need for experimental pain stimuli.

We identified both A and P-type individuals in both the AS group and healthy controls, and there were more A-types than P-types in both groups. This was a surprising revelation, considering that many studies suggest people with chronic pain can demonstrate difficulty attending away from their chronic pain (35–37) and overall impairment in many cognitive domains and tasks (24, 36, 38–40). Therefore, we expected that the AS group's behavioural performance would be affected by other pain experiences like the NI task's concurrent experimental acute pain, and that they would exhibit slower RTs in the task pain condition compared to the no-pain condition. This could have been a result of sampling bias, as A-types might be more likely to volunteer for

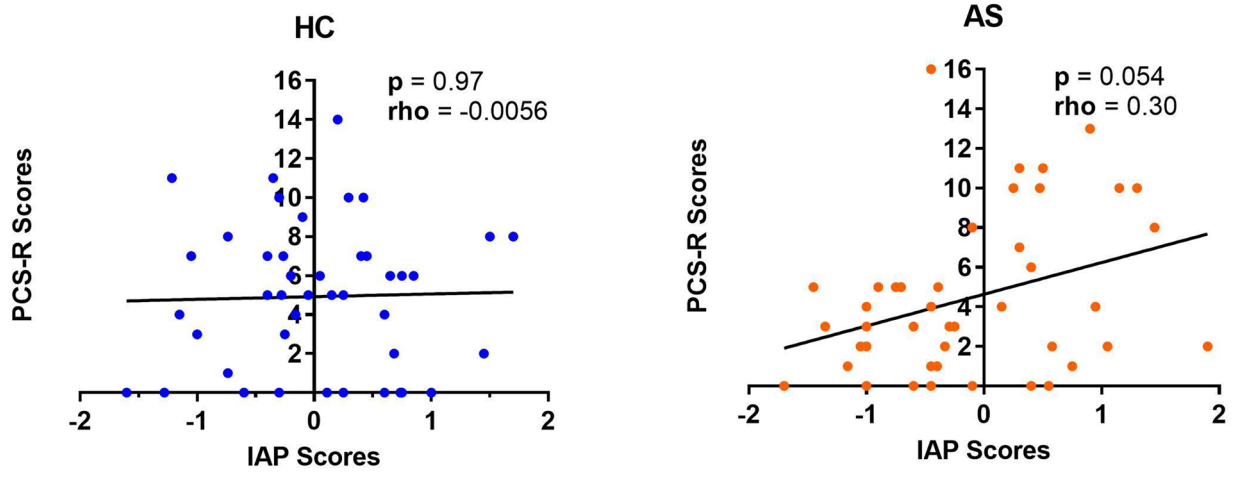


FIGURE 5
Relationship between intrinsic attention to pain (IAP) and pain rumination. The IAP scores were not significantly correlated with the pain rumination scores (derived from the rumination subscale of the pain catastrophizing scale) (PCS-R) in healthy controls (HCs) (left panel) but there was a marginally significant positive correlation between the IAP scores and the PCS-R scores in the ankylosing spondylitis (AS) group (right panel). The HCs are represented by the blue dots. The AS group is represented by the orange dots.

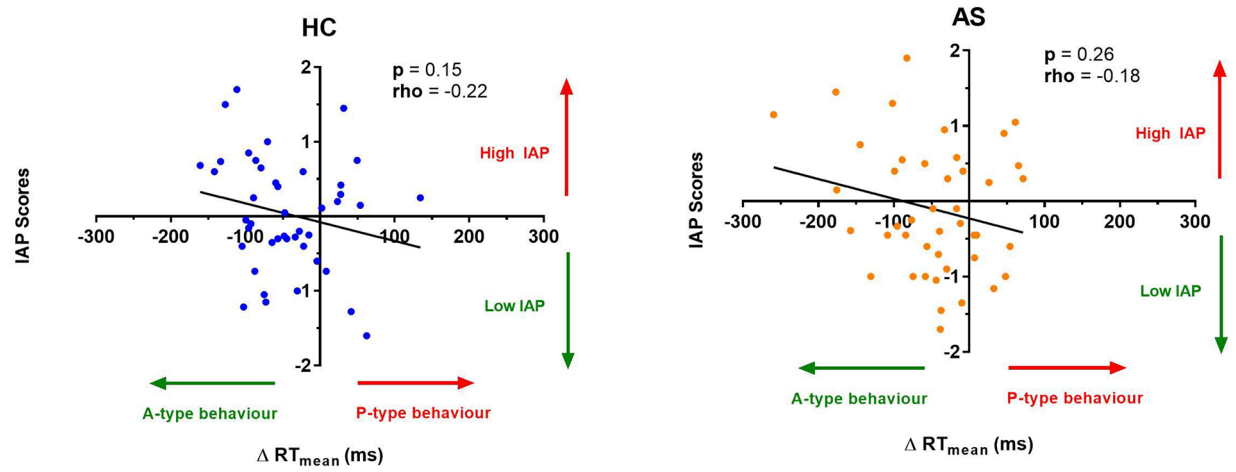


FIGURE 6
Relationship between an individual's intrinsic attention to pain (IAP) and the effect of pain on task reaction time (RT). No significant correlation was found between the IAP scores and the ΔRT mean in the healthy controls (HCs) (left panel) or for the ankylosing spondylitis (AS) group (right panel). The HCs are represented by the blue dots. The AS group is represented by the orange dots. A-type behavior is represented by a negative ΔRT mean value (values that are on the left side of the horizontal axis). *P*-type behavior is represented by a positive ΔRT mean value (values that are on the right side of the horizontal axis).

a pain study than P-types. Another explanation could be that the AS group has built resilience to pain experiences during the course of their disease, thus allowing them to perform better during the NI Task. The AS group may have learned overtime to function normally and accept their pain having been exposed to chronic pain from having AS. Many people with chronic pain build resilience towards their pain, that allows them to operate in their daily lives (41–43). As well, those who have higher acceptance of their chronic pain report lower levels of disability and higher levels of functionality in daily activities than those with lower acceptance (44, 45). The level of resilience and acceptance of chronic pain in the AS group may have supported their ability to perform on the NI task, however we have not tested if this is true.

Similarly to Cheng et al. (12), we used RT variability as a metric of performance to gain insight into inconsistencies in RTs (46), as this measure is considerably understudied in studies looking at pain interference and performance variability (12). Interestingly, there were no differences in RT variances between the HCs and the AS group in the pain condition of the NI task, but there were significant differences in the no-pain condition such that the chronic pain group exhibited overall higher variance compared to the HCs despite there being no experimental pain applied. Since experimental pain is not driving this variance in the AS group, it is possible that the chronic pain experienced by this population could be interfering with the consistency of their RTs on this task. However, our current protocols do not account for fluctuating and

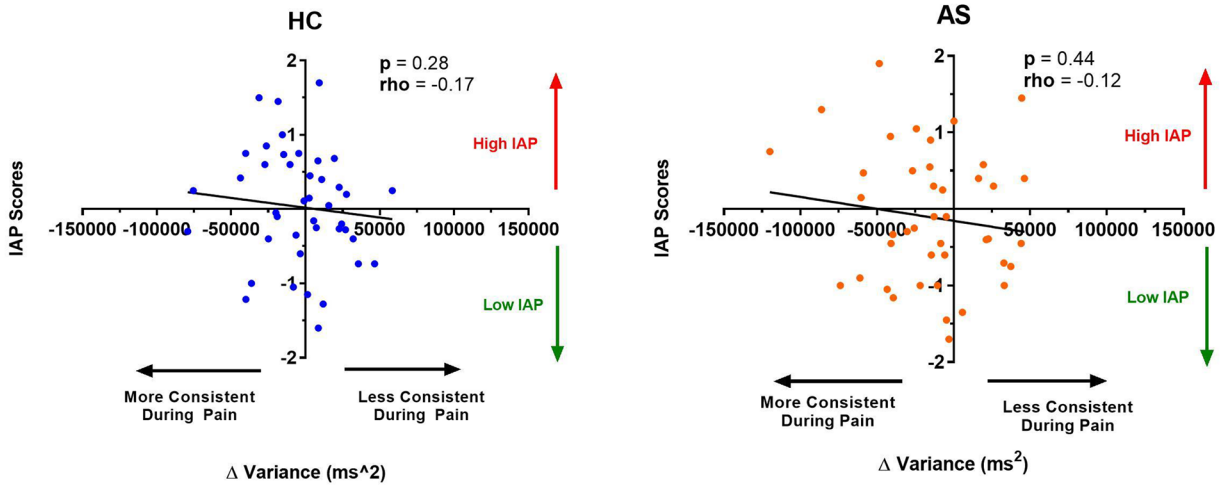


FIGURE 7
Relationship between an individual's intrinsic attention to pain (IAP) and the effect of pain on variance in task reaction time (RTv). No significant correlation was found between the IAP scores and the ΔRTv in the healthy controls (HCs) (left panel) or for the ankylosing spondylitis (AS) group (right panel). The HCs are represented by the blue dots. The AS group is represented by the orange dots. More consistent RTs during pain is represented by a negative ΔRTv value (values that are on the left side of the horizontal axis). Less consistent RTs during pain is represented by a positive ΔRTv value (values that are on the right side of the horizontal axis).

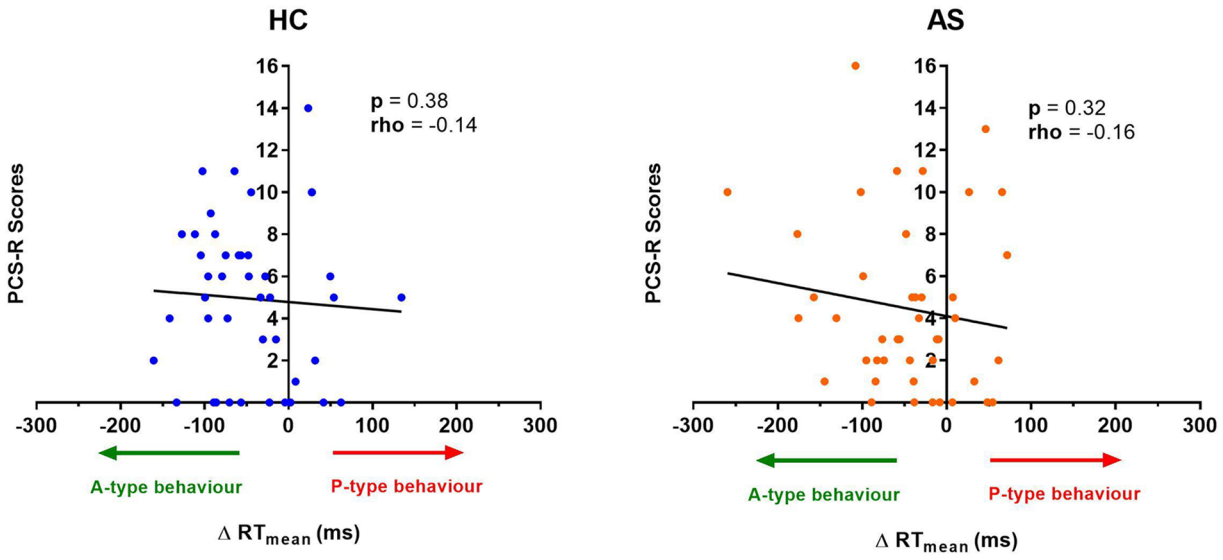


FIGURE 8
Relationship between an individual's pain rumination score and the effect of pain on their task reaction time (RT). No significant correlation was found between the pain rumination scores (derived from the rumination subscale of the pain catastrophizing scale) (PCS-R) and the ΔRT mean in the healthy controls (HCs) or for the ankylosing spondylitis (AS) group. The HCs are represented by the blue dots. The AS group is represented by the orange dots. A-type behavior is represented by a negative ΔRT mean value (values that are on the left side of the horizontal axis). P-type behavior is represented by a positive ΔRT mean value (values that are on the right side of the horizontal axis).

spontaneous chronic pain that the AS group may be experiencing during the NI task. To properly determine if an individual is an A or P type, RTs during the task need to be compared in a condition that has no-pain to a condition where pain is concurrently occurring within the behavioural task. Since we cannot verify the occurrence of chronic pain during the NI task, this suggests that these protocols are not appropriate for people with chronic pain and that further modifications need to be made that take chronic pain into consideration.

It is possible the AS group could have experienced an inhibitory pain modulation or "pain inhibits pain" (47) phenomena during the pain condition of the NI task, such that the experimental acute pain inhibited their chronic pain, which allowed them to produce more consistent RTs in the task pain condition compared to the no-pain condition. Including recordings of chronic pain intensity in future iterations of the A-P/IAP protocols would allow us to further explore whether experimental pain inhibits chronic pain experiences during these behavioral tasks.

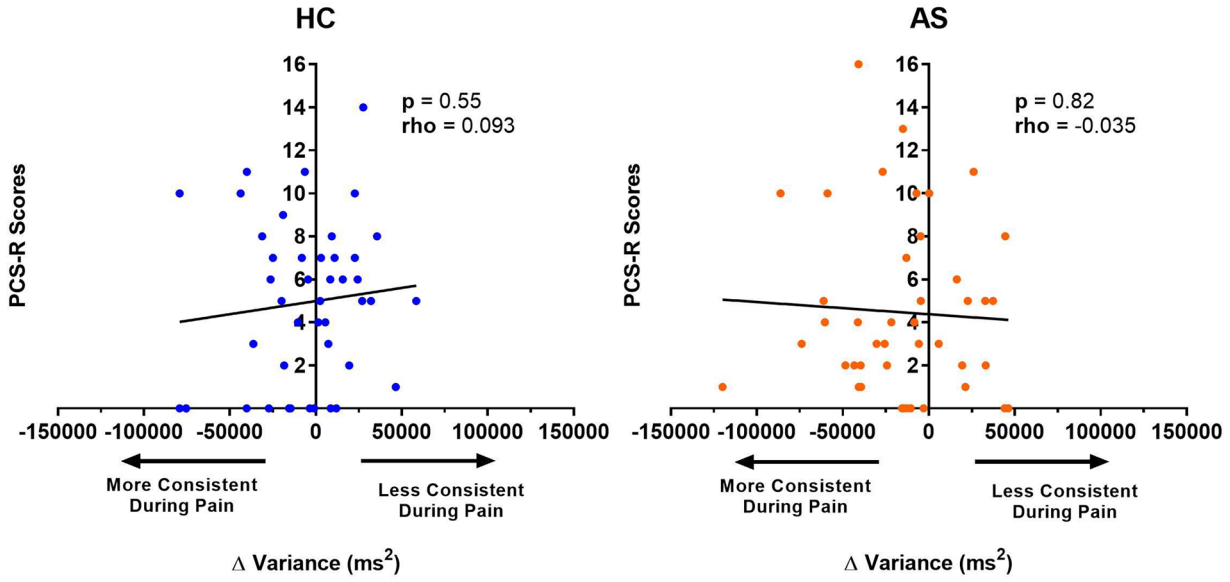


FIGURE 9
Relationship between an individual's pain rumination score and the effect of pain on their variance in task reaction time (RTv). No significant correlation was found between the pain rumination scores (derived from the rumination subscale of the pain catastrophizing scale) (PCS-R) and the ΔRTv in the healthy controls (HCs) or for the ankylosing spondylitis (AS) group. The HCs are represented by the blue dots. The AS group is represented by the orange dots. More consistent RTs during pain is represented by a negative ΔRTv value (values that are on the left side of the horizontal axis). Less consistent RTs during pain is represented by a positive ΔRTv value (values that are on the right side of the horizontal axis).

The marginally significant correlation between IAP scores and PCS-R scores in the AS group suggests PCS-R may be able to quantify "attention to pain" in people with chronic pain, alongside IAP. However, these measures likely quantify different aspects of "attention to pain". IAP has been shown to be a "trait-like" measure of attention to pain (1, 15), and probes participants to think only about whether their attention was directed towards the administered pain or mind-wandering towards something else. PR is defined as "…perseverative negative thinking about pain" (21), and involves characterizing an individual's tendency to negatively ruminate about their pain experiences (21, 48, 49). Unlike IAP which only quantifies whether or not attention is towards pain, PR has a negative affect (50, 51) component towards pain that quantifies negative thinking about pain experiences. The questions presented in the PCS-R have more emotional valence than the questions probed in the IAP protocol. This may explain why these two metrics were only marginally significantly positively correlated in the AS group and not significantly correlated in the HCs. As well, these findings are inconsistent with previous work in our lab that has shown that IAP scores and PCS scores in healthy individuals exhibit a modest positive trend (15). However, it is important to note that this study looked at scores generated from the entire PCS and not the relationship between IAP scores and PCS-R scores alone. The findings in the current study encourages the need for future work to investigate additional metrics that quantify attention to pain similarly to IAP but do not require an experimental pain stimulus and that do not capture pain affect as prominently as the PCS-R.

We also note that the experiences people are reflecting on when they complete the PCS-R or how long ago they occurred are not known. Someone with chronic pain who experiences persistent chronic pain could more readily recall their pain-related cognitions when completing the PCS-R. This assessment of their tendency to catastrophize (52) and/or ruminate about pain, could then better capture in the PCS-R a more accurate representation of their tendency to attend to pain. In contrast, because healthy individuals are not experiencing pain at the time of assessment, they may vary in how difficult it is to recall a pain experience and their pain-related cognitions when completing the PCS-R (52). This may impact how the PCS-R can capture their tendency to ruminate and attend to pain. This issue further highlights the importance to consider other metrics in future studies that can capture attention to pain and be used in both healthy individuals and people with chronic pain.

We did not find a relationship between the measures of attention to pain (IAP and PCS-R scores) and the performance measures of the NI task (ΔRTmean and ΔRTv). This was an unexpected finding considering our previous work has shown a significant positive correlation between IAP scores and ΔRTmean in healthy individuals (15). Both the A-P and IAP protocols are meant to capture an understanding of pain and attention interactions in individuals, but the lack of correlation suggests these two behavioural phenotypes reflect pain and attention interactions in different ways. IAP reflects a trait-like measure of attention to pain (15, 53). It is currently unknown whether the A and P type characterizations are trait or state-like designations of pain and attention interactions. However, consistent evidence of structural and functional brain region differences between A- and P-types (12–15) suggest that these behavioural characterizations are trait-like in nature. It is clear that more work needs to be done to investigate the trait or state like qualities of the A-P and IAP behavioral phenotypes in both healthy and chronic pain populations.

We also note a study limitation that should be addressed arising from examining AS is that it is a condition that is predominately found in males (54–56), and this has limited our ability to include an equal number of female participants and fully address any sex differences. There is a considerable amount of studies that suggest that chronic pain experiences and prevalence are different between males and females (57–63), and so it is important to observe whether there are differences in the reflection of pain and attention interactions between the sexes. Finally, a limitation is that our study did not include a non-painful stimulation control condition. Thus, we cannot factor out the possibility of a non-specific stimulation distraction effect impacting the participants' performance on our behavioural tasks. Our future iterations of these behavioural tasks will include control conditions that take this possibility into consideration. As well, our future studies will use larger sample sizes to explicitly examine sex differences and behavioural phenotypes in greater detail.

In conclusion, the current A-P/IAP behavioural phenotype characterization protocols are likely not appropriate for people with chronic pain as they do not account for the occurrence of chronic pain throughout their behavioural tasks. Although, PR could be used as a supplement to quantify attention to pain alongside IAP, other metrics are needed to be investigated that are more closely related to IAP and circumvent the use of experimental pain stimuli so that IAP phenotypes can be characterized in people with chronic pain. Attention-based therapies of chronic pain such as cognitive behavioural therapy (CBT) provide inconsistent success to help improve chronic pain (16, 17, 64–68), but a greater understanding of behavioural phenotypes of pain and attention interactions may contribute to a better identification of individuals most likely to benefit from attention-based chronic pain therapies such as CBT.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author/s.

Ethics statement

The studies involving human participants were reviewed and approved by University Health Network Research Ethics Board. The patients/participants provided their written informed consent to participate in this study.

Author contributions

KD and VVS conceived this study. Data was collected by JC, KH, AR, NO, RB, and JK. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the Mayday Fund, and by the Canadian Institute of Health Research (operating grant PJT-156409 to KD).

Acknowledgments

The author thanks the Davis lab co-ordinator Ariana Besik for their support and contribution to the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpain.2023.1057659/full#supplementary-material.

References

1. Kucyi A, Davis KD. The dynamic pain connectome. *Trends Neurosci.* (2015) 38 (2):86–95. doi: 10.1016/j.tins.2014.11.006

2. Melzack R. Pain and the neuromatrix in the brain. *J Dent Educ.* (2001) 65 (12):1378–82. doi: 10.1002/j.0022-0337.2001.65.12.tb03497.x

3. Bantick SJ, Wise RG, Ploghaus A, Clare S, Smith SM, Tracey I. Imaging how attention modulates pain in humans using functional MRI. *Brain.* (2002) 125(Pt 2):310–9. doi: 10.1093/brain/awf022

4. Buhle J, Wager TD. Performance-dependent inhibition of pain by an executive working memory task. *Pain.* (2010) 149(1):19–26. doi: 10.1016/j.pain.2009.10.027

5. Lier EJ, Rijn C, Vries M, Goor H, Oosterman JM. The interaction between pain and cognition: on the roles of task complexity and pain intensity. *Scand J Pain.* (2022) 22 (2):385–95. doi: 10.1515/sjpain-2021-0119

6. Miron D, Duncan GH, Bushnell CM. Effects of attention on the intensity and unpleasantness of thermal pain. *Pain.* (1989) 39(3):345–52. doi: 10.1016/0304-3959 (89)90048-1

7. Petrovic P, Petersson KM, Ghatan PH, Stone-Elander S, Ingvar M. Pain-related cerebral activation is altered by a distracting cognitive task. *Pain.* (2000) 85(1–2):19–30. doi: 10.1016/S0304-3959(99)00232-8

8. Rischer KM, González-Roldán AM, Montoya P, Gigl S, Anton F, van der Meulen M. Distraction from pain: the role of selective attention and pain catastrophizing. *Eur J Pain*. (2020) 24(10):1880–91. doi: 10.1002/ejp.1634

9. Seminowicz DA, Davis KD. Interactions of pain intensity and cognitive load: the brain stays on task. *Cereb Cortex*. (2007) 17(6):1412–22. doi: 10.1093/cercor/bhl052

10. Seminowicz DA, Mikulis DJ, Davis KD. Cognitive modulation of pain-related brain responses depends on behavioral strategy. *Pain*. (2004) 112(1-2):48–58. doi: 10.1016/j.pain.2004.07.027

11. Veldhuijzen DS, Kenemans JL, de Bruin CM, Olivier B, Volkerts ER. Pain and attention: attentional disruption or distraction? *J Pain*. (2006) 7(1):11–20. doi: 10.1016/j.jpain.2005.06.003

12. Cheng JC, Bosma RL, Hemington KS, Kucyi A, Lindquist MA, Davis KD. Slow-5 dynamic functional connectivity reflects the capacity to sustain cognitive performance during pain. *NeuroImage*. (2017) 157:61–8. doi: 10.1016/j.neuroimage.2017.06.005

13. Erpelding N, Davis KD. Neural underpinnings of behavioural strategies that prioritize either cognitive task performance or pain. *Pain*. (2013) 154(10):2060–71. doi: 10.1016/j.pain.2013.06.030

14. Rogachov A, Cheng JC, Erpelding N, Hemington KS, Crawley AP, Davis KD. Regional brain signal variability: a novel indicator of pain sensitivity and coping. *Pain*. (2016) 157(11):2483–92. doi: 10.1097/j.pain.0000000000000665

15. Kucyi A, Salomons TV, Davis KD. Mind wandering away from pain dynamically engages antinociceptive and default mode brain networks. *Proc Natl Acad Sci U S A*. (2013) 110(46):18692–7. doi: 10.1073/pnas.1312902110

16. Broderick JE, Keefe FJ, Schneider S, Junghaenel DU, Bruckenthal P, Schwartz JE, et al. Cognitive behavioral therapy for chronic pain is effective, but for whom? *PAIN*. (2016) 157(9):2115–2123. doi: 10.1097/j.pain.0000000000000626

17. Williams AC, Eccleston C, Morley S. Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane Database Syst Rev*. (2012) 11 (11):Cd007407. doi: 10.1002/14651858.CD007407.pub3

18. Arendt-Nielsen L, Curatolo M, Drewes A. Human experimental pain models in drug development: translational pain research. *Curr Opin Investig Drugs*. (2007) 8 (1):41–53. Available from: https://pubmed.ncbi.nlm.nih.gov/17263184/.

19. Campbell CM, Gilron I, Doshi T, Raja S. Designing and conducting proof-of-concept chronic pain analgesic clinical trials. *Pain Rep*. (2019) 4(3):e697. doi: 10.1097/PR9.0000000000000697

20. Schmelz M. Translating nociceptive processing into human pain models. *Exp Brain Res*. (2009) 196(1):173–8. doi: 10.1007/s00221-009-1809-2

21. Kucyi A, Moayedi M, Weissman-Fogel I, Goldberg MB, Freeman BV, Tenenbaum HC, et al. Enhanced medial prefrontal-default mode network functional connectivity in chronic pain and its association with pain rumination. *J Neurosci*. (2014) 34 (11):3969–75. doi: 10.1523/JNEUROSCI.5055-13.2014

22. Rogachov A, Cheng JC, Hemington KS, Bosma RL, Kim JA, Osborne NR, et al. Abnormal low-frequency oscillations reflect trait-like pain ratings in chronic pain patients revealed through a machine learning approach. *J Neurosci*. (2018) 38 (33):7293. doi: 10.1523/JNEUROSCI.0578-18.2018

23. Linden S VD, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum*. (1984) 27(4):361–8. doi: 10.1002/art.1780270401

24. Eccleston C. Chronic pain and distraction: an experimental investigation into the role of sustained and shifting attention in the processing of chronic persistent pain. *Behav Res Ther*. (1995) 33(4):391–405. doi: 10.1016/0005-7967(94)00057-Q

25. Eccleston C. Chronic pain and attention: a cognitive approach. *Br J Clin Psychol*. (1994) 33(4):535–47. doi: 10.1111/j.2044-8260.1994.tb01150.x

26. Morton J. Categories of interference: verbal Mediation and Conflict in Card Sorting. *Br J Psychol*. (1969) 60:329–46. doi: 10.1111/j.2044-8295.1969.tb01204.x

27. Windes JD. Reaction time for numerical coding and naming of numerals. *J Exp Psychol*. (1968) 78(2):318–22. doi: 10.1037/h0026289

28. Luce RD. *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press (1991). Available from: doi: https://doi.org/10.1093/acprof:oso/9780195070019.001.0001.

29. Whelan R. Effective analysis of reaction time data. *Psychol Rec*. (2008) 58(3):475–82. doi: 10.1007/BF03395630

30. Freynhagen R, Baron R, Gockel U, Tölle TR. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. *Curr Med Res Opin*. (2006) 22(10):1911–20. doi: 10.1185/030079906X132488

31. Wu Q, Inman RD, Davis KD. Neuropathic pain in ankylosing spondylitis: a psychophysics and brain imaging study. *Arthritis Rheum*. (2013) 65(6):1494–503. doi: 10.1002/art.37920

32. Apkarian AV, Baliki MN, Geha PY. Towards a theory of chronic pain. *Prog Neurobiol*. (2009) 87(2):81–97. doi: 10.1016/j.pneurobio.2008.09.018

33. Baliki MN, Chialvo DR, Geha PY, Levy RM, Harden RN, Parrish TB, et al. Chronic pain and the emotional brain: specific brain activity associated with spontaneous fluctuations of intensity of chronic back pain. *J Neurosci*. (2006) 26(47):12165–73. doi: 10.1523/JNEUROSCI.3576-06.2006

34. Foss JM, Apkarian AV, Chialvo DR. Dynamics of pain: fractal dimension of temporal variability of spontaneous pain differentiates between pain states. *J Neurophysiol*. (2006) 95(2):730–6. doi: 10.1152/jn.00768.2005

35. McCracken LM. "Attention" to pain in persons with chronic pain: a behavioral approach. *Behav Ther*. (1997) 28(2):271–84. doi: 10.1016/S0005-7894(97)80047-0

36. Viane I, Crombez G, Eccleston C, Devulder J, De Corte W. Acceptance of the unpleasant reality of chronic pain: effects upon attention to pain and engagement with daily activities. *Pain*. (2004) 112(3):282–8. doi: 10.1016/j.pain.2004.09.008

37. Crombez G, Viane I, Eccleston C, Devulder J, Goubert L. Attention to pain and fear of pain in patients with chronic pain. *J Behav Med*. (2013) 36(4):371–8. doi: 10.1007/s10865-012-9433-1

38. Apkarian AV, Sosa Y, Krauss BR, Thomas PS, Fredrickson BE, Levy RE, et al. Chronic pain patients are impaired on an emotional decision-making task. *Pain*. (2004) 108(1–2):129–36. doi: 10.1016/j.pain.2003.12.015

39. Higgins DM, Martin AM, Baker DG, Vasterling JJ, Risbrough V. The relationship between chronic pain and neurocognitive function: a systematic review. *Clin J Pain*. (2018) 34(3):262–75. doi: 10.1097/AJP.0000000000000536

40. Mazza S, Frot M, Rey AE. A comprehensive literature review of chronic pain and memory. *Prog Neuropsychopharmacol Biol Psychiatry*. (2018) 87(Pt B):183–92. doi: 10.1016/j.pnpbp.2017.08.006

41. Goubert L, Trompetter H. Towards a science and practice of resilience in the face of pain. *Eur J Pain*. (2017) 21(8):1301–15. doi: 10.1002/ejp.1062

42. Newton-John TR, Mason C, Hunter M. The role of resilience in adjustment and coping with chronic pain. *Rehabil Psychol*. (2014) 59(3):360–5. doi: 10.1037/a0037023

43. Sturgeon JA, Zautra AJ. Resilience: a new paradigm for adaptation to chronic pain. *Curr Pain Headache Rep*. (2010) 14(2):105–12. doi: 10.1007/s11916-010-0095-9

44. McCracken LM, Eccleston C. A prospective study of acceptance of pain and patient functioning with chronic pain. *Pain*. (2005) 118(1):164–9. doi: 10.1016/j.pain.2005.08.015

45. McCracken LM, Eccleston C. Coping or acceptance: what to do about chronic pain? *Pain*. (2003) 105(1):197–204. doi: 10.1016/S0304-3959(03)00202-1

46. Kofler MJ, Rapport MD, Sarver DE, Raiker JS, Orban SA, Friedman LM, et al. Reaction time variability in ADHD: a meta-analytic review of 319 studies. *Clin Psychol Rev*. (2013) 33(6):795–811. doi: 10.1016/j.cpr.2013.06.001

47. Yarnitsky D. Conditioned pain modulation (the diffuse noxious inhibitory control-like effect): its relevance for acute and chronic pain states. *Curr Opin Anaesthesiol*. (2010) 23(5):611–5. doi: 10.1097/ACO.0b013e32833c348b

48. Craner JR, Gilliam WP, Sperry JA. Rumination, magnification, and helplessness: how do different aspects of pain catastrophizing relate to pain severity and functioning? *Clin J Pain*. (2016) 32(12):1028–35. doi: 10.1097/ajp.0000000000000355

49. Sullivan M, Bishop S, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess*. (1995) 7:524–32. doi: 10.1037/1040-3590.7.4.524

50. Gaskin ME, Greene AF, Robinson ME, Geisser ME. Negative affect and the experience of chronic pain. *J Psychosom Res*. (1992) 36(8):707–13. doi: 10.1016/0022-3999(92)90128-O

51. Price DD, Harkins SW. The affective-motivational dimension of pain A two-stage model. *APS Journal*. (1992) 1(4):229–39. doi: 10.1016/1058-9139(92)90054-G

52. Day MA, Young G, Jensen MP. Differentiating state versus trait pain catastrophizing. *Rehabil Psychol*. (2021) 66(1):39–49. doi: 10.1037/rep0000318

53. Davis KD, Cheng JC. Differentiating trait pain from state pain: a window into brain mechanisms underlying how we experience and cope with pain. *Pain Rep*. (2019) 4(4):e735. doi: 10.1097/PR9.0000000000000735

54. Braun J, Sieper J. Ankylosing spondylitis. *Lancet*. (2007) 369(9570):1379–90. doi: 10.1016/S0140-6736(07)60635-7

55. Feldtkeller E, Khan MA, van der Heijde D, van der Linden S, Braun J. Age at disease onset and diagnosis delay in HLA-B27 negative vs. Positive patients with ankylosing spondylitis. *Rheumatol Int*. (2003) 23(2):61–6. doi: 10.1007/s00296-002-0237-4

56. Lockshin MD. Sex ratio and rheumatic disease. *Isr Med Assoc J*. (2001) 3(7):511–6. Available from: https://www.ima.org.il/MedicineIMAJ/viewarticle.aspx?year=2001&month=07&page=511

57. Berkley KJ. Sex differences in pain. *Behav Brain Sci*. (1997) 20(3):371–80; discussion 435–513. doi: 10.1017/S0140525X97221485

58. Fauchon C, Meunier D, Rogachov A, Hemington KS, Cheng JC, Bosma RL, et al. Sex differences in brain modular organization in chronic pain. *PAIN*. (2021) 162 (4):1188–200. doi: 10.1097/j.pain.0000000000002104

59. Hashmi JA, Davis KD. Deconstructing sex differences in pain sensitivity. *PAIN®*. (2014) 155(1):10–3. doi: 10.1016/j.pain.2013.07.039

60. Mogil JS. Sex differences in pain and pain inhibition: multiple explanations of a controversial phenomenon. *Nat Rev Neurosci*. (2012) 13(12):859–66. doi: 10.1038/nrn3360

61. Racine M, Tousignant-Laflamme Y, Kloda LA, Dion D, Dupuis G, Choinière M. A systematic literature review of 10 years of research on sex/gender and experimental pain perception—part 1: are there really differences between women and men? *Pain*. (2012) 153(3):602–18. doi: 10.1016/j.pain.2011.11.025

62. Racine M, Tousignant-Laflamme Y, Kloda LA, Dion D, Dupuis G, Choinière M. A systematic literature review of 10 years of research on sex/gender and pain perception—

part 2: do biopsychosocial factors alter pain sensitivity differently in women and men? *Pain*. (2012) 153(3):619–35. doi: 10.1016/j.pain.2011.11.026

63. Rosen S, Ham B, Mogil JS. Sex differences in neuroimmunity and pain. *J Neurosci Res*. (2017) 95(1-2):500–8. doi: 10.1002/jnr.23831

64. Bennett R, Nelson D. Cognitive behavioral therapy for fibromyalgia. *Nat Clin Pract Rheumatol*. (2006) 2(8):416–24. doi: 10.1038/ncprheum0245

65. Ehde DM, Dillworth TM, Turner JA. Cognitive-behavioral therapy for individuals with chronic pain: efficacy, innovations, and directions for research. *Am Psychol*. (2014) 69(2):153–66. doi: 10.1037/a0035747

66. Morley S. Efficacy and effectiveness of cognitive behaviour therapy for chronic pain: progress and some challenges. *Pain*. (2011) 152(3 Suppl):S99–s106. doi: 10.1016/j.pain.2010.10.042

67. Sturgeon JA. Psychological therapies for the management of chronic pain. *Psychol Res Behav Manag*. (2014) 7:115–24. doi: 10.2147/PRBM.S44762

68. Veehof MM, Oskam MJ, Schreurs KMG, Bohlmeijer ET. Acceptance-based interventions for the treatment of chronic pain: a systematic review and meta-analysis. *Pain*. (2011) 152(3):533–42. doi: 10.1016/j.pain.2010.11.002

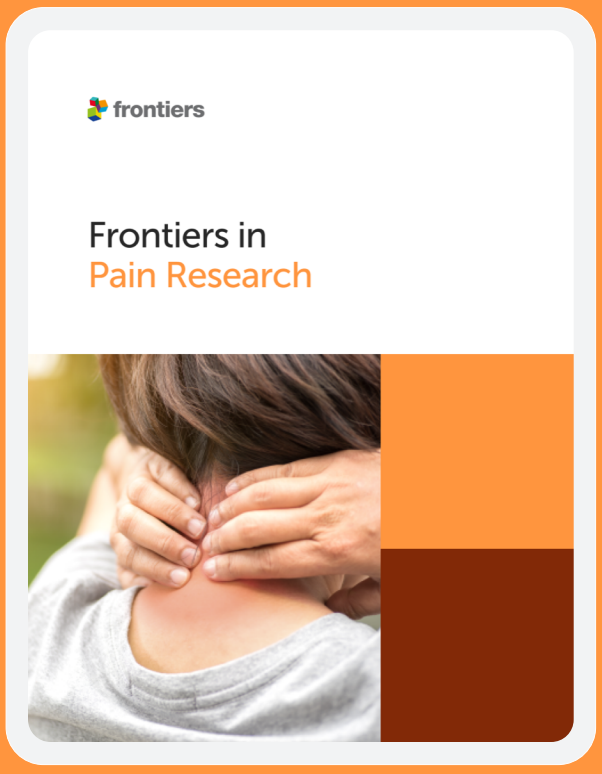
Frontiers in
Pain Research

Investigates new interventions for pain management

An exciting journal in its field, covering mechanisms, treatments, socioeconomics, diagnostics, preventative measures and pain management in a range of medical specialties such as rheumatology and orthopedics.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact