PLOS ONE

RESEARCH ARTICLE

Validity of the cold pressor test and pain sensitivity questionnaire via online self-administration

Matthew H. McIntyre[1], 23andMe Research Team[1¶], Achim Kless[2]*, Peter Hein[2], Mark Field[2], Joyce Y. Tung[1]*

1 23andMe Inc., Mountain View, California, United States of America, 2 Grünenthal Innovation, Grünenthal GmbH, Aachen, Germany

¶ Membership of the 23andMe Research Team is provided in the Acknowledgments.
* Kless@gmx.net (AK); Joyce@23andme.com (JT)

Abstract

To determine the feasibility of complex home-based phenotyping, 1,876 research participants from the customer base of 23andMe completed an online version of a Pain Sensitivity Questionnaire (PSQ) as well as a cold pressor test (CPT) which is used in clinical assessments of pain. Overall our online version of the PSQ performed similarly to the original pen-and-paper version. Construct validity of the PSQ total was demonstrated by internal consistency and consistent discrimination between more and less painful items. Criterion validity was demonstrated by correlation with pain sensitivity as measured by the CPT. Within the same cohort we performed a cold pressor test using a layperson description and household equipment. Comparison with published reports from controlled studies revealed similar distributions of cold pain tolerance times (i.e., time elapsed before removing the hand from the water). Of those who elected to participate in the CPT, a large majority of participants did not report issues with the test procedure or noncompliance with the instructions (97%). We confirmed a large sex difference in CPT thresholds in line with published data, such that women removed their hands from the water at a median of 54.2 seconds, with men lasting for a median time of 82.7 seconds (Kruskal-Wallis statistic, p < 0.0001), but other factors like age or current pain treatment were at most weakly associated, and inconsistently between men and women. We introduce a new paradigm for performing pain testing, called testing@home, that, in the case of cold nociception, showed comparable results to studies conducted under controlled conditions and supervision of a health care professional.

Introduction

Drug development in pain indications has a higher-than-average attrition rate [1]. One possible way to increase success is to identify promising targets based on genetic links between target and disease [2]. However, discovering new genetic links to pain requires large cohorts [3], which may be difficult to obtain using phenotypes obtained under controlled settings (e.g.

OPEN ACCESS

clinical trials). In order to enable collection of larger samples, the goal of this study was to demonstrate whether pain phenotyping in a large cohort of subjects could be accomplished using an internet and home-based approach.

It is well established that individuals experience pain differently [4,5]. A similar sensory input caused by experimental or clinical pain can lead to vastly differing ratings of the pain experience by different subjects, and it is most likely that this is caused by individual differences in both peripheral and central processing of that sensory input. This has significant implications for patients and clinicians who may struggle to establish a treatment that is effective in reducing pain experience (as indicated by pain ratings), but also for drug development where the differences described above introduce another layer of variability, making it more difficult to identify true drug effects. Although the interindividual differences in pain processing are well known, there is only limited data available describing this variability in larger populations. One way of characterizing inter- and intra-individual differences has been as evidence of low rating accuracy [6].

The CPT was developed to measure autonomous responses in the cardiovascular system [7]. The test usually consists of immersion of one hand in ice water for a specified amount of time [8], which induces both pain and a response of the autonomous nervous system. The CPT has been widely adopted as a model for nociceptive pain, and for opioids it is established as a surrogate of clinical efficacy [9]. As such, this test has been used mostly in small populations from clinical trials with two notable exceptions: cohorts from Haifa, Israel [10–12], and unpublished data, and Tromsø, Norway [13], both summarized by Treister et al. [14]. The studies conducted in Haifa included 648 people. The Tromsø Study included 10,486 people enriched for those 40–42 or 60–87 years old. That study found much lower pain sensitivity than seen in the participants from Haifa, in that most participants left their hands in the cold water for longer than 100 seconds. However, they found that participants with chronic pain removed their hands from the water sooner, indicating lower pain tolerance.

Two existing questionnaires, the Central Sensitization Inventory (CSI) [15] and Pain Sensitivity Questionnaire (PSQ) [16], have been designed as instruments for screening patient symptoms like allodynia or hyperalgesia, and whether these are related to central sensitization. Since the CSI Part A has 25 questions that are ranked by categorical values, we have decided to use the PSQ, which is based on fewer but more easily imaginable everyday life situations that are evaluated using a continuous numerical pain scale which fits our data analysis methods.

The PSQ consists of 14 imagined painful situations and 3 non-painful control situations, and subjects are asked to rate their painfulness on a 0–10 numeric rating scale, originally in German [16]. Ruscheweyh et al. showed that the PSQ demonstrated strong internal consistency, supported by high Cronbach's α of both the PSQ-total score and two derived sub-factors labeled as sensitivity to "minor" and "moderate" pain. Moreover, they reported evidence of criterion validity demonstrated by correlation with subjective pain experienced from a range of stimuli, including pinprick, pressure, phasic and tonic heat and cold, and the cold pressor test. However, PSQ-total score did not correlate with response thresholds to any stimulus, including time to withdraw one's hand from cold water in the CPT. A validation in chronic pain patients [16], and a separate English language validation of the PSQ [17] have been published. However, the cohort sizes from which stimulus response correlations with PSQ scores are published are limited (406 subjects in [16], 319 subjects in [18], 136 subjects in [17], 103 [19], 182 [20] and 331 [21]). More recently, Grundström et al. (2019) demonstrated an association between PSQ total score and both temperature and pressure thresholds, not including the CPT, in a sample of 37 women with persistent pelvic pain and 55 control women, though associations were notably stronger in the chronic pain subgroup [22].

Pain sensitivity measured via the PSQ may be associated with a history of chronic pain. In a second study, Ruscheweyh et al. reported significantly elevated PSQ scores in 134 chronic pain patients as compared to 185 healthy controls. A subgroup of 46 chronic pain patients were given experimental pain testing, not including the CPT but including pain ratings, but not thresholds, from another tonic cold stimulus. In that subgroup, there was a strong correlation between PSQ and tonic cold stimulus pain rating [18].

Deep pain phenotyping of subjects is a key prerequisite to identify subgroups for a pain indication.

Owing to high interindividual variability in the perception of pain and response to therapy, deep pain phenotyping of subjects is a key prerequisite to identify subgroups within pain indications. In deep pain phenotyping, details about disease manifestations are gathered in a more individual and finer-grained way for instance by questionnaires [23]. Further, precision medicine approaches need such a better understanding of the precise relationship between genes and phenotype to reveal underlying biological mechanisms. Finally, the discovery of novel subclasses may eventually translate into clinical care [24]. Recent studies have shown progress in the genetic correlations between pain phenotypes and psychiatric traits or using RNA sequence analysis on chondrocytes from osteoarthritis patients [25].

In a laboratory quantitative sensory testing (QST) protocol, subjects may be exposed to multiple types of pain stimulus (heat, cold, pressure, etc.). Applying pain ratings to studies such as genome-wide association studies requires large samples [3]. While large studies of pain sensitivity have been conducted in relatively controlled settings [13,26,27], self-administration allow for large studies to be conduct more quickly and at less expense. Here, we investigate whether online, self-administered versions of the PSQ and CPT demonstrate similar construct and criterion validity as observed in both pen-and-paper and laboratory protocols. We present the results from 1,876 participants asked to perform the cold pressor test (CPT) on themselves at home, via a layperson description, using household equipment, and guided by a web-based workflow. In addition to the CPT, participants self-administered an online version of the Pain Sensitivity Questionnaire (PSQ) and reported about their history of treatment for pain.

Materials and methods

Tools/Measures—web-based phenotyping

23andMe is a direct to consumer personal genomics company with a research platform that allows participants to complete questionnaires for research purposes. We added online versions of the CPT consent and instructions, and questionnaires, to the research area of the 23andMe Personal Genetics Service website (www.23andme.com). Ultimately, 1,876 participants who were at least 20 years old, and consented to participate in the research, completed the questionnaires followed by the CPT. In addition to the PSQ questionnaire, we asked about history of diagnosis with, and treatment for, acute and chronic pain-related conditions. Partly because our pain history questionnaire was unexpectedly long for many participants, and also because many participants were asked to carry out the burdensome cold pressor test protocol after completing the questionnaires, the 1,876 participants who completed the activities reflect substantial attrition from the approximately 10,000 who started the sequence which is certainly a limitation to be considered for future trial designs. Study data collection occurred between June 2017 and February 2018.

Subjects

We recruited participants for both the questionnaires and CPT using email and by participation of research participants on the 23andMe website in the US. Participants provided

informed consent online according to a protocol approved by Ethical & Independent Review Services, a private institutional review board (OHRP/FDA registration number IRB00007807, study number 10044–11), which included separate consent for the CPT. Participants with pain conditions were eligible for this study, though they were excluded from the original validation study of the PSQ by Ruscheweyh et al. [16]. Exclusions for participation are described in Table 1. Participants with self-reported cardiovascular disorders (e.g. high blood pressure, heat diseases, dysrhythmia), a history of Raynaud's phenomenon, any neurological disorders, and/ or pregnancy were not recruited and advised not to participate (Table 1) to minimize the risk of adverse events during the CPT procedure. Especially as no researcher was present in this at-home testing who could provide assistance in such cases, we used this as an additional safety precaution. This resulted in the exclusion of participants with some previously-reported chronic pain conditions, like migraine, but not others, like back pain. Ultimately, of the 1,876 participants, 181 (11.2%) reported being treated in the past 4 months with prescription medication for an acute pain condition, of whom 85 (4.5%) also reported treatment for a chronic pain condition, and another 78 (4.2%) for a chronic pain condition alone.

Those treated for both chronic and acute pain conditions were classified as having acute pain as participants with a chronic pain condition will also have episodes of acute pain [32]. We compared the results of this classification approach with the alternative of considering only treatment for chronic pain itself, regardless of acute pain. We required use of prescription pain medication for classification with acute or chronic pain owing to the high frequency of use of over-the-counter pain medications and their limited mechanisms of action.

Questionnaires

In a sequence of two questionnaires, the first included the English-language version of the PSQ and additional questions about the participant's own memory of painful experiences. The second questionnaire included questions about history of pain conditions and related medications.

Causes of chronic pain included self-reported low back pain; complex regional pain syndrome; joint pain; diabetic neuropathy; endometriosis; cancer; other internal pain not caused by endometriosis or cancer; shingles, cold sores, or herpes; trigeminal neuralgia; migraine; non-migraine headaches that occurred more than half of the days in any given month; pain following amputation; or other pain after injury or surgery that lasted more than 3 months. Acute pain included dental pain, pain after injury or surgery that lasted less than 3 months, or other pain that lasted less than 3 months. Participants were classified as having current acute pain if they had been treated with a prescription pain medication for an acute condition in

**Table 1. Study exclusion criteria, as represented to participants.**

High blood pressure
Heart disease
Dysrhythmia
Any other cardiovascular disorder
History of Raynaud's phenomenon
History of fainting or seizures
History of frostbite
An open cut, sore, or bone fracture on or near your non-dominant hand (the one you do not usually write with)
Any neurological disorder
Are pregnant or think you might be pregnant

https://doi.org/10.1371/journal.pone.0231697.t001

past 4 months. Participants without current acute pain were classified as having current chronic pain if they had been treated with a prescription pain medication for a chronic condition in the past 4 months. Otherwise, they were classified with no current pain condition.

Cold pressor test

In contrast to a controlled clinical or laboratory setting, in our study the CPT was done by the participants themselves at home using a lay description of the procedure (Fig 1). Participants were asked to prepare their own bath of ice water at home. After preparing the bath, the participant was asked to press a button and place her non-dominant hand in the water. Pressing the button started a timer, which was visible to the participant only as a blinking icon without the elapsed time displayed. The participant was asked to press the button again when first feeling pain, and then to press the button a third time after removing her hand from the water. The timer ended at 150s at which point a notification appeared on the screen asking that the participant remove her hand from the water. Participants had two opportunities to report any errors in the test (e.g. distractions), one appearing on the timer pager, and another in the form of a question after the test was complete. Participants were not provided with any guidance about whether to conduct the test alone or with other people present.
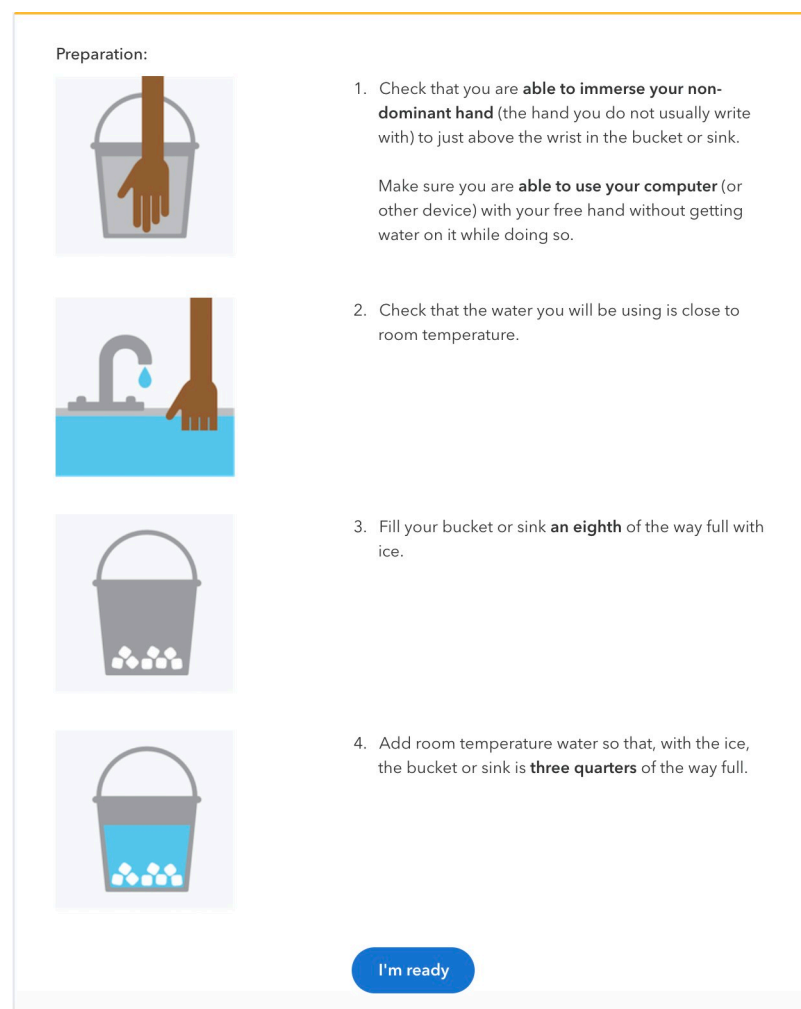
Fig 1. Web-based workflow of cold pressor test.

https://doi.org/10.1371/journal.pone.0231697.g001

Two primary outcomes were assessed: cold pain threshold and cold pain tolerance. Cold pain threshold was the time to the first report of pain and cold tolerance the time to removal of the hand from the water. We additionally asked about maximum pain intensity, but our measure of pain intensity differs from some other laboratory protocols. First, we presented an 11-point visual analog scale immediately after the completion of the task, rather than periodically during the task itself. This was required to minimize the number of simultaneous tasks for participants to perform. Second, we only asked about the maximum pain felt at the end of the test. While at least one large study has measured pain intensity in a similar way [10,11], others score periodic measurements by averaging over all pain intensity measurements and imputing the maximum score 10, for any intervals following the removal of the hand [13]. Studies may also use both methods [28].

Participants with a history of migraine and a number of other chronic conditions that might be directly exacerbated by the activity were not invited to participate in the CPT in order to ensure their safety, while participants with other chronic pain conditions, including chronic back pain, were invited. Moreover, participants reporting neurological or temperature-triggered conditions (e.g. migraine, history of syncope, or Raynaud's phenomenon) or current injuries to their non-dominant hands at the time of recruitment were ineligible to review the supplementary CPT consent document and to participate. The amount of ice in the instructions (1/8 of the container size) would not fully melt at ambient water and air temperatures, and would remain floating at the top of the water. However, if participants did not add enough ice or added warm water and the ice did melt, temperatures would have risen from about 0° at the beginning to a higher temperature at the end of the test. This differs from many laboratory protocols in which refrigeration and water circulation are used to keep the temperature at 2–5° for the duration of the test.

Data analysis

To test the psychometric validity of the online English-language PSQ [17], we compared psychometric properties with those of the original validation of the German-language version of the PSQ [16]. We computed the same measures as described in the original study of the German language version: PSQ-total (all items considered painful), PSQ-minor (the least painful items: 14, 3, 6, 12, 11, 10, and 7, ordered from least to most painful), PSQ-moderate (8, 15, 2, 16, 17, 1, 4 ordered from least to most painful), and calculated Cronbach's α for each measure, a measure of internal consistency.

To further assess construct validity, we conducted factor analysis of the PSQ items. Factor analysis of the PSQ items by Ruscheweyh et al. [16] yielded two factors corresponding to minor and moderate causes of pain. On the basis of this result, they proposed a PSQ-total score (the mean of all painful items) and PSQ-minor and PSQ-moderate subscales (the means of the respective items). Our comparisons with Ruscheweyh et al. [16] include those between (1) item-level means scores, (2) mean scores for the PSQ-total and subscales, (3) varimax factor loading, (4) item correlations with PSQ-total, (5) correlation of CPT pain threshold, tolerance, and intensity with PSQ-total to assess criterion validity of both the PSQ and CPT. We also compare the distribution of CPT pain thresholds with those found in previous studies, including Ruscheweyh et al. [16]. Following the recommendation of Treister et al. [14], we used the Kruskal-Wallis rank test to examine differences between, sex, age, and pain history group in CPT thresholds and pain intensity.

In addition to these validation steps, we also investigated whether participants varied in their reporting accuracy. We defined accuracy as the within-person correlation of PSQ pain ratings with the loadings on the first principal component. Those who rated the items that

were generally considered more painful as more painful, or vice-versa, would have high accuracy scores.

Results

Pain sensitivity questionnaire

Internal consistencies (Cronbach's α) were similar to Ruscheweyh et al. [16] for PSQ total (ours 0.93, theirs 0.92), PSQ minor (ours 0.84, theirs 0.81), and PSQ moderate (ours 0.90, theirs 0.91). Table 2 gives mean scores for the whole group as well as scores stratified by sex, age, and pain history. Our mean PSQ scores were significantly lower than found by Ruscheweyh et al. [16]. We found small sex and age differences, but no differences among those with current treatment for acute pain, for chronic pain only, and without current treatment for pain. Likewise, we found no difference between those with treatment for chronic pain and without chronic pain, regardless of treatment for acute pain (F = 0.1, p = 0.7761). Items for the PSQ-moderate scale (1, 2, 4, 8, 15, 16, and 17) were each rated as more painful than items for the PSQ-minor scale (3, 6, 7, 10, 11, 12, and 14) in both the present study and Ruscheweyh et al. [16] (see Tables 3 and 4). Three items (5, 9, 13) that describe normally non-painful situations have been excluded.

While we found only one principal component with an eigenvalue over 1, the second was close to 1 (7.25 and 0.97), so we compared with the varimax-rotated two-factor solution presented by Ruscheweyh et al. [16]. Varimax rotation results in two factors with a similar structure to that observed by Ruscheweyh et al. [16], as shown in Table 3, but with several notable differences. These two rotated factors explain about 59% of the total variance, compared to

Table 2. Mean PSQ total, minor and moderate scores for the total study population as well as stratified scores for sex, age, and current pain, and F-tests of group differences.

	Count	PSQ total	PSQ minor	PSQ moderate
		mean (SD)	mean (SD)	mean (SD)
Ruscheweyh et al. (2009)	354	3.6 (1.2)	2.5 (1.1)	4.7 (1.6)
23andMe/Grünenthal online cohort Total	1876	3.1 (1.4)	2.2 (1.2)	4.0 (1.6)
		F = 10.2	F = 7.1	F = 14.8
		p < 0.0001	p < 0.0001	p < 0.0001
SEX: Male	679	3.0 (1.3)	2.1 (1.1)	3.9 (1.6)
SEX: Female	1197	3.2 (1.4)	2.3 (1.3)	4.0 (1.7)
		F = 5.5	F = 10.0	F = 2.4
		p = 0.0188	p = 0.0016	p = 0.1248
AGE: 20–29	356	3.3 (1.3)	2.4 (1.2)	4.2 (1.5)
AGE: 30–39	426	3.1 (1.3)	2.2 (1.2)	4.0 (1.6)
AGE: 40–49	297	3.0 (1.4)	2.1 (1.3)	3.9 (1.6)
AGE: 50–59	285	3.1 (1.4)	2.2 (1.3)	4.1 (1.7)
AGE: 60–69	332	2.9 (1.3)	2.0 (1.2)	3.8 (1.7)
AGE: >70	180	3.1 (1.4)	2.1 (1.2)	4.1 (1.7)
		F = 3.2	F = 5.3	F = 1.9
		p = 0.0070	p = 0.0001	p = 0.0986
PAIN: Acute	181	3.2 (1.5)	2.2 (1.3)	4.1 (1.8)
PAIN: Chronic	78	3.2 (1.5)	2.3 (1.4)	4.0 (1.7)
PAIN: None	1617	3.1 (1.3)	2.2 (1.2)	4.0 (1.6)
		F = 0.5	F = 0.4	F = 0.7
		p = 0.6058	p = 0.6636	p = 0.5027

Table 3. Ratings for individual PSQ items and factor structures observed in online sample vs. original sample.

Item	Mean (SD)	Factor loading		Factor loading		Correlation with PSQ total	Correlation with PSQ total Ruscheweyh et al. [12]
		Online PSQ-E		Ruscheweyh et al. [12] Table 2		Online	
				n = 354			
		n = 1,876					
		Moderate (2)	Minor (1)	Moderate (1)	Minor (2)		
1 bump shin	**4.3 (2.0)**	**0.83**	**0.19**	**0.77**	**0.12**	**0.66**	**0.62**
2 burn tongue	**3.4 (1.9)**	**0.80**	**0.24**	**0.78**	**0.21**	**0.68**	**0.69**
3 sore muscles	2.3 (1.7)	0.62	0.33	0.12	0.78	0.61	0.51
4 trap finger	**4.5 (2.1)**	**0.77**	**0.34**	**0.77**	**0.28**	**0.75**	**0.73**
6 sunburn shoulders	1.8 (1.5)	0.33	0.55	0.08	0.77	0.57	0.47
7 graze knee	3.1 (1.8)	0.49	0.58	0.42	0.53	0.70	0.59
8 bite cheek	**3.4 (2.0)**	**0.58**	**0.54**	**0.72**	**0.36**	**0.77**	**0.74**
10 cut finger	2.5 (1.8)	0.25	0.76	0.41	0.65	0.70	0.67
11 prick finger	2.0 (1.5)	0.30	0.76	0.48	0.56	0.74	0.66
12 hands in snow	2.6 (2.1)	0.19	0.71	0.32	0.51	0.65	0.51
14 strong hand shake	1.1 (1.5)	0.17	0.53	0.25	0.53	0.53	0.47
15 hot handle	**4.6 (2.1)**	**0.43**	**0.63**	**0.71**	**0.35**	**0.72**	**0.73**
16 crush foot	**4.2 (2.1)**	**0.53**	**0.58**	**0.77**	**0.30**	**0.77**	**0.75**
17 funny bone	3.5 (2.1)	0.56	0.52	0.79	0.31	0.74	0.76

Items used to calculate the PSQ moderate subscale are indicated in **bold**. Others comprise the moderate subscale.

https://doi.org/10.1371/journal.pone.0231697.t003

55% in Ruscheweyh et al. [16]. Our Factor 1 loads mostly minor causes of pain, similar to Factor 2 in Ruscheweyh et al. [16] and our Factor 2 loads more moderate causes of pain. The biggest difference is seen for item 3 ("Imagine your muscles are slightly sore as the result of

Table 4. Imagined situations for pain ratings used in the questionnaire (Ruscheweyh et al. [16]).

1	Imagine you bump your shin badly on a hard edge, for example, on the edge of a glass coffee table.
2	Imagine you burn your tongue on a very hot drink.
3	Imagine your muscles are slightly sore as the result of physical activity.
4	Imagine you trap your finger in a drawer.
5	Imagine you take a shower with lukewarm water.
6	Imagine you have mild sunburn on your shoulders.
7	Imagine you grazed your knee falling off your bicycle.
8	Imagine you accidentally bite your tongue or cheek badly while eating.
9	Imagine walking across a cool tiled floor with bare feet.
10	Imagine you have a minor cut on your finger and inadvertently get lemon juice in the wound.
11	Imagine you prick your fingertip on the thorn of a rose.
12	Imagine you stick your bare hands in the snow for a couple of minutes or bring your hands in contact with snow for some time, for example, while making snowballs.
13	Imagine you shake hands with someone who has a normal grip.
14	Imagine you shake hands with someone who has a very strong grip.
15	Imagine you pick up a hot pot by inadvertently grabbing its equally hot handles.
16	Imagine you are wearing sandals and someone with heavy boots steps on your foot.
17	Imagine you bump your elbow on the edge of a table ("funny bone").

https://doi.org/10.1371/journal.pone.0231697.t004

physical activity"), which loaded on the minor pain factor in Ruscheweyh et al. [16], but more heavily load on the moderate pain factor in our results, despite being rated as less painful. Conversely, items 15 ("Imagine you pick up a hot pot by inadvertently grabbing its equally hot handles") and 16 ("Imagine you wear sandals and someone with heavy boots steps on your foot") more heavily loaded the moderate pain factor in Ruscheweyh et al. [16], but more heavily load the minor pain factor in our results, despite being rated as relatively painful.

Items correlated similarly with the PSQ total score in the two studies (correlation of correlations = 0.92, see Table 3). Pain attributed to a strong handshake (#14) was the weakest correlate of total PSQ score in both studies, and, as a non-painful item, is also not a component of the PSQ scores. The three strongest correlates with PSQ total score in both studies were pain attributed to hitting one's "funny bone" (#17), to having one's foot stepped on (#16), and to biting one's cheek (#8).

Cold pressor test

Cold pain tolerance distribution is presented in Fig 2. This figure also shows results from two other large cohorts (data from [14]), obtained in a controlled laboratory setting with a standard methodology (e.g. use of a circulating, refrigerated water bath kept at a constant temperature). Cold pressor pain thresholds differed substantially between the Haifa and Tromsø cohorts and our cohort threshold fell between these two as shown in Fig 3.

CPT tolerance time distributions differed substantially by sex (Kruskal-Wallis statistic = 71, df = 1, p<0.0001, Fig 4), but not by age (Kruskal-Wallis statistic = 13, df = 5, p = 0.0241, Fig 5) or pain history (Kruskal-Wallis statistics = 4, df = 2, p = 0.1591, Fig 6). Women report a median tolerance of 54.2 seconds (IQR 30.4–116.5), which was 31.0 seconds earlier than the median of 82.7 seconds (IQR 43.6–150.0) reported by men. With respect to pain history, comparing those taking medication for a chronic pain condition to all others also showed no significant difference (Kruskal-Wallis statistic = 2, df = 1, p = 0.1318) in CPT tolerance.
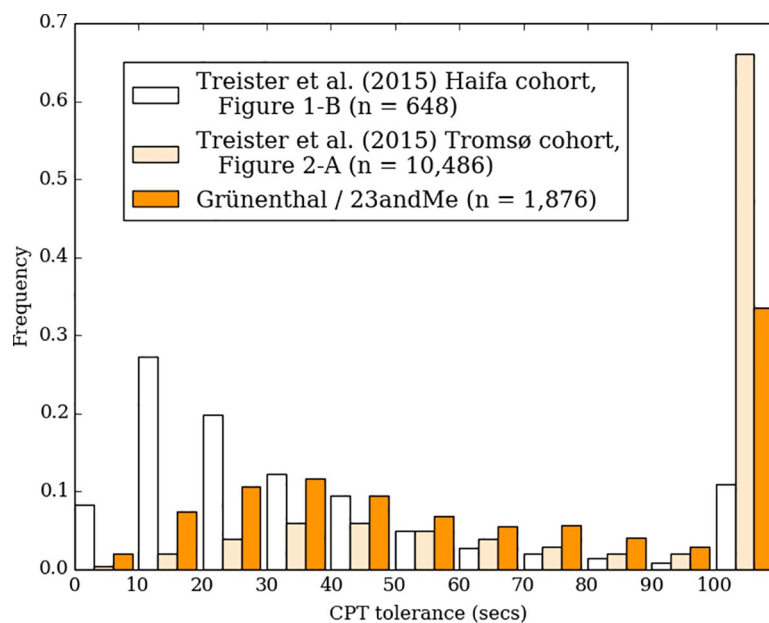
Fig 2. Comparison of the pain tolerance time (time to withdrawal of hand from cold water) of the CPT in our study compared to the data from two other large cohorts (as reviewed by Treister et al. [14]).

https://doi.org/10.1371/journal.pone.0231697.g002
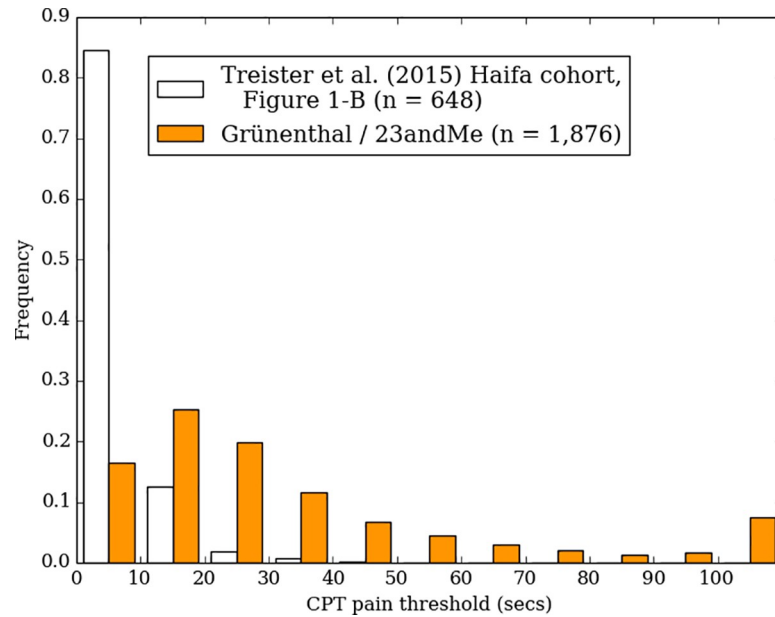
Fig 3. Comparison of the pain threshold time (time to initial report of pain) of the CPT in our study compared to the data from two other large cohorts (as reviewed by Treister et al. [14]).

https://doi.org/10.1371/journal.pone.0231697.g003

PSQ total score was somewhat less correlated with retrospective pain intensity rating just after the CPT (r = 0.30 [95% CI: 0.26, 0.34], Spearman rho = 0.30) than found by Ruscheweyh et al. [16] who asked about pain intensity during the task (r = 0.56), but more similar to the correlation found in the validation of the Norwegian-language PSQ (r = 0.36). [21] Conversely, Ruscheweyh et al. [16] found no significant correlation of PSQ total with CPT pain threshold (r = 0.03, p = 0.86, n = 47), whereas we found small but significant correlations with both time
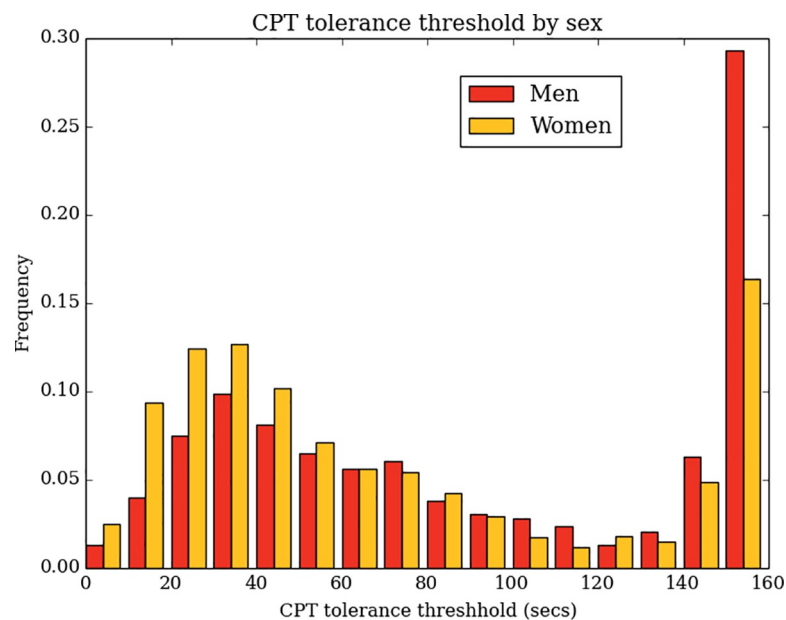
Fig 4. Distributions of CPT tolerance times by sex.

https://doi.org/10.1371/journal.pone.0231697.g004
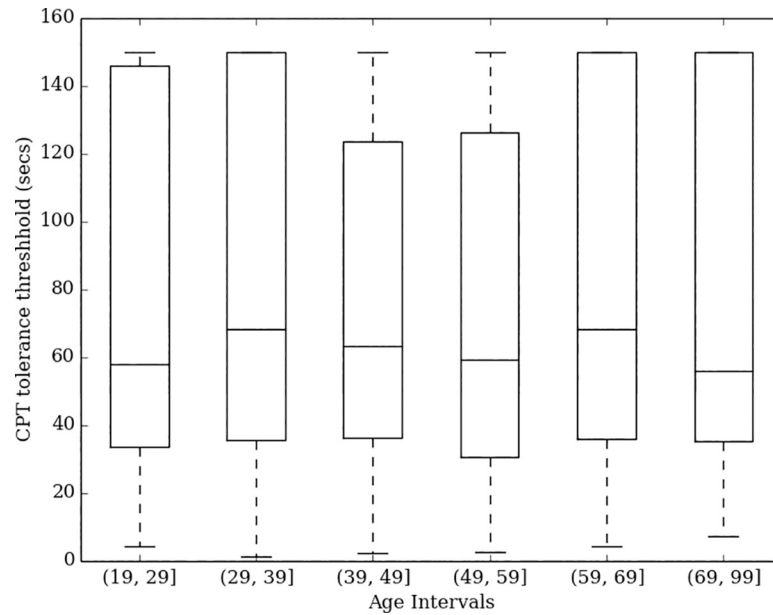
Fig 5. Box plot of CPT tolerance times by age.

to the first report of pain in the CPT (r = -0.14 [-0.19, -0.09], Spearman rho = -0.15) and time to withdrawal of the hand, or "tolerance", (r = -0.22 [-0.27, -0.17], Spearman rho = -0.22). Ruscheweyh et al. [16] found no significant correlation of PSQ total with CPT pain threshold (r = 0.03, p = 0.86, n = 47), whereas the validation of the Norwegian-language PSQ [21] found a somewhat stronger correlation (r = -0.30, p<0.05, n = 48).

469 participants (25%) reported maximum, retrospective pain intensities during the CPT below 3 on a scale of 0–10, where 0 is no pain at all and 10 is the worst pain imaginable. While we hypothesized that these participants might have prepared the CPT test incorrectly, they
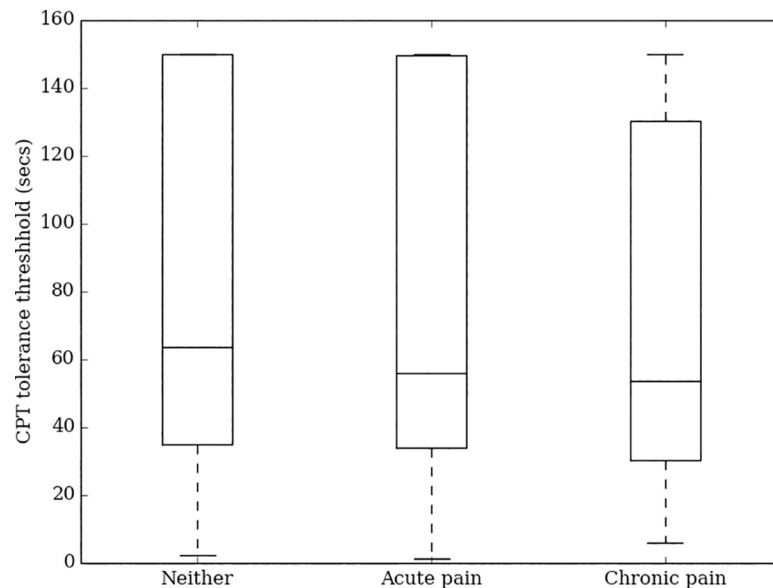
Fig 6. Box plot of CPT tolerance times by current pain status.
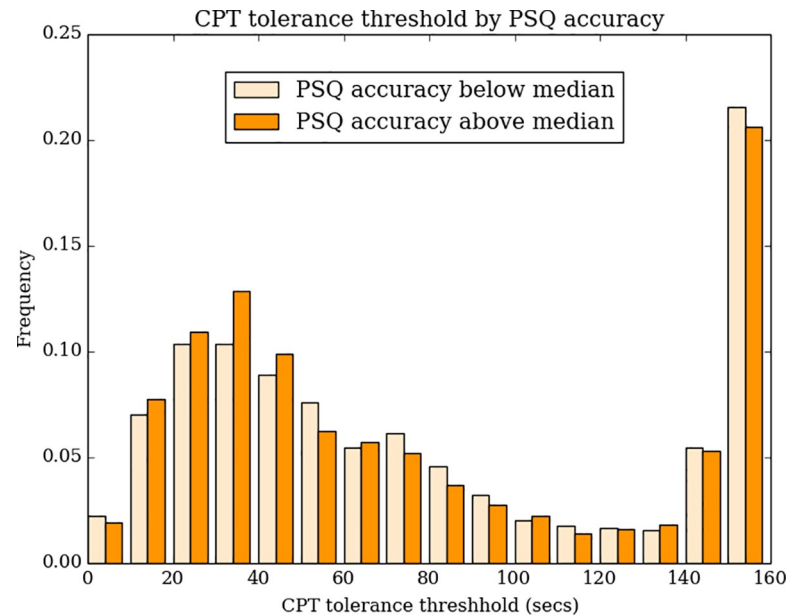
Fig 7. Distributions of CPT tolerance times by PSQ rating accuracy above or below the median.

also reported lower PSQ total score (mean 2.55, SD 1.19 among those with CPT pain rating < = 3.0, mean 3.29, SD 1.35 among those with pain rating > 3.0, t = 11.3, df = 900.5, p<0.0001). Removing these participants from the analysis did not substantially change associations between the CPT and PSQ total score. The largest change after removal of these participants was a decrease in the correlation between CPT pain threshold and PSQ total score (from r = -0.14 [-0.19, -0.09] to r = -0.10 [-0.15, -0.05]).

Accuracy.   The 14 painful items of the PSQ can be ordered and weighted according to their painfulness. We therefore analyzed how accurately participants were in assigning pain scores to the individual items. Within-person correlations between PSQ item ratings and loadings on the first principal component had a median of 0.62 (IQR: 0.31–0.81), with a long-left tail stretching into negative correlations. Accuracy scores showed small but significant correlations with both PSQ-total score (Spearman r = 0.14, p<0.0001) and CPT tolerance (Spearman r = -0.08, p = 0.0002), suggesting that pain sensitivity and accuracy of pain rating related (Fig 7).

Conclusions

The primary goal of this study was to assess the quality of internet-mediated, self-administered pain sensitivity measurements, the PSQ and CPT, in a large population and an uncontrolled, at-home setting, relative to published studies that used laboratory modalities. We also investigated a secondary question of whether participants vary in the accuracy of their pain assessments.

The online PSQ performed differently, in some respects, from the original German-language version and sample, but similarly to other versions and samples. The average PSQ-total score that we observed (3.3) was somewhat lower than that measured in small studies of several translations and in samples with and without chronic pain conditions [17,18,29,30], but a population study of 4,979 German-speaking Italians [27] also observed an average pain score of 3.3.

While we find strong psychometric support for the PSQ-total score, we show only weak replication of distinct minor and moderate PSQ scales, relative to a single PSQ scale. Specifically, several items loaded most heavily on the incorrect factor. In this study, we see stronger overall evidence for a one-factor than a two-factor solution of the PSQ, consistent with results from the large study of the German-speaking Italians [27]. The PSQ factor structure that we observed was similar to that seen for the Polish and French-language versions of the PSQ, which showed weak distinction between the minor and moderate pain factors, particularly for questions 3, 6, 7, 8, 16, and 17 [29,30]. The Polish-language version was validated in a sample of 161 lower back pain patients, and the French-language version in two samples, one of 146 pre-surgical patients and the other of 85 health controls. The factor structure for the English-language PSQ has not been reported previously [17]. With the exception of the two-factor structure described in the original validation study, we replicate the psychometric findings of these previous studies.

Among the participants who completed both the PSQ and an online self-administration of the CPT, the distribution of CPT scores was largely within the rather broad range of those found in laboratory studies, with the exception of more frequent intermodal thresholds in the online cohort. Whereas Ruscheweyh et al. [16] found no significant association between PSQ and CPT thresholds, we found small but significant negative correlations between PSQ score and both time to report of pain and to removal of the hand. The original validation study may have been underpowered to detect these associations. Power to detect the correlation of 0.22 in a sample of 47 is only 32%. Alternatively, participants who report lower pain sensitivity might over-report their pain threshold in the online design, but not in a laboratory design. Future studies should directly compare CPT performance in laboratory and self-administered settings.

These results support the notion that, when properly instructed, subjects are capable of self-administering the CPT in the absence of trained staff providing individual instructions. Earlier studies have already shown that web-based phenotyping can produce a phenotype similar enough to physician-obtained phenotypes to yield similar results in genome-wide association studies [31]. Our study adds to this in showing that web-based, self-phenotyping appears to be a valid approach not only when considering questionnaires, but that some test procedures can be followed to yield results similar to results from laboratory settings. Since we are using a web-based approach we do not have a familiarization session at the beginning which is usually done in a supervised laboratory test like QST [32]. Instead we implemented a dry test before the actual time recordings started so that the participants were prepared to use the web interface and knew about the upcoming sequential steps.

In addition to supporting the validity of subjective and self-administered measures of pain sensitivity, our results also suggest that individuals differ in their ability to precisely and accurately rate pain, the latter measured by consistency with the observed factor structure, which aligns with other recent research [33,34]. However, apparent differences might also be explained by, for example, differences in general attentiveness or conscientiousness, rather than pain rating ability as such. Accuracy scores had a long-left tail stretching into negative scores (i.e. scoring less painful items as more painful). Such a left tail is also observed in measures of person fit to item response theory models of the participant's knowledge of the correct answers to a set of questions. These have been interpreted to reflect, for example, lack of attention to the task, rather than lack of ability, which is thought to have a more symmetric distribution [35]. Factors like attentiveness are difficult to assess in a non-supervised setting like the one used for this study.

A limitation of the study is that we did not directly compare the results of self-administration of the CPT to results from a laboratory test within the same cohort. As we recruited

participants via an online platform throughout the United States, it was impractical for us to repeat the same study with the same cohort under laboratory conditions. Instead, we compared the results of self-administration, in aggregate, to results reported from laboratory tests. For the critical measures of pain sensitivity, including cold pressor test tolerance, our results are similar to results obtained in controlled laboratory settings. Because of self-administration at home we had to adapt precise CPT protocol descriptions like temperature to simpler terms like room temperature. This also includes the amount of ice, which we described by $1/8^{th}$ of the water volume. Overall, this likely led to considerable variability in the actual stimulus that was received by each participant but allowed us to conduct a web-based instruction workflow to self-administer a CPT in the absence of trained staff. Variation in the how participants prepare the test apparatus, for example, the temperature of the water [36], certainly affects the reported pain thresholds. We anticipate future studies directly comparing online and lab implementations.

Population-based, multi-omics studies will grow in the future and therefore a deep phenotypic characterization of datasets is key to deriving maximum amount of insight out of these efforts [37]. Our approach of at-home testing and questionnaires is a first attempt to fill this gap. An online approach, by broadening participation in multi-omics studies, might help to further fill the gap of understanding pain phenotype subgroups, and discovering new pain-related biological pathways.

Upcoming investigations will include the identification of patients whose pain perception is different from the average, especially those able to discriminate better between more and less painful stimuli. Furthermore, we will continue to investigate additional pain testing@home possibilities and will identify pain relevant genes and pathways in order to derive targets using genetic association studies to relevant pain measures. This may allow us to adjust pain scales to standardize individual responses and to identify groups of patients for precision medicine approaches.

Supporting information

S1 Fig. High resolution pictures of the web-based version of the CPT in addition to Fig 1. (PDF)

Acknowledgments

Author Contributions

Conceptualization: Achim Kless, Peter Hein.

Methodology: Matthew H. McIntyre, Peter Hein.

Supervision: Achim Kless, Joyce Y. Tung.

Writing – original draft: Matthew H. McIntyre, Achim Kless, Peter Hein.

Writing – review & editing: Mark Field.

References

1. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov. 2004; 3: 711–715. https://doi.org/10.1038/nrd1470 PMID: 15286737

2. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, et al. The support of human genetic evidence for approved drug indications. Nat Genet. 2015; 47: 856–860. https://doi.org/10.1038/ng.3314 PMID: 26121088

3. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 years of GWAS discovery: Biology, function, and translation. Am J Hum Genet. 2017; 101: 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005 PMID: 28686856

4. Coghill RC, McHaffie JG, Yen Y-F. Neural correlates of interindividual differences in the subjective experience of pain. Proc Natl Acad Sci U S A. 2003; 100: 8538–8542. https://doi.org/10.1073/pnas.1430684100 PMID: 12824463

5. Fillingim RB. Individual differences in pain responses. Curr Rheumatol Rep. 2005; 7: 342–347. https://doi.org/10.1007/s11926-005-0018-7 PMID: 16174481

6. Krebs EE, Carey TS, Weinberger M. Accuracy of the pain numeric rating scale as a screening test in primary care. J Gen Intern Med. 2007; 22: 1453–1458. https://doi.org/10.1007/s11606-007-0321-2 PMID: 17668269

7. Hines EA, Brown GE. Standard stimulus for measuring vasomotor reactions. Its application in the study of hypertension. Proc Staff Meet Mayo Clin. 1932; 7: 332–335.

8. Modir JG, Wallace MS. Human experimental pain models 2: the cold pressor model. Methods Mol Biol Clifton NJ. 2010; 617: 165–168. https://doi.org/10.1007/978-1-60327-323-7_13 PMID: 20336421

9. Staahl C, Olesen AE, Andresen T, Arendt-Nielsen L, Drewes AM. Assessing analgesic actions of opioids by experimental pain models in healthy volunteers—an updated review. Br J Clin Pharmacol. 2009; 68: 149–168. https://doi.org/10.1111/j.1365-2125.2009.03456.x PMID: 19694733

10. Pud D, Eisenberg E, Sprecher E, Rogowski Z, Yarnitsky D. The tridimensional personality theory and pain: harm avoidance and reward dependence traits correlate with pain perception in healthy volunteers. Eur J Pain Lond Engl. 2004; 8: 31–38. https://doi.org/10.1016/S1090-3801(03)00065-X

11. Pud D, Golan Y, Pesta R. Hand dominancy—a feature affecting sensitivity to pain. Neurosci Lett. 2009; 467: 237–240. https://doi.org/10.1016/j.neulet.2009.10.048 PMID: 19853018

12. Treister R, Pud D, Ebstein RP, Eisenberg E. Dopamine transporter genotype dependent effects of apomorphine on cold pain tolerance in healthy volunteers. PloS One. 2013; 8: e63808. https://doi.org/10.1371/journal.pone.0063808 PMID: 23704939

13. Johansen A, Schirmer H, Stubhaug A, Nielsen CS. Persistent post-surgical pain and experimental pain sensitivity in the Tromso study: comorbid pain matters. Pain. 2014; 155: 341–348. https://doi.org/10.1016/j.pain.2013.10.013 PMID: 24145207

14. Treister R, Nielsen CS, Stubhaug A, Farrar JT, Pud D, Sawilowsky S, et al. Experimental comparison of parametric versus nonparametric analyses of data from the cold pressor test. J Pain Off J Am Pain Soc. 2015; 16: 537–548. https://doi.org/10.1016/j.jpain.2015.03.001 PMID: 25801300

15. Mayer TG, Neblett R, Cohen H, Howard KJ, Choi YH, Williams MJ, et al. The development and psychometric validation of the central sensitization inventory. Pain Pract Off J World Inst Pain. 2012; 12: 276–285. https://doi.org/10.1111/j.1533-2500.2011.00493.x PMID: 21951710

16. Ruscheweyh R, Marziniak M, Stumpenhorst F, Reinholz J, Knecht S. Pain sensitivity can be assessed by self-rating: Development and validation of the Pain Sensitivity Questionnaire. Pain. 2009; 146: 65–74. https://doi.org/10.1016/j.pain.2009.06.020 PMID: 19665301

17. Sellers AB, Ruscheweyh R, Kelley BJ, Ness TJ, Vetter TR. Validation of the English language pain sensitivity questionnaire. Reg Anesth Pain Med. 2013; 38: 508–514. https://doi.org/10.1097/AAP.0000000000000007 PMID: 24141873

18. Ruscheweyh R, Verneuer B, Dany K, Marziniak M, Wolowski A, Colak-Ekici R, et al. Validation of the pain sensitivity questionnaire in chronic pain patients. Pain. 2012; 153: 1210–1218. https://doi.org/10.1016/j.pain.2012.02.025 PMID: 22541722

19. Coronado RA, George SZ. The Central Sensitization Inventory and Pain Sensitivity Questionnaire: An exploration of construct validity and associations with widespread pain sensitivity among individuals with shoulder pain. Musculoskelet Sci Pract. 2018; 36: 61–67. https://doi.org/10.1016/j.msksp.2018.04.009 PMID: 29751194

20. Quan X, Fong DYT, Leung AYM, Liao Q, Ruscheweyh R, Chau PH. Validation of the Mandarin Chinese Version of the Pain Sensitivity Questionnaire. Pain Pract Off J World Inst Pain. 2017. https://doi.org/10.1111/papr.12587 PMID: 28422444

21. Valeberg BT, Pedersen LM, Girotto V, Christensen VL, Stubhaug A. Validation of the Norwegian Pain Sensitivity Questionnaire. J Pain Res. 2017; 10: 1137–1142. https://doi.org/10.2147/JPR.S129540 PMID: 28553134

22. Grundstrom H, Larsson B, Arendt-Nielsen L, Gerdle B, Kjolhede P. Associations between pain thresholds for heat, cold and pressure, and Pain Sensitivity Questionnaire scores in healthy women and in women with persistent pelvic pain. Eur J Pain Lond Engl. 2019; 23: 1631–1639. https://doi.org/10.1002/ejp.1439 PMID: 31192501

23. Edwards RR, Dworkin RH, Turk DC, Angst MS, Dionne R, Freeman R, et al. Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. Pain. 2016; 157: 1851–1871. https://doi.org/10.1097/j.pain.0000000000000602 PMID: 27152687

24. Meng W, Adams MJ, Reel P, Rajendrakumar A, Huang Y, Deary IJ, et al. Genetic correlations between pain phenotypes and depression and neuroticism. Eur J Hum Genet EJHG. 2020; 28: 358–366. https://doi.org/10.1038/s41431-019-0530-2 PMID: 31659249

25. Ji Q, Zheng Y, Zhang G, Hu Y, Fan X, Hou Y, et al. Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. Ann Rheum Dis. 2019; 78: 100–110. https://doi.org/10.1136/annrheumdis-2017-212863 PMID: 30026257

26. Hermesdorf M, Berger K, Baune BT, Wellmann J, Ruscheweyh R, Wersching H. Pain sensitivity in patients with major depression: Differential effect of pain sensitivity measures, somatic cofactors, and disease characteristics. J Pain Off J Am Pain Soc. 2016; 17: 606–616. https://doi.org/10.1016/j.jpain.2016.01.474 PMID: 26867484

27. Melotti R, Ruscheweyh R, Pramstaller PP, Hicks AA, Pattaro C. Structural consistency of the Pain Sensitivity Questionnaire in the Cooperative Health Research In South Tyrol (CHRIS) population-based study. J Pain. 2018; 19: 1424–1434. https://doi.org/10.1016/j.jpain.2018.06.007 PMID: 30017960

28. von Baeyer CL, Piira T, Chambers CT, Trapanotto M, Zeltzer LK. Guidelines for the cold pressor task as an experimental pain stimulus for use with children. J Pain. 2005; 6: 218–227. https://doi.org/10.1016/j.jpain.2005.01.349 PMID: 15820909

29. Dualé C, Bauer U, Storme B, Eljezi V, Ruscheweyh R, Eschalier S, et al. Transcultural adaptation and French validation of the Pain Sensitivity Questionnaire. Can J Anaesth J Can Anesth. 2019. https://doi.org/10.1007/s12630-019-01377-w PMID: 31020630

30. Latka D, Miekisiak G, Kozlowska K, Olbrycht T, Chowaniec J, Latka K, et al. Translation, validation, and cross-cultural adaptation of the Polish version of the pain sensitivity questionnaire. J Pain Res. 2019; 12: 969–973. https://doi.org/10.2147/JPR.S189427 PMID: 30936737

31. Tung JY, Do CB, Hinds DA, Kiefer AK, Macpherson JM, Chowdry AB, et al. Efficient Replication of over 180 Genetic Associations with Self-Reported Medical Data. PLoS ONE. 2011; 6: e23473. https://doi.org/10.1371/journal.pone.0023473 PMID: 21858135

32. Perry BG, Bear TLK, Lucas SJE, Mundel T. Mild dehydration modifies the cerebrovascular response to the cold pressor test. Exp Physiol. 2016; 101: 135–142. https://doi.org/10.1113/EP085449 PMID: 26374269

33. Treister R, Eaton TA, Trudeau JJ, Elder H, Katz NP. Development and preliminary validation of the focused analgesia selection test to identify accurate pain reporters. J Pain Res. 2017; 10: 319–326. https://doi.org/10.2147/JPR.S121455 PMID: 28243138

34. Treister R, Lawal OD, Shecter JD, Khurana N, Bothmer J, Field M, et al. Accurate pain reporting training diminishes the placebo response: Results from a randomised, double-blind, crossover trial. PloS One. 2018; 13: e0197844. https://doi.org/10.1371/journal.pone.0197844 PMID: 29795665

35. Felt JM, Castaneda R, Tiemensma J, Depaoli S. Using person fit statistics to detect outliers in survey research. Front Psychol. 2017; 8: 863. https://doi.org/10.3389/fpsyg.2017.00863 PMID: 28603512

36. Mitchell LA, MacDonald RAR, Brodie EE. Temperature and the cold pressor test. J Pain Off J Am Pain Soc. 2004; 5: 233–237. https://doi.org/10.1016/j.jpain.2004.03.004 PMID: 15162346

37. Robinson PN, Mungall CJ, Haendel M. Capturing phenotypes for precision medicine. Cold Spring Harb Mol Case Stud. 2015; 1: a000372. https://doi.org/10.1101/mcs.a000372 PMID: 27148566