General section
Research Paper

OPEN

PAIN REPORTS®

A tonic heat test stimulus yields a larger and more reliable conditioned pain modulation effect compared to a phasic heat test stimulus

Marie Udnesseter Lie[a,b,*], Dagfinn Matre[c], Per Hansson[d,e], Audun Stubhaug[e,f], John-Anker Zwart[a,f,g], Kristian Bernhard Nilsen[a,c,g,h]

Abstract

Introduction: The interest in conditioned pain modulation (CPM) as a clinical tool for measuring endogenously induced analgesia is increasing. There is, however, large variation in the CPM methodology, hindering comparison of results across studies. Research comparing different CPM protocols is needed in order to obtain a standardized test paradigm.

Objectives: The aim of the study was to assess whether a protocol with phasic heat stimuli as test-stimulus is preferable to a protocol with tonic heat stimulus as test-stimulus.

Methods: In this experimental crossover study, we compared 2 CPM protocols with different test-stimulus; one with tonic test-stimulus (constant heat stimulus of 120-second duration) and one with phasic test-stimuli (3 heat stimulations of 5 seconds duration separated by 10 seconds). Conditioning stimulus was a 7°C water bath in parallel with the test-stimulus. Twenty-four healthy volunteers were assessed on 2 occasions with minimum 1 week apart. Differences in the magnitude and test–retest reliability of the CPM effect in the 2 protocols were investigated with repeated-measures analysis of variance and by relative and absolute reliability indices.

Results: The protocol with tonic test-stimulus induced a significantly larger CPM effect compared to the protocol with phasic test-stimuli ($P < 0.001$). Fair and good relative reliability was found with the phasic and tonic test-stimuli, respectively. Absolute reliability indices showed large intraindividual variability from session to session in both protocols.

Conclusion: The present study shows that a CPM protocol with a tonic test-stimulus is preferable to a protocol with phasic test-stimuli. However, we emphasize that one should be cautious to use the CPM effect as biomarker or in clinical decision making on an individual level due to large intraindividual variability.

Keywords: Conditioned pain modulation, Experimental pain, Reliability

1. Introduction

Pain experiences can be altered by the central nervous system through endogenous modulatory systems, including through activity set up by the nociceptive system itself.[2,14,32] Although the theoretical framework of pathways affected by such activity seems by many authors to be limited to one fairly well researched circuitry, ie, "diffuse noxious inhibitory controls" (DNIC), this is just one of several possible competing systems that may facilitate or inhibit the nociceptive system at different central nervous system levels. Conditioned pain modulation (CPM) is a test introduced to survey the net inhibitory and excitatory effect of activated nociceptive pathways on the pain sensitivity in a body part other

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] Research and Communication Unit for Musculoskeletal Health (FORMI), Oslo University Hospital, Oslo, Norway, [b] Faculty of Health Science, Oslo and Akershus University College of Applied Sciences, Oslo, Norway, [c] Department of Work Psychology and Physiology, National Institute of Occupational Health, Oslo, Norway, [d] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden, [e] Department of Pain Management and Research, Division of Emergencies and Critical Care, Oslo University Hospital, Oslo, Norway, [f] Faculty of Medicine, University of Oslo, Oslo, Norway, [g] Department of Neurology, Oslo University Hospital, Oslo, Norway, [h] Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway

**Corresponding author. Address: FORMI, Oslo University Hospital, Ullevål, PO 4956, Nydalen 0424, Oslo, Norway. Tel.: +47 22117768; fax: +47 23016150. E-mail address: maulie@ous-hf.no (M.U. Lie).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PR9 2 (2017) e626

http://dx.doi.org/10.1097/PR9.0000000000000626

than the one being conditioned. In this test paradigm, the perceived pain intensity of a painful test-stimulus is measured once and thereafter again during or immediately after a painful conditioning stimulus in a remote body area. A reduction in the pain perception of a test-stimulus or an increase in a pain threshold induced by a test-stimulus in response to a conditioning stimulus is considered as a net inhibitory CPM effect.[26]

A meta-analysis has shown that individuals with chronic pain disorders have impaired CPM effect compared to healthy controls.[15] Studies have also shown that pain-free individuals with impaired CPM effect have higher risk of developing chronic pain after surgery than individuals with a stronger CPM effect.[31,36] These findings suggest that impaired CPM effect may be a trait that is a risk factor for the development of chronic pain. However, there is large variation in the CPM methodology, which makes it difficult to reach firm conclusions.[35] A recent systematic review suggests that CPM is a reliable measure, but the degree of reliability depends among other factors on stimulation parameters.[13] A standardized and evidence-based CPM protocol is needed.[35]

One of the commonly used test-stimulus in the CPM paradigm is contact heat stimulation.[26] Both tonic heat test-stimulus (eg, Refs. 11, 28) and phasic heat test-stimuli (eg, Refs. 29, 30) are in use, but the 2 have not been compared to establish the optimal modality of heat test-stimulus in CPM testing. Brief phasic heat stimulations may be more reliable,[10] more efficient, and convenient in a clinical setting than tonic heat stimulations. Thus, the main aim of the present study was to assess whether a protocol with phasic heat stimuli is preferable to a protocol with tonic heat stimulus. Secondly, the present study aimed to identify the optimal stimulation time or number of stimulations in the tonic and phasic test-stimuli, respectively. Assuming CPM is a trait, the optimal protocol should be based on methods that evoke statistically significant and reliable CPM effects.

2. Methods

2.1. Subjects

Self-reported healthy men and women aged 18 to 45 years were recruited by advertisement at local hospitals and colleges or universities in Oslo, Norway. Exclusion criteria were presence of chronic pain, somatic or psychiatric disease, headache for more than 2 days a month, hypertension (>140/90 mm Hg), pregnancy, acquaintances with the experimenter, and regular use of medication (including nonprescription pain killers), except oral contraceptives. Subjects were requested not to work night shifts 48 hours before the experiment, not to consume alcohol or pain killers 24 hours before the experiment, or caffeine and tobacco the last hour before the experiment. Blood pressure measurements were obtained prior to the experiments after minimum 5 minutes of rest. A written informed consent was obtained prior to participation. The study was approved by the regional committee for medical and health research ethics (REK nr 2010/2927) and conducted in accordance with the Declaration of Helsinki. Subjects received a gift certificate of NOK 500 for participation.

A priori power analysis based on previous studies from our laboratory[18,20] showed that 20 subjects were needed to detect a difference of 1 cm on a 10-cm visual analogue scale (VAS) in the CPM effect between the 2 protocols with a SD of 1.5, assuming a 2-sided significance level of 5% and 80% power.

2.2. Design

The present study was an experimental crossover study comparing 2 different heat test-stimuli. Each experiment started with the test-stimulus alone, either tonic or phasic. After a 5-minute break, an identical test-stimulus in parallel with the conditioning stimulus was applied. A parallel CPM design was used to maximize the CPM effect.[26] Since the CPM effect is short lived,[35] a 30-minute break followed to eliminate any carryover CPM effects before the other protocol was conducted with the same procedure at the opposite arm (**Figure 1**). A computerized block-randomization for both the order of protocol and the test arm was conducted prior to the experiments. Subjects were blinded for the study hypothesis and temperatures of the painful stimuli. A female experimenter (M.U.L.) carried out all experiments. Room temperature (21°C–23°C), placement of instruments, instructions, and the experimenter's clothes were standardized. The experiment was repeated after 14.2 (6.8) (mean [SD]) days with a minimum interval of 7 days, where the second session was identical to the first session.

2.3. Test-stimuli

Test-stimulus was contact heat induced by a 30 × 30-mm Peltier thermode (Medoc, Ramat Yishai, Israel) applied on the proximal volar aspect of the forearm. Thermode temperature increased from baseline (32°C, 2°C per second) to a temperature equivalent to a pain intensity of 6/10 cm on a VAS (see Pretest) and kept constant for a given period. Temperature decreased to baseline at 8°C per second. The tonic heat test-stimulus was constant for 120 seconds, while the phasic test-stimuli consisted of 3 heat plateaus of 5-second duration separated by 10 seconds with baseline temperature. The pain intensity of the test-stimulus was continuously rated on a computerized 10-cm horizontal VAS, where the left end of the scale represented "no pain" and the right end represented "worst pain imaginable." An average pain score of the test-stimulus was calculated for the total time of the test-stimulus and for different periods during the test-stimulus (0–30, 31–60, 61–90, and 91–120 seconds for the tonic heat test-stimulus and first, second, and third stimulation for the phasic heat test-stimuli) (**Figure 2**). The average pain score during the

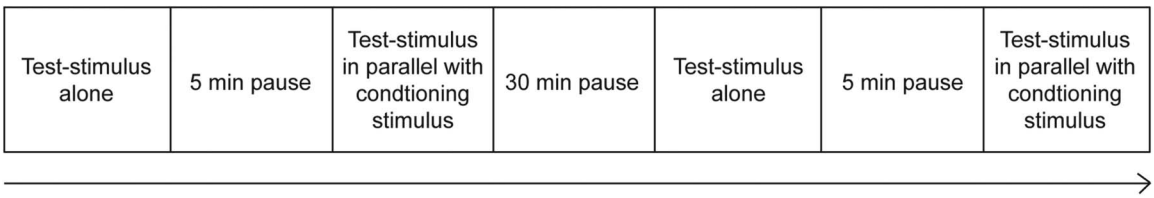

Figure 1. Experiment timeline. The main experiment started with the test-stimulus alone, either the tonic heat test-stimulus or the phasic heat test-stimuli. After a 5-minute break, an identical test-stimulus in parallel with the conditioning stimulus was applied. A 30-minute pause followed before the other protocol was conducted with the same procedure at the opposite arm. An identical experiment was conducted after a minimum of 7 days.

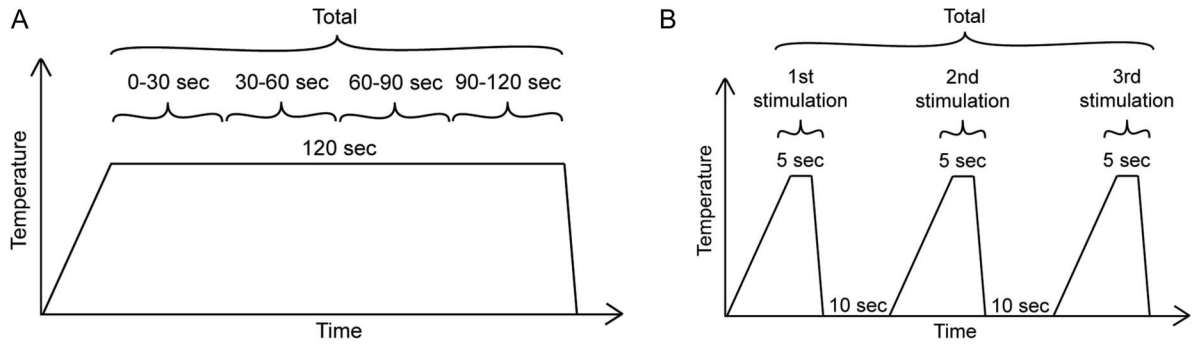

Figure 2. Illustration of the test-stimulus in 2 different conditioned pain modulation protocols. In both protocols, the test-stimulus was a contact heat stimulus with a temperature equivalent to a pain intensity level of 6/10 on a visual analogue scale (VAS) applied on the proximal volar aspect of the subjects' forearm. (A) The protocol with the tonic heat test-stimulus consisted of a constant heat stimulus for 120 seconds. An average test-stimulus pain score was calculated for the total time of the test-stimulus and for 4 different periods along the test-stimulus (0–30, 31–60, 61–90, and 91–120 seconds). (B) The protocol with the phasic stimuli consisted of 3 heat plateaus of 5 seconds duration separated by 10 seconds. An average test-stimulus pain score based on the highest VAS score in each of the 3 stimulations was calculated for the total time of the test-stimulus and for the different stimulations (first, second, and third stimulations).

phasic heat test-stimuli was calculated based on the highest VAS score in each of the 3 stimulations.

2.4. Conditioning stimulus

A 7°C cold-pressor test (CPT) (LAUDA Alpha RA8; LAUDA-Brinkman LP, New Jersey, NJ) was used as conditioning stimulus at the contralateral hand to the test-stimulus side. Subjects were instructed to hold their hand wide open and steady in the container with circulating water up to the wrist. The hand was immersed at the same time as the test-stimulus started and withdrawn after the test-stimulus was terminated or until the pain forced the subject to remove the hand from the container. Thereafter, the overall conditioning stimulus pain intensity was rated using a verbal numerical rating scale with the endpoints; 0 = "no pain," 10 = "worst pain imaginable."

Two subjects withdrew their hand from the cold water before completing the 2 minutes during both protocols and both sessions. A third subject was also unable to complete the conditioning stimulus, but only in the first session during the tonic heat test-stimuli.

2.5. Pretest

A pretest was performed to familiarize subjects with the pain intensity ratings and to determine an individual test-stimulus temperature for each subject to eliminate floor and ceiling effects due to individual differences in heat pain perception. A novel technique was used to calculate the test-stimulus temperature (see supplemental figure, available at http://links.lww.com/PR9/A13), which was aimed to reflect a pain intensity of 6/10 cm on a VAS, but a score within the range of 4–9/10 cm was considered acceptable.

2.6. Data analysis and statistics

The absolute CPM effect was defined as the difference in pain ratings during the test-stimulus in parallel with the conditioning stimulus compared to the pain ratings during the test-stimulus alone. Additionally, the CPM effect was calculated as a percent change CPM effect (absolute CPM effect/test-stimulus pain × 100).

Statistical analyses were conducted using SPSS Statistics v. 21 (IBM, Armonk, NY). *P*-values ≤0.05 were regarded as significant. The distribution of data was assessed in preliminary analyses by a Shapiro–Wilk test and inspection of descriptive statistics, histograms, boxplots, and Q-Q plots. These analyses did not indicate any extreme departures from normality that would affect the planned parametric analysis.

Pain ratings of the test-stimulus alone were compared with pain ratings during conditioning stimulus in paired sample Student *t* tests. Bonferroni correction for multiple testing was applied. Differences between the 2 protocols and sessions regarding pain ratings of the test-stimulus, test-stimulus temperature, and pain ratings of the conditioning stimulus were investigated with repeated-measures analysis of variance (RM ANOVA), with session (levels: first session vs second session) and protocol (levels: tonic stimulus vs phasic stimuli) as factors.

Differences in CPM effect between the 2 protocols were investigated with RM ANOVA, with session (levels: first session vs second session) and protocol (levels: tonic stimulus vs phasic stimuli) as factors. To assess the relative reliability, intraclass correlation coefficients with a 2-way random-effect model ($ICC_{2,1}$) and absolute agreement definition for single measures were used (<0.4: poor reliability; 0.4–0.59: fair reliability; 0.6–0.75: good reliability; >0.75: excellent reliability).[28] Several indices were used to assess the absolute reliability. The mean difference ($\bar{d}$) was calculated by subtracting the mean CPM effect in the first session from the second session and evaluated with a 1-sample Student *t* test. Ninety-five percent limits of agreement (LoA) was calculated as $\bar{d} \pm 1.96 \times SD_{diff}$ (SD_{diff} = SD of the difference).[5] Coefficient of variation (CV%) was calculated as $(SD_{diff}/\sqrt{2}) \times 100$/mean of the 2 sessions.

An RM ANOVA with Bonferroni correction was conducted to determine differences in CPM effect during different periods. Session (levels: first session vs second session) and periods (levels: 0–30 vs 31–60 seconds vs 61–90 vs 91–120 seconds or first stimulation vs second stimulation vs third stimulation) were used as factors. Greenhouse-Geisser correction was used if sphericity was violated.

Subjects were dichotomized into responders and nonresponders (responder = CPM effect <0), and the 95% confidence interval (95% confidence interval [CI]) of the estimated proportion was calculated.

3. Results

3.1. Sample

Twenty-seven subjects were included in the study. One subject was excluded due to hypertension, one subject was unable to

tolerate the tonic heat test-stimulus for the required period of time, and one subject chose not to participate in the second session because the subject found the conditioning stimulus too uncomfortable. Thus, a total of 24 subjects (10 females) were included in the analysis. For sample characteristic see **Table 1**.

3.2. Conditioned pain modulation effect

A CPM effect was observed with the tonic heat test-stimulus for each period in both sessions (**Figures 3A and B**). A CPM effect was also observed in the protocol with phasic heat test-stimuli, although not for all stimulations when analyzed separately (**Figures 3C and D**). When combining data from both sessions, the mean absolute CPM effect (95% CI) for the protocol with tonic heat test-stimulus was −2.8 cm (95% CI: −3.4 to −2.2), corresponding to a percent change of −47.5%, while the mean absolute CPM effect for the other protocol was −1.4 cm (95% CI: −2.0 to −0.8), corresponding to a percent change of −27.0%. The protocol with tonic heat test-stimulus yielded a significantly larger absolute CPM than the protocol with phasic heat test-stimuli ($P < 0.001$; **Figure 4**). A significant increase in absolute CPM effects was observed over time in the protocol with tonic heat test-stimulus ($P < 0.001$) and between the different stimulations in the other protocol ($P = 0.006$). However, this difference was only significant between the first (0–30 seconds) and the 3 other periods (30–60 seconds: $P < 0.001$, 60–90 seconds: $P < 0.001$, 90–120 seconds: $P = 0.004$), indicating that there was no further increase in absolute CPM effect after 30 seconds. In the protocol with phasic heat test-stimuli, the significant difference was found between the first and third stimulations ($P = 0.023$).

Test-stimuli temperatures did not differ between the protocols ($P = 0.421$). Slightly higher test-stimulus temperature was required to achieve a pain intensity of 6/10 cm on a VAS in the second session than in the first session ($P = 0.049$). There was a trend towards higher pain ratings of the test-stimulus during the tonic heat test-stimulus than the phasic heat test-stimuli ($P = 0.062$), but no difference when comparing the test-stimuli ratings between the first and second sessions ($P = 0.386$). Neither were there any significant differences in pain rating in the 2 protocols ($P = 0.218$) nor any differences in such pain ratings between sessions ($P = 0.703$). No interactions were found between sessions and protocols in any of the variables tested.

Table 1

Sample characteristics.

Variable	Value
Sex	
Male, n (%)	14 (58.3)
Female, n (%)	10 (41.7)
Age, y, mean (SD)	25.8 (3.8)
BMI, kg/m^2, mean (SD)	23.7 (3.1)
Education	
Primary school 7–10 y, n (%)	0
Vocational high school, n (%)	0
General high school, n (%)	7 (29.2)
College or university <4 y, n (%)	15 (62.5)
College or university >4 y, n (%)	2 (8.3)
Dominant hand	
Right, n (%)	19 (79.2)
Left, n (%)	5 (20.8)
CPM responders	
Protocol with tonic heat test-stimulus first session, n (%, 95% CI)	22 (92, 74–97)
Protocol with tonic heat test-stimulus second session, n (%, 95% CI)	24 (100, 86–100)
Protocol with phasic heat test-stimuli first session, n (%, 95% CI)	20 (83, 64–93)
Protocol with phasic heat test-stimuli second session, n (%, 95% CI)	18 (75, 55–88)

BMI, body mass index; CI, confidence interval; CPM, conditioned pain modulation; CPM responder, subjects with an absolute CPM effect <0.

3.3. Test–retest reliability

The ICC values of the total CPM effect suggest good relative reliability in the protocol with tonic test-stimulus and fair reliability in the protocol of phasic stimuli (**Tables 2 and 3**). Only the second stimulation in the latter protocol showed a significant difference in mean difference ($\bar{d}$). Large LoA was observed for CPM effects in both protocols, which indicates large intraindividual differences between sessions. Also, high CV% was found in CPM effects in both protocols. An increase in ICC and a decrease in CV% were observed for CPM effects during each period in both protocols, indicating increasing reliability with increasing duration of the stimulus. However, as demonstrated by the ICC's overlapping 95% CI, this increase was not significant. The reliability of the test-stimulus alone in both protocols was also overall poor, while the conditioning stimulus showed excellent reliability (**Tables 2 and 3**).

4. Discussion

The present study showed that a protocol using tonic heat as test-stimulus resulted in a significantly larger CPM effect compared to a protocol using phasic heat as test-stimulus. Fair and good relative reliability was observed for the CPM effect in the protocol with phasic test-stimuli and tonic test-stimulus, respectively. In both protocols, the absolute reliability indices displayed good agreement in the mean CPM effect between the 2 sessions. However, high intraindividual variability was observed for both protocols.

4.1. Conditioned pain modulation effect

The 2 protocols compared in the present do not differ in test-stimulus temperature, test-stimulus pain ratings, or conditioning stimulus pain ratings. This suggests that the observed difference in the CPM effect between the 2 protocols is most likely due to the differences in design of the 2 protocols. Based on the resulting increase in CPM effects during the different periods, one could argue that tonic stimulation longer than 30 seconds as well as phasic stimulations with more than 2 stimulations is preferable to those less than 30 seconds duration or with less than 2 stimulations. However, further research specially designed for this purpose is needed to confirm such assumptions.

Currently, only a few studies have compared various test-stimuli using the CPM paradigm,[11,12,17,19,22] and no studies have compared protocols with tonic and phasic heat test-stimuli. The CPM effect observed using the protocol with tonic heat test-stimulus is consistent with findings from several other studies using 120 seconds of heat stimulation as test-stimulus and 7°C–12°C CPT as conditioning stimulus.[18,20,24,28] In these studies, the percent change CPM effect ranges between −29.1% and −39.5%, ie, slightly lower than the percent change observed in the present study at −47.5%. A possible explanation for this difference may be differences in design (parallel vs sequential CPM testing) or related to the intensity of the test-stimulus as the

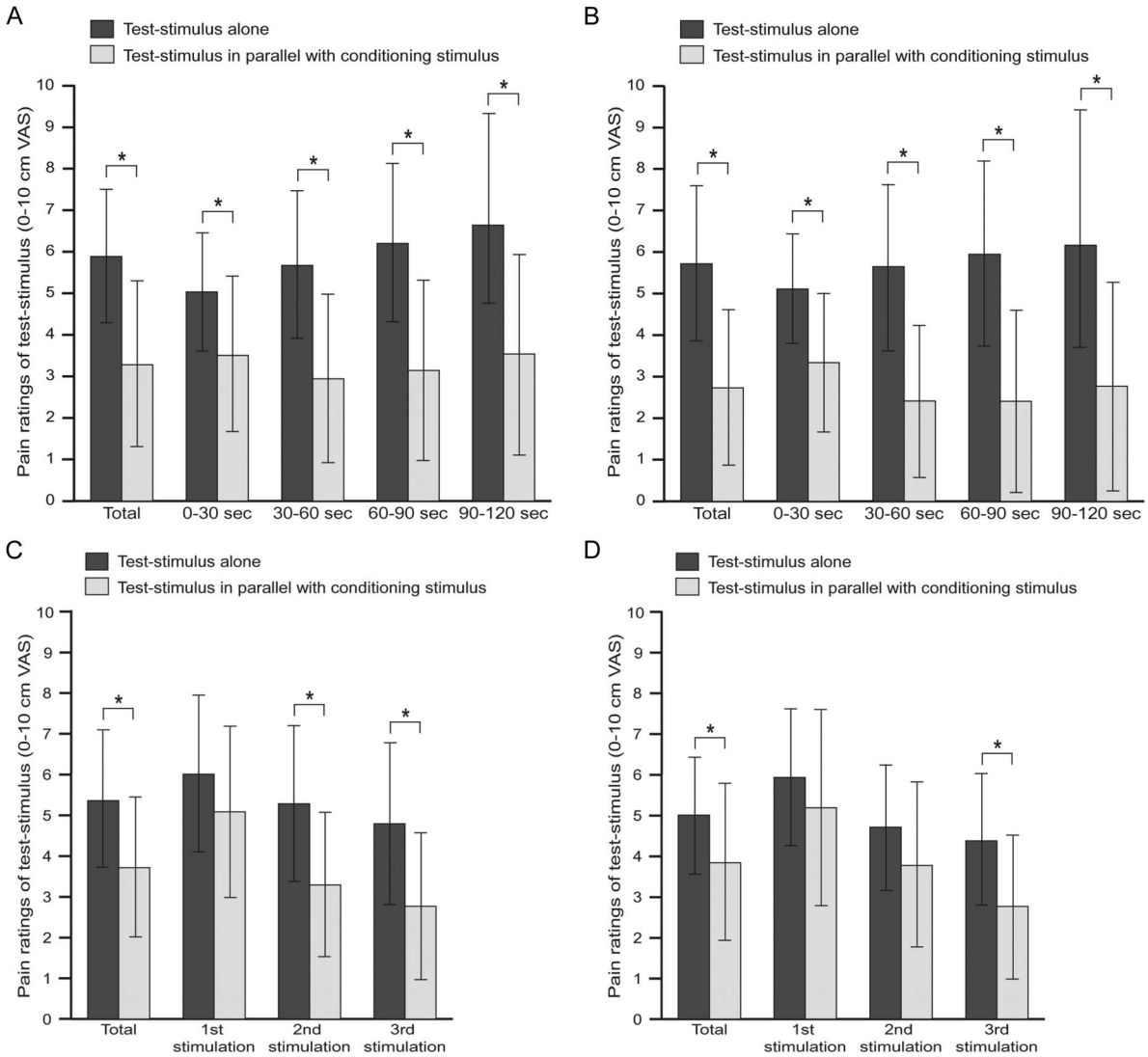

Figure 3. Mean test-stimulus pain ratings during test-stimulus alone and during test-stimulus in parallel with conditioning stimulus for both protocols divided in the different periods at both sessions. A conditioned pain modulation effect was observed in both protocols and sessions as test-stimulus pain ratings were significantly lower during test-stimulus in parallel with conditioning stimulus than during test-stimulus alone. (A) The protocol with the tonic heat test-stimuli, the first session. (B) The protocol with the tonic heat test-stimuli, the second session. (C) The protocol with the phasic heat test-stimulus, the first session. (D) The protocol with the phasic heat test-stimulus, the second session. * = $P \leq 0.05$ (paired sample Student *t* tests with Bonferroni correction). Error bars = ± 1 SD.

test-stimulus is rated slightly lower in comparable studies. However, such assumptions are not clearly supported in the literature.[21] Furthermore, although the present study was not designed for investigating such assumptions, we found no correlation between pain ratings of the test-stimulus and CPM effect in post-hoc analysis.

A study by Treister et al.[29] used 5 heat stimulations of 3-second duration with 12-second intervals as test-stimulus and a 30-second CPT at 12°C as conditioning stimulus, which resulted in a percent change CPM effect of −43%. An absolute CPM effect of −2.9 cm on a 10-cm VAS (percent change not reported) was reported by Demeter et al.[8] when using a comparable design as Treister et al.[29] The CPM effect in these studies is somewhat higher than the CPM effect in the protocol with phasic test-stimuli in the present study (−27.0% or −1.4 cm). As the CPM effect seems to increase from the first to the third stimulation in the present study, similar increases could also be expected for subsequent stimulations, which may serve as an explanation for larger CPM effects in the comparable studies with 5 stimulations. Also, differences in duration of the stimulation and the length of intervals between each of the stimulations may contribute to differences in CPM effects between studies.

4.2. Test-retest reliability

To our knowledge, 13 studies have attempted to analyze the test–retest reliability of the CPM paradigm, and 10 of them in healthy volunteers.[6,9,12,13]

Five studies have investigated relative reliability in different protocols using heat as test-stimulus with either CPT or heat as conditioning stimulus.[9,11,12,30,33] Granovsky et al.[11] reported poor reliability (ICC 0.21 and 0.34) in 2 protocols with tonic heat test-stimulus of 30 and 45 seconds duration, respectively, in a traditional parallel CPM testing with heat as conditioning stimulus, while fair reliability (ICC 0.59) was observed when test-stimulus was delivered for 45 seconds accompanied by a heat conditioning stimulus during the last 25 seconds. Poor reliability (ICC 0.39) was reported by Wilson et al.,[33] who used a tonic heat

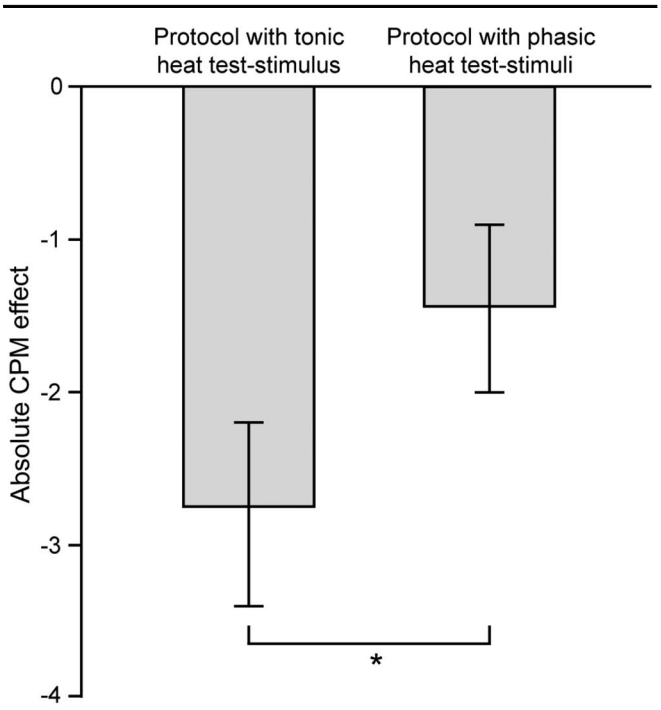

Figure 4. The absolute mean CPM effect for the protocol with tonic heat test-stimulus and the protocol with phasic heat test-stimuli when sessions are combined. The absolute CPM effect in the protocol with the tonic heat test-stimulus was significantly larger compared to the absolute CPM effect in the protocol with the phasic heat test-stimuli ($P < 0.001$). CPM, conditioned pain modulation. *P-value (repeated-measures analysis of variance). Error bars = 95% confidence intervals.

test-stimulus of 30 seconds duration in parallel with a heat as conditioning stimulus. The ICC values in these studies are somewhat lower than the ICC value observed for the protocol with tonic heat test-stimulus in the present study. A possible explanation for this could be the duration of the test-stimulus, since the ICC values in the present study seems to increase with time. The differences could also be due to the use of different conditioning stimulus (heat vs CPT). A recent study by Gehling et al.[9] used 30 seconds of tonic heat as test-stimulus in parallel with a 10°C CPT as conditioning stimulus, which resulted in good reliability (ICC 0.62), similar to the present study. It should, however, be noted that Gehling et al.[9] investigated the short-term test–retest reliability within 72 hours. Valencia et al.[30] reported good reliability (ICC 0.71) when using 5 phasic heat stimulations of 1 second duration as test-stimulus and an 8°C CPT as conditioning stimulus in a sequential test paradigm. Good reliability (ICC 0.60) was also found in another recent study using heat pain threshold as test-stimulus in parallel with a 0°C–4°C CPT.[12]

Although ICC is a common method to analyze test–retest reliability, it is not a sufficient measurement.[1,16] High correlation may be observed as long as individuals have maintained their position in the sample across repeated measures even though they may have changed essentially from session to session. In addition, ICC strongly depends on the sample's heterogeneity, leaving a homogenous group with lower ICC values than a heterogenous group, although the difference between sessions are the same in both groups.[1] Based on these limitations, using ICC alone may lead to false conclusions and may have contributed to the inconsistencies between prior CPM reliability studies. Thus, it is recommended to include measures of absolute reliability in test–retest reliability studies.[1,7,13]

The absolute reliability in the present study did not differ substantially between the protocols. The only apparent difference was the CV%, where the protocol using tonic heat test-stimulus was superior to the protocol using phasic heat test-stimuli. The CV% in the present study is in line with other studies using CV% as an indicator of absolute reliability. Oono et al.[22] reported a CV % of 41.1% using CPT as conditioning stimulus and pressure pain tolerance as test-stimulus, while Biurrun Manresa et al.[3] reported a CV% of 64.1% to 76.2% when using the nociceptive withdrawal reflex as test-stimulus and a CPT below 2°C as conditioning stimulus. A recent study by Imai et al.[12] found a CV% of 2.3%, when using heat pain perception level as test-stimulus.

Table 2

Test–retest indices for the test-stimulus, conditioning stimulus, and absolute CPM effect in the protocol with tonic heat test-stimulus.

Variables	Session I (cm; mean, SD)	Session II (cm; mean, SD)	$\bar{d}$ (LoA LB to UB)	P*	CV (%)	$ICC_{2,1}$ (95% CI)
Total pain ratings during test-stimulus (0–10 VAS)	5.9 (1.6)	5.7 (1.9)	−0.2 (−4.0 to 3.6)	0.63	23.7	0.42 (0.02 to 0.70)
Pain ratings during test-stimulus: 0–30 s (0–10 VAS)	5.0 (1.5)	5.1 (1.3)	0.1 (−3.1 to 3.2)	0.87	22.8	0.33 (−0.09 to 0.65)
Pain ratings during test-stimulus: 30–60 s (0–10 VAS)	5.7 (1.8)	5.6 (2.0)	0.0 (−4.4 to 4.3)	0.91	27.6	0.36 (−0.05 to 0.67)
Pain ratings during test-stimulus: 60–90 s (0–10 VAS)	6.2 (1.9)	5.9 (2.3)	−0.3 (−4.8 to 4.3)	0.56	27.2	0.41 (0.01 to 0.69)
Pain ratings during test-stimulus: 90–120 s (0–10 VAS)	6.6 (1.9)	6.1 (2.5)	−0.5 (−4.8 to 3.8)	0.27	24.2	0.53 (0.17 to 0.76)
Pain ratings of conditioning stimulus (0–10 NRS)	8.3 (1.8)	8.3 (2.0)	−0.1 (−2.0 to 1.8)	0.63	8.3	0.87 (0.73 to 0.94)
Total absolute CPM effect	−2.6 (1.6)	−3.0 (2.1)	−0.4 (−3.6 to 2.9)	0.29	42.5	0.60 (0.27 to 0.80)
Absolute CPM effect during 0–30 s	−1.5 (1.6)	−1.8 (1.6)	−0.2 (−3.4 to 3.0)	0.49	70.3	0.51 (0.14 to 0.76)
Absolute CPM effect during 30–60 s	−2.7 (1.6)	−3.2 (2.6)	−0.5 (−4.7 to 3.7)	0.27	51.0	0.51 (0.14 to 0.75)
Absolute CPM effect during 60–90 s	−3.1 (1.9)	−3.5 (2.5)	−0.5 (−4.5 to 3.6)	0.28	44.5	0.56 (0.21 to 0.78)
Absolute CPM effect during 90–120 s	−3.1 (2.3)	−3.4 (2.6)	−0.3 (−4.1 to 3.5)	0.49	42.2	0.69 (0.40 to 0.85)

* *P* value obtained with 1-sample Student *t* test of $\bar{d}$.

CI, confidence interval; CPM effect, conditioned pain modulation effect = test-stimulus pain alone − test-stimulus pain during conditioning stimulus; CV, coefficient of variation = $(SD_{diff}/\sqrt{2}) \times 100$/mean of the 2 sessions; $ICC_{2,1}$, intraclass correlation coefficient with a 2-way random-effect model; LoA (LB–UB), limits of agreement (upper boundary–lower boundary) = $\bar{d} \pm 1.96 \times SD_{diff}$, where SD_{diff} is the SD of the difference; mean difference ($\bar{d}$), mean difference in absolute CPM effect between the first session and the second session$\bar{d}\bar{d}$; VAS, visual analogue scale.

Table 3

Test–retest indices for test-stimulus, conditioning stimulus, and the absolute CPM effect in the protocol with phasic heat test-stimuli.

Variables	Session I (cm; mean, SD)	Session II (cm; mean, SD)	$\bar{d}$ (LoA LB to UB)	*P**	CV (%)	ICC (95% CI)
Total pain ratings during test-stimulus (0–10 VAS)	5.3 (1.7)	5.0 (1.5)	−0.3 (−3.6 to 2.9)	0.33	22.5	0.47 (0.09 to 0.73)
Pain ratings during test-stimulus: first stimulation (0–10 VAS)	6.0 (2.0)	5.9 (1.7)	−0.4 (−3.4 to 3.3)	0.39	20.0	0.59 (0.25 to 0.80)
Pain ratings during test-stimulus: second stimulation (0–10 VAS)	5.3 (2.0)	4.7 (1.6)	−0.6 (−4.3 to 3.2)	0.17	26.9	0.43 (0.05 to 0.70)
Pain ratings during test-stimulus: third stimulation (0–10 VAS)	4.8 (2.0)	4.4 (1.6)	−0.1 (−4.7 to 3.9)	0.88	34.0	0.29 (−0.12 to 0.61)
Pain ratings of conditioning stimulus (0–10 NRS)	8.1 (1.8)	8.1 (2.1)	0.03 (−1.5 to 1.5)	0.87	6.6	0.93 (0.84 to 0.97)
Total absolute CPM effect	−1.6 (1.4)	−1.2 (1.6)	0.5 (−2.3 to 3.2)	0.11	70.9	0.55 (0.21 to 0.77)
Absolute CPM effect during first stimulation	−0.9 (1.7)	−0.7 (1.8)	0.2 (−3.9 to 4.2)	0.68	177.2	0.29 (−0.13 to 0.62)
Absolute CPM effect during second stimulation	−2.0 (1.7)	−0.9 (2.0)	1.0 (−3.1 to 5.2)	0.02	101.8	0.34 (−0.03 to 0.63)
Absolute CPM effect during third stimulation	−2.0 (1.9)	−1.6 (2.0)	0.4 (−2.7 to 3.5)	0.22	62.2	0.65 (0.35 to 0.83)

* *P* value obtained with 1-sample Student *t* test of $\bar{d}$.

CI, confidence interval; CPM effect, conditioned pain modulation effect = test-stimulus pain alone − test-stimulus pain during conditioning stimulus; CV, coefficient of variation = $(SD_{diff}/\sqrt{2}) \times 100$/mean of the 2 sessions; $ICC_{2,1}$, intraclass correlation coefficient with a 2-way random-effect model; LoA (LB–UB), limits of agreement (upper boundary–lower boundary) = $\bar{d} \pm 1.96 \times SD_{diff}$, where SD_{diff} is the SD of the difference; mean difference ($\bar{d}$), mean difference in absolute CPM effect between the first session and the second session; NRS, numerical rating scale; VAS, visual analogue scale$\bar{d}$.

However, even though CV% is a dimensionless parameter, it is highly dependent on the employed scale. With a mean close to zero (eg, with the CPM effect assessed as a change in cm on the VAS), CV% would be large. In contrast, the CV% would be small if the variation is considerably lower than the mean (eg, with the CPM effect assessed as a change in heat pain perception levels, with a mean of 47°C).

No evident change in the mean difference ($\bar{d}$) in CPM effect was found between sessions, suggesting absence of learning effects and a reliable CPM effect on a group level. However, large intraindividual variability was observed in both protocols, which indicate that neither of the protocols evokes a reliable CPM effect in healthy adults on an individual level. The LoA in the present study are somewhat higher compared to other studies; Imai et al.[12] reports a LoA of −1.7 to 2.3 and Biurrun Manresa et al.[3] reports an LoA of −2.5 to 2.7, indicating somewhat lower absolute reliability.

The poor absolute reliability observed in the present study complicates the appraisal of the clinical value of CPM. For instance, in the protocol with tonic heat test-stimulus, the LoA shows that an healthy individual can have a CPM effect of −3.6 in the first session (ie, an excellent CPM effect), while 14 days later, the CPM effect could be 0 (ie, no CPM effect at all). Still, the change is lower than the LoA and should not be considered a clinical relevant change. The findings in the present study indicate that CPM varies from time to time and may not be a stable trait after all. Thus, there is reason to exercise caution when using measures of CPM protocol as a basis for clinical decision making on an individual level.

4.3. Strength and limitations

Gender and handedness may possibly influence the CPM effect.[23,25] However, we found no indications that gender or handedness influenced our findings as post-hoc analyses showed no differences in CPM effects between men and women nor any difference in CPM effect when all 5 left-handed subjects were excluded (data not reported). Moreover, we performed a block randomization considering handedness. We did not control or consider the effects of menstrual cycle, but there are no clear evidence that menstrual cycle affects the CPM effect.[33] Both protocols in the present study consisted of parallel testing, which is criticized for being biased by other modulation systems such as distraction.[35] However, one could argue that rating the pain intensity of the conditioning stimulus after the test-stimulus was terminated reduce the distraction effect of the concurrent conditioning stimulus in the present study. In the present study, 2 upper limbs were used, which may result in a segmental spinal inhibitory effect rather than reflecting an inhibitory effect of the DNIC.[34] Importantly, none of these limitations can explain the poor absolute reliability. The poor reliability could, however, be related day-to-day differences in expectation[4] or mood, which needs to be addressed by future research. Although the conditioning stimulus had excellent reliability, 7°C water in 2 minutes may be too painful in a clinical setting, as 3 subjects withdrew their hand before the 2 minutes ended and 1 subject chose not to participate due to the discomfort during the conditioning stimulation.

5. Conclusions

The present study suggests that due to larger CPM effects and somewhat better test–retest reliability, the protocol with the tonic heat test-stimulus is preferable to the protocol with the phasic heat test-stimuli. More importantly, the present study indicates that the CPM effect may not be a stable trait in healthy adults.

Disclosures

The authors have no conflicts of interest to declare.

The present study received funding from the Oslo University Hospital and Oslo and Akershus College University. M. U. Lie received funding from FysioFondet.

Acknowledgments

The authors are most grateful for invaluable assistance from Ingrid Fjeldheim Bånerud and Monica Wigemyr for coordinating the experiments.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PR9/A13.

Article history:
Received 28 April 2017
Received in revised form 15 August 2017
Accepted 13 September 2017

References

[1] Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998; 26:217–38.
[2] Basbaum AI, Bautista DM, Scherrer G, Julius D. Cellular and molecular mechanisms of pain. Cell 2009;139:267–84.
[3] Biurrun Manresa JA, Fritsche R, Vuilleumier PH, Oehler C, Morch CD, Arendt-Nielsen L, Andersen OK, Curatolo M. Is the conditioned pain modulation paradigm reliable? A test-retest assessment using the nociceptive withdrawal reflex. PLoS One 2014;9:e100241.
[4] Bjorkedal E, Flaten MA. Expectations of increased and decreased pain explain the effect of conditioned pain modulation in females. J Pain Res 2012;5:289–300.
[5] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
[6] Costa YM, Morita-Neto O, de Araujo-Junior EN, Sampaio FA, Conti PC, Bonjardim LR. Test-retest reliability of quantitative sensory testing for mechanical somatosensory and pain modulation assessment of masticatory structures. J Oral Rehabil 2017;44:197–204.
[7] de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. New York: Cambridge University Press, 2011.
[8] Demeter N, Josman N, Eisenberg E, Pud D. Who can benefit from virtual reality to reduce experimental pain? A crossover study in healthy subjects. Eur J Pain 2015;19:1467–75.
[9] Gehling J, Mainka T, Vollert J, Pogatzki-Zahn EM, Maier C, Enax-Krumova EK. Short-term test-retest-reliability of conditioned pain modulation using the cold-heat-pain method in healthy subjects and its correlation to parameters of standardized quantitative sensory testing. BMC Neurol 2016;16:125.
[10] Granot M, Granovsky Y, Sprecher E, Nir RR, Yarnitsky D. Contact heat-evoked temporal summation: tonic versus repetitive-phasic stimulation. PAIN 2006;122:295–305.
[11] Granovsky Y, Miller-Barmak A, Goldstein O, Sprecher E, Yarnitsky D. CPM test-retest reliability: "standard" vs "single test-stimulus" protocols. Pain Med 2015;17:521–9.
[12] 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. Somatosensory Mot Res 2016;33:169–77.
[13] Kennedy DL, Kemp HI, Ridout D, Yarnitsky D, Rice AS. Reliability of conditioned pain modulation: a systematic review. PAIN 2016;157:2410–19.
[14] Le Bars D, Dickenson AH, Besson JM. Diffuse noxious inhibitory controls (DNIC). I. Effects on dorsal horn convergent neurones in the rat. PAIN 1979;6:283–304.
[15] 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.
[16] Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil 2005;84:719–23.
[17] Lindstedt F, Berrebi J, Greayer E, Lonsdorf TB, Schalling M, Ingvar M, Kosek E. Conditioned pain modulation is associated with common polymorphisms in the serotonin transporter gene. PLoS One 2011;6:e18252.
[18] Matre D, Andersen MR, Knardahl S, Nilsen KB. Conditioned pain modulation is not decreased after partial sleep restriction. Eur J Pain 2016;20:408–16.
[19] Nahman-Averbuch H, Yarnitsky D, Granovsky Y, Gerber E, Dagul P, Granot M. The role of stimulation parameters on the conditioned pain modulation response. Scand J Pain 2013;4:10–14.
[20] Nilsen KB, Olsen IC, Solem AN, Matre D. A large conditioned pain modulation response is not related to a large blood pressure response: a study in healthy men. Eur J Pain 2014;18:1271–9.
[21] Oono Y, Fujii K, Motohashi K, Umino M. Diffuse noxious inhibitory controls triggered by heterotopic CO2 laser conditioning stimulation decreased the SEP amplitudes induced by electrical tooth stimulation with different intensity at an equally inhibitory rate. PAIN 2008;136:356–65.
[22] Oono Y, Nie H, Matos RL, Wang K, Arendt-Nielsen L. The inter- and intra-individual variance in descending pain modulation evoked by different conditioning stimuli in healthy men. Scand J Pain 2011;2:162–9.
[23] Popescu A, LeResche L, Truelove EL, Drangsholt MT. Gender differences in pain modulation by diffuse noxious inhibitory controls: a systematic review. PAIN 2010;150:309–18.
[24] Potvin S, Stip E, Tempier A, Pampoulova T, Bentaleb LA, Lalonde P, Lipp O, Goffaux P, Marchand S. Pain perception in schizophrenia: no changes in diffuse noxious inhibitory controls (DNIC) but a lack of pain sensitization. J Psychiatr Res 2008;42:1010–16.
[25] Pud D, Golan Y, Pesta R. Hand dominancy–a feature affecting sensitivity to pain. Neurosci Lett 2009;467:237–40.
[26] Pud D, Granovsky Y, Yarnitsky D. The methodology of experimentally induced diffuse noxious inhibitory control (DNIC)-like effect in humans. PAIN 2009;144:16–19.
[27] Shrout PE, Fleiss JL. Intracalss correlations: uses in assesing rater reliability. Psychol Bull 1979;86:420–8.
[28] Tousignant-Laflamme Y, Page S, Goffaux P, Marchand S. An experimental model to measure excitatory and inhibitory pain mechanisms in humans. Brain Res 2008;1230:73–9.
[29] Treister R, Eisenberg E, Gershon E, Haddad M, Pud D. Factors affecting—and relationships between-different modes of endogenous pain modulation in healthy volunteers. Eur J Pain 2010;14:608–14.
[30] Valencia C, Kindler LL, Fillingim RB, George SZ. Stability of conditioned pain modulation in two musculoskeletal pain models: investigating the influence of shoulder pain intensity and gender. BMC Musculoskelet Disord 2013;14:182.
[31] 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.
[32] Willis WD, Westlund KN. Neuroanatomy of the pain system and of the pathways that modulate pain. J Clin Neurophysiol 1997;14:2–31.
[33] Wilson H, Carvalho B, Granot M, Landau R. Temporal stability of conditioned pain modulation in healthy women over four menstrual cycles at the follicular and luteal phases. PAIN 2013;154:2633–8.
[34] Yarnitsky D. Role of endogenous pain modulation in chronic pain mechanisms and treatment. PAIN 2015;156(suppl 1):S24–31.
[35] Yarnitsky D, Bouhassira D, Drewes AM, Fillingim RB, Granot M, Hansson P, Landau R, Marchand S, Matre D, Nilsen KB, Stubhaug A, Treede RD, Wilder-Smith OH. Recommendations on practice of conditioned pain modulation (CPM) testing. Eur J Pain 2015;19:805–6.
[36] Yarnitsky D, Crispel Y, Eisenberg E, Granovsky Y, Ben-Nun A, Sprecher E, Best LA, Granot M. Prediction of chronic post-operative pain: pre-operative DNIC testing identifies patients at risk. PAIN 2008; 138:22–8.