Reliability of subjective pain ratings and nociceptive flexion reflex responses as measures of conditioned pain modulation

C Jurth1, Benno Rehberg2, Falk von Dincklage MD1

BACKGROUND: The endogenous modulation of pain can be assessed through conditioned pain modulation (CPM), which can be quantified using subjective pain ratings or nociceptive flexion reflexes. However, to date, the test-retest reliability has only been investigated for subjective pain ratings.

OBJECTIVE: To compare the test-retest reliability of CPM-induced changes, measured using subjective pain ratings and nociceptive flexion reflexes, to provide a reliable scoring parameter for future studies.

METHOD: A total of 40 healthy volunteers each received painful electrical stimuli to the sural nerve to elicit nociceptive flexion reflexes. Reflex sizes and subjective pain ratings were recorded before and during the immersion of the contralateral hand in hot water to induce CPM as well as innocuous water as control. Measurements were repeated in a retest 28 days later.

RESULTS: Intraclass correlation coefficients showed good test-retest reliabilities of CPM during the hot water stimulus for both scoring parameters. Subjective pain ratings also correlated between test and retest during the control stimulus.

CONCLUSIONS: Subjective pain ratings and nociceptive flexion reflexes show comparable test-retest reliabilities, but they reflect different components of CPM. While subjective pain ratings appear to incorporate cognitive influences to a larger degree, reflex responses appear to reflect spinal nociception more purely.

Key Words: Conditioned pain modulation; Diffuse noxious inhibitory controls; Nociceptive flexion reflex; Test-retest reliability

The concept of modulating a painful sensation by inducing a painful conditioning stimulus is often used to investigate pain pathophysiology (1). Altered pain modulation has been shown in various pain disorders (2,3) and recent research indicates that it may identify patients at risk for developing chronic pain because impaired pain modulation correlates with a greater chance of acquiring pain disorders (4,5).

The experimental paradigm to investigate this concept of an endogenous modulation of pain by conditioning stimuli is termed ‘conditioned pain modulation’ (CPM), for which a standardized conditioning stimulus is applied to inhibit an experimental painful test stimulus (6). Via spinal-superior pain feedback loops, painful stimuli at any part of the body inhibit nociceptive transmission distant from these stimuli (7). CPM is believed to be mediated, in part, by the descending inhibition of nociceptive transmission of wide dynamic range neurons in the dorsal horn as well as in the trigeminal nuclei, a system known as diffuse noxious inhibitory controls (8,9). Furthermore, modulation on the cerebral level via cortico-cortical interactions has been shown during CPM (10).

Several modalities, such as cold, ischemic, chemical or heat pain, have been used as the conditioning stimulus to inhibit the test stimulus, which is commonly applied as a thermal, mechanical or electrical pain stimulus (1). To quantify the inhibitory effect of the conditioning stimulus on the test stimulus, typically each test stimulus is rated subjectively on a pain rating scale. Another method to quantify pain modulation is the nociceptive flexion reflex (NFR) (11,12). The NFR is a withdrawal reflex of the leg, typically recorded using a biceps femoris muscle electromyogram following stimulation of the sural nerve at the foot. The reflex threshold and amplitude are closely related to those of the subjective pain ratings, making it an objective and reliable tool for pain assessment in CPM paradigms, which may be differentially affected by cognitive influences (13,14). The NFR has been shown in various pain paradigms to quantify the effects of CPM, and it may be used as an objective and reliable parameter for pain assessment in CPM paradigms, which may be differentially affected by cognitive influences.

The aim of the present study was, therefore, to investigate and compare the test-retest reliability of subjective pain ratings and nociceptive flexion reflex responses as instruments to quantify CPM in healthy volunteers to provide a reliable scoring parameter for CPM in future studies.
METHODS

Participants
After approval by the ethics committee of the Charité – Universitätsmedizin Berlin (Berlin, Germany) and written informed consent, the study was conducted with 40 healthy volunteers (20 men and 20 women). Exclusion criteria were chronic or acute pain conditions. Each volunteer was required to abstain from analgesics for 48 h, any drugs, caffeine and alcohol for 24 h as well as strenuous exercise for 6 h before measurements. Participants received €50 as compensation for the participation.

Experimental procedure
The study was performed during two sessions, 28 days apart. An interval of 28 days was chosen to reduce variability due to the menstrual cycle (19,20), although endogenous inhibition of the NFR may not be altered by the menstrual phases (12). The measurements were conducted in a quiet, air-conditioned room and began in the morning to reduce variability due to circadian changes (21). To determine individual levels of anxiety, Spielberger’s State/Trait Anxiety Inventory scores were evaluated (22). To control for differences in subjective pain perception, the volunteers completed the Pain Sensitivity Questionnaire (23). The volunteer’s expectations regarding the effect of the conditioning stimuli on the level of pain elicited by the test stimulus were grouped as “increase in pain”, “no change” or “decrease in pain”. All assessments were conducted before both sessions.

Before commencement of the electrophysiological measurements, the volunteers were positioned in a comfortable flexible chair, including a foot rest, with 120° flexion in the hip, 160° in the knee and 110° in the ankle. They were instructed to keep their eyes closed throughout the measurements.

Electrocutaneous stimulation was applied using surface electrodes below the left lateral malleolus to stimulate the sural nerve at its retro-malleolar pathway. Stimuli were applied repeatedly at interstimulus intervals randomized between 8 s and 12 s to minimize the likelihood of stimulus predictability (24). Each stimulus consisted of five rectangular pulses of 1 ms duration at 200 Hz using the computer-controlled DS5 constant current stimulator (Digitimer, United Kingdom).

To evaluate the NFR responses evoked by these stimuli, the biceps femoris electromyogram was recorded using surface electrodes placed over the biceps femoris muscle (25). Electromyographic activity was amplified by a factor of 10,000, hand-pass filtered between 2 Hz and 1 kHz (g.bsamp, gTech, Austria), digitized at 2 kHz (Micro1401, Cambridge Electronic Design Ltd, United Kingdom), rectified and analyzed using Matlab (Matlab R2010b, MathWorks Inc, USA). The reflex interval mean z score in the interval between 90 ms and 180 ms after the stimulus was used to score the NFR magnitude (26).

Experimental sequence
To quantify the subjective pain intensity evoked by the stimuli, the volunteers were instructed to rate each perceived stimulus on a numerical rating scale (NRS). This scale ranged from 0 to 100 and was anchored at either end by 0 = ‘no pain’ and 100 = ‘strongest pain imaginable’.

Before the measurements, all volunteers were habituated to stimuli up to an intensity of NRS 70. Following a resting period, the stimulation intensity corresponding to a pain level of NRS 50 was determined using an up-down staircase method (27). A sequence of stimuli using the predetermined individual intensity evoking NRS 50 followed, which consisted of 15 stimuli to reach a steady state (28), 20 stimuli for the state ‘before conditioning stimulus’ and 20 stimuli for the state ‘during conditioning stimulus’. Throughout the 200 s during which the 20 stimuli of the state ‘during conditioning stimulus’ were applied, the volunteers immersed their contralateral hand in a stirred water bath for which the temperature was set either to 46.5°C to induce CPM or to 33°C as innocuous control (29). After a resting period of 30 min, the stimulation intensity corresponding to a pain level of NRS 50 was again determined (16). Another sequence of stimuli as described above followed, during which the water bath was set to whichever temperature had not been used previously. Volunteers were randomly assigned into couples starting with either hot or innocuous water, and rated the painfulness of each temperature using the NRS. The water’s temperature was electronically regulated and double-checked manually (Haake C10-K10, Thermo Fisher Scientific, USA; GTH 175/Pt, Greisinger Electronic, Germany).

Data analysis and statistical analysis
CPM was quantified for each individual at each session and for each conditioning stimulus as the relative change of the individual mean of the 20 values for the subjective pain ratings and the reflex responses during the conditioning stimulus compared with the individual mean of the 20 respective values before the conditioning stimulus. Outliers that deviated by >2 SDs from the means were removed.

To investigate the test-retest reliability of the effect of the conditioning stimuli on the subjective pain ratings and the reflex size, intra-class correlation coefficients (ICCs) were calculated (30). In contrast to simple correlation coefficients, ICCs take the intraindividual and interindividual variabilities into account to allow for an unbiased comparison of the test-retest reliability between the subjective pain ratings and the reflex size. ICC values <0.4 were considered to be poor, values between 0.4 and 0.75 were considered to be good and values >0.75 were considered to be excellent (31).

In addition to the ICC, standard correlations were performed to investigate whether the test-retest reliabilities of the measures are significantly influenced by the interindividually variability. ICCs account for such variability, while standard correlations do not.

All statistical analyses were performed using the R software package (R Foundation for Statistical Computing, Austria).

RESULTS

General results
Of the 40 volunteers included in the study, four were excluded from analysis: two volunteers did not return for the retest measurements; for one volunteer, the study had to be aborted due to intolerance of the hot water bath; and for one participant, the study had to be aborted due to technical problems with the electrophysiological equipment.

Questionnaire results, pain ratings of the conditioning stimuli, baseline responses to the test stimuli and the changes in subjective pain ratings and reflex size during the conditioning stimuli are shown in Table 1.

Test-retest reliability of the effect of the CPM
To compare the test-retest reliabilities of the changes in subjective pain ratings and reflex size during the conditioning stimuli, ICCs were calculated. Changes in subjective pain ratings showed ICCs of 0.54 (95% CI 0.26 to 0.74) for the hot water condition and 0.42 (95% CI 0.12 to 0.66) for the control condition. Calculation of ICCs for the changes in reflex size resulted in 0.61 (95% CI 0.36 to 0.78) for the hot water condition and 0.07 (95% CI −0.26 to 0.39) for the control condition.

The individual changes in subjective pain ratings and reflex size during the conditioning stimuli at test and retest and their correlations are shown in Figure 1.

DISCUSSION

The aim of the present study was to investigate and compare the test-retest reliability of subjective pain ratings and the NFR to score the inhibitory effect by a conditioning stimulus on an electric test stimulus in healthy volunteers, a phenomenon termed CPM. Our data show a good test-retest reliability for the subjective pain rating reductions during the hot water conditioning stimulus, consistent with previous findings (16-18). Similarly, we observed a good test-retest reliability for NFR reductions during the hot water conditioning stimulus, comparable with that of the subjective pain ratings, although the NFR reduction was not significant for the whole group. However, the two scoring parameters differed regarding the effect of the control conditioning stimulus.
During the control conditioning stimulus, our data revealed a good correlation between the individual changes in subjective pain ratings at the test session compared with the retest session. For the NFR, in contrast, we found no test-retest correlation for the individual effects during the control conditioning stimulus.

The most parsimonious explanation for this difference would be to assume an influence of confounders that affect the individual change in subjective pain ratings during the conditioning stimulus while having no effect on the NFR, which is well established for several cognitive influences (3,32-34). It has been shown that CPM involves mechanisms on several levels, including modulation on the cerebral level via cortico-cortical interactions as well as modulation on the cerebro-spinal level via descending pathways (7,10,35,36). Especially for the cognitive modulation of CPM through changes in attention, expectation and emotion, it has been shown that modulation at the cerebral level plays an important role (32,37-39). Because the NFR can only reflect the effect of the descending pathways while the subjective pain ratings incorporate all mechanisms, cognitive modulation can be expected to influence NFR and subjective pain ratings to a different extent.

Such cognitive influences may, therefore, account for the test-retest correlation of subjective pain rating changes during the control conditions, by creating a confounder that remains at a stable degree and direction between the test and the retest session in each individual, while varying to a larger extent among individuals.

However, if such confounders influence the subjective pain ratings during the innocuous control stimulus measurements, they are likely to have an influence during the hot water conditioning stimulus as well. Therefore, the good test-retest reliability observed for the changes in subjective pain ratings during the hot water conditioning stimulus could, in part, be ascribed to these cognitive influences. The NFR, in contrast, for which no test-retest correlation of its changes during the control conditioning stimulus was observed, appears to be less affected by cognitive influences. The comparable reliabilities for the subjective pain ratings and the NFR may, therefore, be based on at least partially different mechanisms. While the subjective pain ratings incorporate the associated cognitive mechanisms to a larger extent, the NFR instead reflects the nociceptive processes at spinal level.

In our study design, we used pain rating tailored test stimuli for better comparability with previous studies because all previous studies on the reliability of CPM investigated pain ratings only. However, our design increases the population variability of the reflex responses in comparison to the subjective pain ratings. Because the different population variabilities could influence the test-retest reliability as measured by the ICCs, we additionally performed a correlation analysis, showing that the overall results stay the same, whether the intraindividual variability is (ICCs), or is not (standard correlation) taken into account.

The different interindividual variabilities due to the study design may also explain why we found a significant reduction of subjective pain ratings during the conditioning hot water stimulus, while NFR reductions did not reach statistical significance. Because the aim of our study was to investigate the reliability of the individual inhibition, the significance of the inhibition of the whole group is only of interest to confirm that the overall setup is valid to induce inhibition. That the NFR, similar to the subjective pain ratings, is reduced during a hot water conditioning stimulus has been shown in numerous studies. Therefore, the sample size calculation was optimized for the parameter with the lower variability (subjective pain ratings), which already confirms the validity of the setup.

However, because our data show no significant inhibition of the NFR, it is difficult to draw conclusions regarding the reliability of the NFR. Possibly this was only due to the sample size, but it is also possible that our setup was not effective in inducing a CPM effect on the NFR, in which case the reliabilities reported here could not be regarded as reliability of the CPM effect.

For the conditioning stimulus, we chose tonic heat pain set to 46.5°C, as previously suggested (29). The tonic heat stimulus appears to be more constant and less confounded by changes in cardiovascular activity compared with the cold water stimulus (40).

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Questionnaire, median (range)</th>
<th>Test</th>
<th>Retest</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAI-State</td>
<td>35 (22–68)</td>
<td>34 (21–68)</td>
</tr>
<tr>
<td>STAI-Trait</td>
<td>35 (21–66)</td>
<td>34 (20–66)</td>
</tr>
<tr>
<td>PSQ</td>
<td>3.8 (1.1–6.2)</td>
<td>3.8 (1.0–6.5)</td>
</tr>
</tbody>
</table>

Expectations regarding the effects of stimuli on pain levels (increased/no change/decreased):

- Hot water: 2/2/32
- Control: 5/29/2

Pain ratings of the conditioning stimulus, median (range):

- Hot water (46°C): 53 (40–63)
- Control (30°C): 0 (0–0)

Baseline responses to the test stimuli, median (range):

- Pain ratings: 52 (48–55)
- Reflex response: 7.3 (4.8–8.1)

Changes during conditioning stimulus, %, median (95% CI):

- Hot water (46°C): Pain ratings 5.2 (4.8–5.5), Reflex response 7.3 (4.8–8.1)
- Control (33°C): Pain ratings 5.2 (49–55), Reflex response 7.5 (5.6–8.6)

**PSQ** Pain Sensitivity Questionnaire; **STAI** Spielberger’s State/Trait Anxiety Inventory

**Figure 1** Correlations between the individual changes in subjective pain ratings and reflex size during the hot water (46.5°C) and control (33°C) conditioning stimuli at the test session compared with the retest session. Subjective pain ratings are shown for the hot water condition (A) and for the control condition (B). Reflex responses are shown for the hot water condition (C) and for the control condition (D).
SUMMARY

We have shown a good test-retest reliability for both the subjective pain ratings and the NFR reduction to quantify the effect of a painful hot water conditioning stimulus on a painful electric test stimulus. Conclusions regarding the NFR reliability remain difficult because the reflex reduction did not reach significance on the group level, most likely due to its larger variability in our study design. However, the good reliability of the subjective pain ratings may, at least in part, be ascribed to cognitive influences. These influences appear to account for the test-retest correlation of subjective pain rating changes during the control conditions, by creating a confounder that remains at a stable degree and direction between the test and the retest session in each individual while varying among the individuals in a larger extent. Reflex responses appear to be less influenced by these cognitive influences, leading to the assumption that the NFR provides a tool that more purely reflects the CPM effecting conditioning stimulation. Therefore, if cognitive influences are considered to be a relevant part of a specific research question, subjective pain ratings appear to be the method of choice to score the effect of CPM. However, if the primary research question is the effect of CPM on nociception, with less interest in the cognitive influences, the NFR appears to be the method of choice to score the CPM.

FUNDING AND DISCLOSURES: This study was funded through the Internal Research Funding of the Department of Anesthesiology and Intensive Care Medicine, Charité – Universitätsmedizin Berlin. Dr von Dincklage is a participant in the Friedrich C Lauf Clinical Scientist Pilot Program funded by Volkswagen Foundation and Charité Foundation. The authors have no conflicts of interest or financial interests regarding the study.