Edith Cowan University

Research Online

4-17-2020

Considerations for multi-centre conditioned pain modulation (CPM) research; an investigation of the inter-rater reliability, level of agreement and confounders for the Achilles tendon and Triceps Surae

Myles Murphy

William Gibson

Paola Chivers
Edith Cowan University

Sean Docking

Ebonie Rio

Full Title: Considerations for multi-centre conditioned pain modulation (CPM) research; an investigation of the inter-rater reliability, level of agreement and confounders for the Achilles tendon and Triceps Surae

Short Title: inter-rater reliability of conditioned pain modulation (CPM)

ABSTRACT

Objective: This study aimed to investigate the inter-rater reliability of the conditioned pain modulation (CPM) effect.

Methods: The reliability between two examiners assessing the CPM effect via pressure pain thresholds and induced using the cold pressor test of 28 healthy volunteers at the mid-portion Achilles tendon (AT) and Triceps Surae musculotendinous junction (TS) was performed. Reliability was calculated using intraclass correlation (ICC). Confounders were assessed using multivariable generalised estimating equations (GEE). Bias in the level of agreement was assumed if the confidence intervals of the mean difference in Bland-Altman plots did not cross the line of equality.

Results: The inter-rater reliability of the CPM effect was poor to moderate in the AT (ICC 95%CI= 0.00-0.66) and TS (ICC 95%CI= 0.00-0.69). However, when accounting for confounders within the GEE there were no differences between testers and Bland-Altman plots reported good agreement between testers. Habitual completion of running-related physical activity was a confounder for both the AT parallel-paradigm (p= 0.017) and sequential-paradigm (p= 0.029). Testing order was a confounder for the AT (p= 0.023) and TS (p= 0.014) parallel-paradigm.

Conclusion: This study suggests the CPM effect may be site specific (i.e. differences between the AT and TS exist). Additionally, differences in the reliability between examiners is likely due to the influence of confounders and not examiner technique and therefore confounders should be considered in research investigating the CPM effect.

KEYWORDS

INTRODUCTION

Conditioned pain modulation (CPM) is an experimental procedure that attempts to measure the endogenous analgesic effect in response to a noxious conditioning stimulus.(1) Assessment of this phenomenon is increasingly utilised in pain research due to its reduction in certain chronic pain conditions.(2) It has been suggested that CPM investigations may provide key insights regarding pain modulatory mechanisms and assist in understanding persistent chronic pain and inform better management.(3)

To measure CPM, a testing stimulus (TS) is performed at baseline and repeated following the application of a noxious conditioning stimulus (CS).(4) Multiple TS and CS paradigms exist.(5) Pressure pain threshold (PPT) assessment (a proxy of mechanical sensitivity) as a TS paired with the cold pressor test (CPT) as the CS is a common model of CPM induction.(2) The TS can be assessed whilst exposure to the CS is occurring (parallel-paradigm) or can be assessed following the removal of the CS (sequential-paradigm).(6)

Different methods exist to report the CPM effect.(2) It is common practice to report on the absolute change of the TS following application of the CS (e.g. Newtons of pressure difference pre-to-post),(4) relative change of the TS following application of the CS (e.g. percentage difference pre-to-post)(4) and whether a meaningful CPM effect was elicited.(7) Meaningful CPM effects have been suggested to be defined as the difference in pre-to-post measurements that are larger than the TS standard error of the measurement.(7)

Kennedy et al. reviewed the CPM effect (10 studies, median sample= 34) and concluded that sufficient evidence exists that CPM is reliable in both painful and healthy volunteers.(1) Most studies were at moderate to high risk of bias with only one study out of the ten not at high risk of bias in at least one domain and there was significant heterogeneity within included studies that could impact the findings.(5) The studies included within Kennedy et al. used the intra-class correlation co-

efficient of within-subject variability to determine reliability.(5) Vaegter et al. assessed the inter-session reliability of CPM in 26 healthy men, mean (SD) age of 25.3 (5.6) with the CPM effect measures 1-3 weeks apart using 10 different protocols.(8) This study showed that using PPTs as the TS and CPT as the CS a CPM effect was able to be elicited in most patients. However, the agreement on classification of those who elicited a CPM effect and those who did not was poor.(8)

Larger samples are likely to provide more accurate estimates of the true effect size(9) and therefore many studies have now commenced multi-site involvement.(10) While this assists researchers in recruitment of larger samples there are problems when outcome measures do not have acceptable reliability. The ability to reliably detect and quantify meaningful CPM effects are fundamental to the success of multi-centre studies. Ideally, reporting of this phenomenon should include absolute, relative and dichotomously identified CPM effects. Specifically, whether the relative CPM effect is reliable independent of levels of agreement in absolute PPT, is of particular interest given intra-rater reliability of PPTs are consistently better than inter-rater reliability.(11, 12)

OBJECTIVES

The objectives of this study were to:

1. Determine the inter-rater reliability of the absolute, relative and dichotomously classified CPM effect assessed at the Achilles tendon and Triceps Surae musculotendinous junction.

2. Determine the level of agreement of the CPM effect assessed at the Achilles tendon and Triceps Surae musculotendinous junction.

3. Investigate potential confounders of the CPM effect assessed at the Achilles tendon and Triceps Surae musculotendinous junction.

METHODS

Ethical considerations

This study was approved by the La Trobe University Australia's Human Research Ethics Committee (Reference number: 07167F) and The University of Notre Dame Australia's Human Research Ethics Committee (Reference number: 017067F).

Study design

An intra- and inter-rater reliability study of PPTs and an inter-rater reliability study of the CPM effect elicited using a CPT and assessed via PPT test site assessment at the mid-portion Achilles tendon and the musculotendinous junction of the Triceps Surae. The study was conducted between June and September 2019.

Data storage

All data will be made available upon reasonable request of the authors.

Setting

All recruitment and data collection occurred at the La Trobe Sport and Exercise Medicine Research Centre in Bundoora, Victoria, performed within a quiet, distraction free room.

Participants

Participants were recruited from staff at the La Trobe Sport and Exercise Medicine Research Centre in Bundoora, Victoria. Staff included within this study were not pain researchers and were unaware of the methodology and expected results of CPM testing. Participants were included provided they met the required inclusion and exclusion criteria:

Inclusion Criteria

- Age range 18-60 years

- Ability to provide informed consent

- Able to complete questionnaires in English

Exclusion Criteria

- Corticosteroid injection to the Achilles tendon within the past 12 months

- Mid-portion or insertional Achilles tendinopathy

- Foot or ankle surgery within previous six months

- Fracture of the lower limb within previous 12 months

- Systemic disease (diabetes, rheumatic disease, circulatory disorders, neurological disorders)

- Regular intake of medications that may impact outcomes (statins or fluroquinolones)

- Confounding injury for CPM (chronic lumbar pain, chronic or recurrent tension headache, migraine, whiplash or painful OA)

- Current radiculopathy

- Contra-indications to cold application (cold uticaria, Raynaud's phenomenon or skin conditions)

Variables

Baseline demographics

We recorded all participants self-reported age (years), sex (male/female), height (cm), and weight (kg) as well as whether they currently completed running related sport (yes/no).

Test Stimulus

The test stimulus used (PPT) was obtained using a manual algometer with a 1cm diameter, circular tip (Force Ten™ FDX digital force gage, Wagner Instruments, Greenwich, CT) that is annually calibrated by two examiners (one Caucasian female examiner and one Caucasian male examiner who both have over five years of experience as physiotherapists and had both received training in

assessment of the CPM effect prior to this study). The mid-portion of the Achilles tendon was determined at a site 2cm proximal to the most superior aspect of the calcaneus, marked using a permanent marker (Figure 1). The musculotendinous junction of the Triceps Surae was determined by palpation and was again marked using a permanent marker (Figure 1). These locations were chosen due to their close proximity to each other and that they represent two different types of tissue; muscle and tendon. The locations were marked with indelible marker in order to ensure both examiners applied the algometer over the same area thus reducing measurement location as a confounder. Standardised instructions were given to the participants of *"We are going to assess the pressure pain thresholds of your Achilles tendon and the muscle-tendon junction of your calf. We are going to increase the pressure using this device until you first experience pain and at that point tell us so we can stop the test. This is not a test to see how much pain you can tolerate we are purely looking for when the sensation of pressure first becomes a one out of ten pain."* Participants were assessed in prone on a portable physiotherapy table with the ankle resting freely over the edge. The ankle was placed into plantar grade by the examiner when performing the PPTs (Figure 1). The algometer was applied perpendicular to the skin and pressure aimed to be increased at a rate of 50 kPa/s until the participant reported they could feel pain. The point at which the participant recorded the first onset of pain (1 on an 11-point numerical rating scale of pain) was recorded and only one test was performed to determine a PPT.

Conditioning Stimulus

The CS used was the CPT and involved a bucket of ice water that was placed next to the participant with the temperature aimed at being maintained as close to five degrees Celsius as possible. Participants immersed their contralateral hand to their testing side to the level of the wrist crease. Their hand remained immersed until instructed to remove the hand by the examiner (180 seconds) or until they were unable to tolerate the feeling of cold.

Assessment of conditioned pain modulation

The assessment procedure for eliciting the CPM effect in the Achilles tendon and Triceps Surae musculotendinous junction is described (figure 2). To our knowledge no consensus exists which describes either the parallel or sequential paradigm as superior and given this is the first work to explore the reliability of the CPM effect of the Achilles tendon Triceps Surae musculotendinous junction we opted for the conservative approach of including both paradigms. The CPM effect at the Achilles tendon and Triceps Surae musculotendinous junction were assessed consecutively by the same examiner within a single CPT. The parallel paradigm was assessed 60 seconds into application of the CPT while the sequential paradigm was assessed immediately after removal of the CPT at 180 seconds. Following completion of assessment by the first examiner the participants were required to have a 30 minute wash out period in order to avoid confounding due to any temporal summation.(13) Following this period the identical testing procedure was performed by the other examiner. The testing order was randomised via a coin flip and examiners were blinded to the results of the other examiners assessment.

Bias

This study was specifically designed to limit the extent of any biases likely to impact analysis of the reliability of CPM. These biases were determined based on the six domains of the Quality in Prognostic Studies risk of bias tool modified by Kennedy et al. when they performed a systematic review on the reliability of CPM.(5) The six domains considered were study participation, study attrition, prognostic factor measurement, outcome measurement, confounding and statistical analysis.(5)

Study Size

A priori power calculation was performed in accordance to the recommendations by Walter et al. 1998 for sample size calculations in reliability studies.(14) Therefore, with two observations per subject, significance set at 0.05, power set at 0.8, an accepted reliability of 0.7, an expected

reliability of 0.9 and a 0% drop out it was calculated that 18 participants were required to be adequately powered to assess reliability.

Statistical methods

Data analysis were conducted using SPSS v25 (IBM Corp. released 2017).

The absolute CPM effect was calculated as:

$$Absolute\ CPM\ effect = Baseline\ PPT - PPT\ during\ CS$$

The relative CPM effect was calculated as:

$$Relative\ CPM\ effect = \frac{Absolute\ CPM\ effect}{Baseline\ PPT} \times 100$$

A CPM effect which increased the PPT value following the CPT was classified as negative for the absolute and relative classification of CPM as an increase in the PPT represents a decrease in mechanical sensitivity. The overall reliability was rated in accordance to the recommendations by Koo and Li, that intraclass correlation coefficient (ICC) values <0.50 indicate of poor reliability, values from 0.50 to <0.75 indicate moderate reliability, values from 0.75 to <0.9 indicate good reliability, and values ≥0.90 indicate excellent reliability.(15) Furthermore, we classified the reliability based on the range of the 95% confidence interval and not the ICC value.(15) For calculating ICC, SPSS uses the method of moments for calculating variance components which can lead to possible negative ICC values which are meaningless, although theoretically possible.(16) Therefore, any negative ICC values in the lower bounds for CI were reported in this study as zero.

Inter-rater reliability of the CPM effect elicited using a cold pressor test

Participant age, sex, height, weight, body mass index (BMI) and whether they performed regular running-related physical activity were described using count, mean and SD, where appropriate. The testing variables which had the potential to differ between examiners (temperature of the CPT, pain

induced from the CPT, Achilles tendon PPT, Triceps Surae musculotendinous junction PPT, absolute CPM effect, relative CPM effect and meaningful CPM effect) were also described. An assessment of normality was performed using the Shapiro-Wilk test. T-tests, or the non-parametric alternative (Mann-Whitney U test), were used to determine differences between these variables for each examiner. Additionally, dependant t-tests were used to determine if there were differences in the baseline PPT for test one versus test two, irrespective of the examiner, to ensure PPTs had returned to baseline before subsequent CPM testing.

The inter-reliability for both the parallel and sequential CPM effect was calculated in three different ways to determine the reliability of the absolute CPM effect, the reliability of the relative CPM effect and the reliability of the meaningful CPM effect:

1. The absolute CPM effect represents the absolute difference of the test stimulus between the baseline value and value following application of the conditioning stimulus. The ICC and 95% confidence interval of a single measurement ICC (3, 1) and SEM for the inter-rater reliability of the absolute CPM effect of both the Achilles tendon and the musculotendinous junction of the Triceps Surae were calculated using a two-way, consistency, mixed-effects model.

2. The relative CPM effect represents the percentage difference of the test stimulus between the baseline score and score following application of the conditioning stimulus. The ICC and 95% confidence interval of a single measurement ICC (3, 1) and SEM for the inter-rater reliability of the relative CPM effect of both the Achilles tendon and the musculotendinous junction of the Triceps Surae were calculated using a two-way, consistency, mixed-effects model.

3. The meaningful CPM effect represents an outcome of whether a CPM effect was achieved (Intact CPM effect, absent CPM effect, inverse CPM effect). A CPM effect was classified as having been achieved when the relative CPM effect exceeded the relative reliability SEM as

calculated in appendix A. The reliability of the meaningful CPM effect was calculated using

Kappa Chi-square analysis.

The level of agreement of agreement between two examiners of the CPM effect is presented

graphically using Bland Altman plots.(17) The y-axis represents the difference between the absolute

CPM effect from examiner one and examiner two divided by the mean percentage and the x-axis

represents the mean of the absolute CPM effect between examiner one and examiner two for each

participant.

We included a variety of factors to explore whether they were confounders of the relative CPM

effect (even if they were unlikely to influence the reliability of the CPM effect) as to our knowledge

this is the first assessment of the reliability of the CPM effect performed in the Achilles tendon and

Triceps Surae musculotendinous junction. To assess for confounding factors, a univariable

generalised estimating equation (GEE) was used to examine the relationship for each of the

dependant variables (Relative CPM effect of the Achilles tendon parallel paradigm, Achilles tendon

sequential paradigm, Triceps Surae musculotendinous junction parallel-paradigm and Triceps Surae

musculotendinous junction sequential paradigm) while controlling for the examiner (factor).

Confounding dichotomous variables (participant sex,(18) examiner was of the opposite sex as the

participant, whether the participant completed regular running-related physical activity,(19) order of

testing or whether the examiner slipped off the Achilles tendon during assessment requiring a

second PPT which may result in additional sensitivity(20)) were treated as factors, with linear

variables (age,(21) BMI,(22) temperature of the CPT, pain induced from the CPT) treated as

covariates in the GEE. Statistically significant variables ($p<0.05$) or clinically significant variables

(Achilles tendon $\beta>5.82$, Triceps Surae musculotendinous junction $\beta>9.34$) were included in a final

GEE model for each of the dependant variables (Appendix A). Clinical significance was determined as

the largest SEM of the intra-rater reliability and for the β of covariates we multiplied by the

difference in the covariate from test one and test two. As per the recommendation of a

biostatistician, only including significant variables within the GEE was determined as the best approach to examine the effect of confounders given our small sample size. Goodness of fit was examined using the Quasi likelihood under Independence Model Criterion (QIC) with a lower the value representing a better fit.

RESULTS

Participants

A total of 28 participants (eight women and 20 men) were recruited with a mean (SD) age of 32.5 (7.9) years, height of 174.0 (10.2) cm, weight of 76.6 (16.8) kg and BMI of (25.1 (4.2) kg/m^2. Twenty (71%) participants reported completing habitual running-related physical activity and eight did not.

Conditioning stimulus

The mean (SD) temperature of the CPT was 4.1 (1.0) degrees Celsius. The mean (SD) NRS pain rating at 60 seconds into the CPT was 6.1 (1.3) points and at 180 seconds was 5.7 (1.8) points as measured on an 11-point NRS-P. No participants removed their hand from the CPT prior to instructions by the examiner at 180 seconds.

Reliability of the Conditioned Pain Modulation Effect

Within-group analysis based on examiners

Data on the CPM effect are included (table one). There were no significant differences between examiners for the temperature of the CS (t= 0.13, p= 0.897), pain due to the CS at either 60 seconds (t= 0.50, p=0.619) or 180 seconds (t= 0.53, p= 0.598) or in baseline Triceps Surae musculotendinous junction PPTs (t= -0.19, p=0.853). There was a significant difference between examiners for the baseline Achilles tendon PPT with examiner one consistently recording a higher PPT value (t= 3.12, p= 0.003). There were no significant differences between baseline PPTs of the Achilles tendon (t=

1.335, p= 0.193) or Triceps Surae musculotendinous junction (t=-0.549, p=0.588) between test one and test two suggesting PPT's had returned to baseline levels following assessment of the CPM effect the first time.

Inter-rater reliability of the Conditioned Pain Modulation Effect

The inter-rater reliability of the absolute and relative CPM effect is detailed in table two with the inter-rater reliability of the dichotomously classified CPM effect detailed in table three. The 95% confidence intervals of the ICCs of the inter-rater reliability of the absolute and relative CPM effect was poor to moderate in the Achilles tendon (ICC= 0.00-0.66). The 95% confidence intervals of the ICCs of the inter-rater reliability of the absolute and relative CPM effect was poor to moderate in the Triceps Surae musculotendinous junction for (ICC= 0.00-0.69). The Kappa of the inter-rater reliability of the dichotomously classified CPM effect were poor for all assessment (Kappa < 0.4).

Levels of agreement

The agreement between examiners for the inter-rater reliability of the CPM effect for the Achilles tendon and Triceps Surae musculotendinous junction is shown in figure three and the data are presented within appendix C. The difference between examiners, and the mean difference between examiners were normally distributed. Bland Altman plots using percentages were used as there was an increase in the variability of the differences between examiner one and examiner two as the magnitude of the measurement increases.(23) These figures demonstrate that there is no significant bias between examiners in regards to agreement with all measures as the line of equality ("0") lies within the confidence intervals of the mean difference. Finally, there was an obvious outlier for each assessment though this was not the same participant in each assessment.

Analysis of Confounders

The influence of confounders on the CPM effect is detailed in appendix B (univariable GEE) and table four (multivariable GEE). Across the different testing paradigms and regions there were several confounders which were significant within the univariate modelling and included within the multi-variate model; examiner, testing order, participant sex and sex differences between examiner and participant, completing regular running related physical activity, pain induced from the CS and temperature of the CS.

There were three testing errors in the Achilles tendon parallel-paradigm (5%), five testing errors in the Achilles tendon sequential-paradigm (9%) and no errors when assessing the Triceps Surae (0%) therefore due to low numbers these were included within the univariable analysis but excluded from any further multivariable analysis. When analysed within a multivariable model, age, BMI, sex, differences in the sex between examiner and the participant, CPT temperature and pain induced by the CPT were not significant. Completing running related physical activity significantly decreased the CPM effect for both the Achilles tendon parallel-paradigm (p= 0.017) and the sequential-paradigm (p= 0.029) but not for the Triceps Surae musculotendinous junction. Testing order showed that being tested second significantly enhanced the CPM effect for parallel-paradigm of the Achilles tendon (p= 0.023) and diminished the parallel-paradigm of the Triceps Surae musculotendinous junction (p= 0.014) but not for sequential-paradigm testing of the Achilles tendon or Triceps Surae musculotendinous junction. Finally, the goodness of fit for all multivariable GEE models were better than the individual univariable GEE models.

DISCUSSION

Investigating the reliability and agreement of the CPM effect is vital in enabling researchers to determine whether multi-centre studies are feasible and compare between studies. This study suggests that multi-centre studies investigating the relative CPM effect may be feasible with differences between examiners not being significant within multivariable GEEs and no bias in the level of agreement between testers. However, due to poor reliability in the dichotomous

classification of the CPM effect this may not be appropriate for multi-centre studies. Additionally, special consideration needs to be given to controlling for potential confounders within statistical modelling(5) and ensuring that the sample size is large enough to be powered to detect confounders (and that these confounders are recorded). This study determined the reliability and level of agreement of the CPM effect for both the Achilles tendon and Triceps Surae musculotendinous junction as well as multiple confounders which can influence the CPM effect.

Reliability and levels of agreement of the conditioned pain modulation effect

The inter-rater reliability ICC 95% confidence intervals for the absolute and relative CPM effect were poor to moderate for both the Achilles tendon and Triceps Surae musculotendinous junction. The inter-rater reliability Kappa for the dichotomously classified CPM effect were poor for both the Achilles tendon and Triceps Surae musculotendinous junction. However, there was no significant bias of the agreement between examiners with the line of equality falling within the confidence intervals of the mean difference. This means that while the CPM effect of the two examiners was different (as shown by the poor reliability with ICCs) the differences were consistent (as shown by the non-significant differences within the Bland Altman plots) with examiner one consistently recording a larger PPT than examiner two. This is important; because the examiners are consistent the differences between measures can be accounted for within statistical modelling as a confounder. However, it was clear from the presence of an outlier within each Bland Altman plot that outliers within the level of agreement exist.

Conditioned pain modulation effect confounders

This study design enabled the observation of several different confounders of the CPM effect when analysed in a multivariable generalised estimating equation. Running related physical activity and testing order were shown to confound the CPM effect and have implications for future study design consideration.

Completing running-related physical activity resulted in a reduction of 22% and 19% for the parallel and sequential-paradigm CPM effect of the Achilles tendon, respectively. Running-related physical activity had no influence on the CPM effect of the Triceps Surae musculotendinous junction. These findings suggest that confounders of the CPM effect may be different depending on the tissue being assessed. This may be due to running, as opposed to other forms of physical activity, such as walking, relying on energy storage and release from the Achilles tendon placing it under more load.(24)

Testing order was also a significant confounder for the parallel-paradigm CPM effect with the second assessment having an increase in the Achilles tendon by 14% and a reduction in the Triceps Surae musculotendinous junction by 19%. Testing order was not a significant confounder for the sequential-paradigm CPM effect though the trends were in the same direction for each test site. These results are interesting given that there were no significant differences between the baseline PPT of test one or test two that would indicate that the CPM effect had washed out. While this may be a spurious finding these results could also represent different tissue responses (muscle versus tendon) which occur as a result of repeated CPT. Physiologically muscle and tendon do respond differently in the presence of pain with tendon pain decreasing with loading exercise (warming up)(24) versus the worsening of muscle injury pain with loading exercise. Given there may be the potential for repeated CPT exposures enhance/ inhibit the CPM effect it is important to therefore account for this finding within any statistical analysis.

While not statistically significant within the multivariate GEE model the pain induced by the CPT and temperature of the CPT may be potentially clinically significant given that for every one-point increase or decrease in pain or temperature there was an associated increase/ decrease of the CPM effect beyond what has been reported as the measurement error.(7) These findings warrant further investigation on how the relationship between the degree of pain, temperature of the CPT and the size of the CPM effect influences test reliability.

Interestingly, when confounders were included within the multivariable GEE differences between examiners were no longer statistically significant. The analysis of confounders is important for studies calculating ICCs of the CPM effect. If you are able to account for different confounders (such as testing order or gender of the tester) it is possible that differences between the CPM effect, as assessed by different examiners, may not be due to different examiners PPT techniques allowing a more accurate explanation of results. Therefore, multicentre CPM research may be possible as differences between examiners were not significant when confounders are controlled for within statistical modelling. Specifically, multivariable models, as opposed to univariable between group comparisons, which account for potential confounders of the CPM effect must be included to be confident that any results are not influenced other variables. Further research with larger sample sizes so that statistical models controlling for confounders can be adequately powered, do not standardise the location of assessment and includes the ethnicity of participants (and subsequently includes them as confounders) are necessary before confidence in multi-site CPM effect testing can be established.

This study was limited by a small sample size (n= 28) and while adequately powered for reliability and comparable to other studies in this field, median sample size of 34 in Kennedy et al. 2016.(5) this study was underpowered to include all variables within the GEE and stepwise elimination of variables based on significance within the model that would have been the preferred method. Finally, this study did standardise the assessment site by marking the participant which decreases it generalisability to multi-site testing.

CONCLUSION

The reliability of the CPM effect between two examiners in this study was poor, however bias in the level of agreement was not significant. Furthermore, when confounders were accounted for, differences between examiners were not significant. This means that while the ICCs were poor, this

is unlikely to be due to assessment technique as no difference between examiners existed in the multivariable model controlling for confounders.

Performing regular running related physical activity was a significant confounder for the Achilles tendon but not the Triceps Surae musculotendinous junction suggesting the CPM effect may differ in people who perform more tissue specific loading (e.g. running to load the Achilles tendon) compared to controls. Testing order was a significant confounder on the parallel paradigm, but not the sequential-paradigm for both the Achilles tendon and Triceps Surae musculotendinous junction even though baseline PPTs had returned to normal levels after the washout and before the second test. This suggests that while PPTs return to baseline levels within 30 minutes of the CS the subsequent parallel-paradigm CPM effect elicited was different to the previous parallel-paradigm CPM effect elicited. Caution should be used when interpreting these findings however as this study had a small sample size. Future studies powered to complete multivariable models using a stepwise elimination of potential confounding variables (including habitual running-related physical activity, testing order, different assessors, sex differences between the assessor and participant, CPT temperature and pain from the CPT) based on significance would assist understanding contributors of the CPM effect.

REFERENCES

1.   Pud D, Granovsky Y, Yarnitsky D. The methodology of experimentally induced diffuse noxious inhibitory control (DNIC)-like effect in humans. Pain. 2009;144(1-2):16-9.

2.   Lewis GN, Rice DA, McNair PJ. Conditioned pain modulation in populations with chronic pain: a systematic review and meta-analysis. The journal of pain : official journal of the American Pain Society. 2012;13(10):936-44.

3.   Damien J, Colloca L, Bellei-Rodriguez C-É, Marchand S. Pain Modulation: From Conditioned Pain Modulation to Placebo and Nocebo Effects in Experimental and Clinical Pain. Int Rev Neurobiol. 2018;139:255-96.

4.   Yarnitsky D, Arendt-Nielsen L, Bouhassira D, Edwards RR, Fillingim RB, Granot M, et al. Recommendations on terminology and practice of psychophysical DNIC testing. European journal of pain (London, England). 2010;14(4):339.

5.   Kennedy DL, Kemp HI, Ridout D, Yarnitsky D, Rice ASC. Reliability of conditioned pain modulation: a systematic review. Pain. 2016;157(11):2410-9.

6.   Nahman-Averbuch H, Granovsky Y, Coghill RC, Yarnitsky D, Sprecher E, Weissman-Fogel I. Waning of "conditioned pain modulation": a novel expression of subtle pronociception in migraine. Headache. 2013;53(7):1104-15.

7.   Locke D, Gibson W, Moss P, Munyard K, Mamotte C, Wright A. Analysis of meaningful conditioned pain modulation effect in a pain-free adult population. The journal of pain : official journal of the American Pain Society. 2014;15(11):1190-8.

8.   Vaegter HB, Petersen KK, Morch CD, Imai Y, Arendt-Nielsen L. Assessment of CPM reliability: quantification of the within-subject reliability of 10 different protocols. Scandinavian journal of pain. 2018;18(4):729-37.

9.   Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ (Clinical research ed). 2013;346:f2304.

10.   Harden M, Friede T. Sample size calculation in multi-centre clinical trials. BMC Medical Research Methodology. 2018;18(1):156.

11.   Jakorinne P, Haanpaa M, Arokoski J. Reliability of pressure pain, vibration detection, and tactile detection threshold measurements in lower extremities in subjects with knee osteoarthritis and healthy controls. Scandinavian journal of rheumatology. 2018;47(6):491-500.

12.   Walton D, MacDermid J, Nielson W, Teasell R, Chiasson M, Brown L. Reliability, Standard Error, and Minimum Detectable Change of Clinical Pressure Pain Threshold Testing in People With and Without Acute Neck Pain. Journal of Orthopaedic & Sports Physical Therapy. 2011;41(9):644-50.

13.   Eckert NR, Vierck CJ, Simon CB, Calderon S, Cruz-Almeida Y, Staud R, et al. Methodological Considerations for the Temporal Summation of Second Pain. The journal of pain : official journal of the American Pain Society. 2017;18(12):1488-95.

14.   Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Statistics in Medicine. 1998;17(1):101-10.

15.   Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63.

16.   Demetrashvili N, Wit EC, van den Heuvel ER. Confidence intervals for intraclass correlation coefficients in variance components models. Statistical methods in medical research. 2016;25(5):2359-76.

17.   Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet (London, England). 1986;1(8476):307-10.

18.   Martel MO, Wasan AD, Edwards RR. Sex differences in the stability of conditioned pain modulation (CPM) among patients with chronic pain. Pain medicine (Malden, Mass). 2013;14(11):1757-68.

19.   Umeda M, Lee W, Marino CA, Hilliard SC. Influence of moderate intensity physical activity levels and gender on conditioned pain modulation. Journal of sports sciences. 2016;34(5):467-76.

20.   Valencia C, Gay C, Zeppieri G, Jr., George S. Inducing temporal summation: the importance of relevant stimulus parameters. The Journal of Pain. 2013;14(4):S14.

21.   Grashorn W, Sprenger C, Forkmann K, Wrobel N, Bingel U. Age-dependent decline of endogenous pain control: exploring the effect of expectation and depression. PLoS One. 2013;8(9):e75629.

22.   Tashani OA, Astita R, Sharp D, Johnson MI. Body mass index and distribution of body fat can influence sensory detection and pain sensitivity. European journal of pain (London, England). 2017;21(7):1186-96.

23.   Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb). 2015;25(2):141-51.

24.   Cook JL, Rio E, Purdam CR, Docking SI. Revisiting the continuum model of tendon pathology: what is its merit in clinical practice and research? Br J Sports Med. 2016;50(19):1187-91.

FIGURE LEGENDS

Figure 1. Assessment landmarks

Figure 2. Conditioned pain modulation effect assessment procedure

Figure 3. Bland Altman plots of the level of agreement for the conditioned pain modulation effect of the Achilles tendon where the short dashed line represents the mean percentage difference, the long dashed line represents the confidence intervals for the mean percentage difference and the dotted line represented the 95% standard deviation of the mean percentage difference. A) Achilles tendon parallel-paradigm CPM effect, B) Achilles tendon sequential-paradigm CPM effect, C) Triceps Surae musculotendinous junction parallel-paradigm CPM effect, and D) Triceps Surae musculotendinous junction sequential-paradigm CPM effect.