Pain pressure threshold algometry in knee osteoarthritis: intra- and inter-rater reliability

Martin Bjørn Stausholm, Jan Magnus Bjordal, Rolf Moe-Nilssen & Ingvill Fjell Naterstad

Published online: 12 Jan 2022.

Submit your article to this journal ⬈

View related articles ⬈

View Crossmark data ⬈

Taylor & Francis
Taylor & Francis Group

Pain pressure threshold algometry in knee osteoarthritis: intra- and inter-rater reliability

Martin Bjørn Stausholm, MSc, PT [ID], Jan Magnus Bjordal, PhD, PT [ID], Rolf Moe-Nilssen, PhD, PT [ID], and Ingvill Fjell Naterstad, MSc, PT [ID]

Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

ABSTRACT

Introduction: Synovitis and effusion can cause pain sensitization in persons with knee osteoarthritis (KOA). Pain Pressure Threshold (PPT) algometry is a means to quantify somatosensory abnormalities, including inflammatory-mediated pressure hyperalgesia. We investigated the reliability of PPT algometry with three raters.

Methods: Twenty-seven persons (50 knees) with KOA, according to the American College of Rheumatology criteria, were included. The PPT of the most tender spot in the joint line of each knee, identified by palpation, was assessed using a digital pressure algometer with a round 1 cm^2 rubber tip. The algometer was applied three times with at least twenty-second intervals by three physiotherapists each in a single session. Two of the physiotherapists had no experience with the procedure prior to the study. We estimated the Intraclass Correlation Coefficient (ICC) model 2.1, 95% within-subject standard deviation (s_w), and Minimal Detectable Difference (MDD).

Results: The mean PPTs ranged from 39.94 to 41.81 Newton (N), the intra-rater ICC ranged from 0.909 to 0.956, the s_w ranged from 6.44 to 10.77 N, and the related MDD ranged from 9.11 to 15.23 N. The three raters achieved an inter-rater ICC of 0.707, an s_w of 17.68 N, and an MDD of 25.01 N. The results were homoscedastic.

Conclusions: Our results indicate that PPT algometry is a suitable method for assessment of pain in osteoarthritic knees. After a short session of PPT procedure training, good intra-rater and acceptable inter-rater ICCs were achieved.

Introduction

Pain is the dominating knee osteoarthritis (KOA) complaint. The presence of inflammation, meniscal extrusion (i.e., pathologically displaced medial meniscus), osteophytes, and bone marrow lesions of the knee are associated with more intense KOA pain (Cicuttini, Baker, Hart, and Spector, 1996; Heidari, 2011; Hunter et al., 2013; Roubille et al., 2014; Yusuf et al., 2011). Furthermore, persistent inflammation can cause both local and widespread pain sensitization in persons with KOA (Neogi et al., 2016; Suokas et al., 2012). Therefore, therapeutically targeting inflammation early could prove valuable in the management of the disease (Neogi et al., 2016).

Palpation tenderness can provide information about physical damage and level of inflammation (Bjordal, Lopes-Martins, and Iversen, 2006). Unfortunately, finger palpation is difficult to standardize and has moderate sensitivity (Cook et al., 2001; Ramos et al., 2009). However, the Pain Pressure Threshold (PPT) can be quantified using an algometer device. A numerical value is displayed on the algometer with a lower value representing less pressure (Maquet, Croisier, Demoulin, and Crielaard, 2004). Pain is subjective and dependent on individual differences in physiological, emotional, and cognitive states. Nevertheless, somatosensory abnormalities, including inflammatory-mediated pressure hyperalgesia, in knees can potentially be detected with PPT algometry. In a cohort of 1,111 persons with or at risk of KOA, Neogi et al. (2016) found that knee inflammation as evidenced by synovitis and effusion identified with Magnetic Resonance Imaging was associated with lower PPTs. Furthermore, Neogi et al. (2016) discovered that the presence of synovitis was a predictor of decreased PPT two years later. In line with these findings, Dina, Green, and Levine (2008) found that higher levels of intramuscular interleukin-6 and prostaglandin E$_2$ (markers of inflammation) are associated with lower PPTs in vivo. Furthermore, low pre-operative PPTs seem to be associated with more intense pain after knee replacement (Arendt-Nielsen et al., 2018; Leung et al., 2019; Wylde, Palmer, Learmonth, and Dieppe, 2013). There is also

---

evidence that in persons with KOA, lower PPTs are associated with increased self-reported disability (Imamura et al., 2008; Kuni et al., 2015); pain (Imamura et al., 2008); anxiety (Urban, Eyles, Hunter, and Mills, 2018); and reduced quality of life (Imamura et al., 2008). It is also noteworthy that lower PPT has been found to be associated with higher age (Lautenbacher et al., 2005) and female gender (Chesterton et al., 2003).

Assessment of reliability is a necessary first step in the validation procedures of clinical tests. The reliability of the PPT measurement is susceptible to the influence of rater behavior and judgment, such as the instructions to the participant, rate of force application, and reaction time of the rater (Moe-Nilssen, Nordin, and Lundin-Olsson, 2008). The reliability of PPT in persons with KOA has been investigated in several studies. The intra-rater reliability was found to be good (Interclass Correlation Coefficient (ICC) ≥0.900) by Alahmari et al. (2020), Osgood et al. (2015), Mutlu and Ozdincler (2015), and Wessel (1995); and acceptable (ICC ≥ 0.700) by Jakorinne, Haanpaa, and Arokoski (2018). However, no attempt to manage rater blinding during each measurement has been described in any of the reports and only Mutlu and Ozdincler (2015) specified the ICC model used.

Inter-rater reliability of PPT algometry in persons with KOA has, to our knowledge, only been investigated by Alahmari et al. (2020), Osgood et al. (2015), and Jakorinne, Haanpaa, and Arokoski (2018) and never with more than two raters per study. In the study by Jakorinne, Haanpaa, and Arokoski (2018), the PPT values decreased significantly during the sessions and the authors hypothesized that this was caused by a relatively short (≥10 seconds) pause between each measurement. Therefore, we opted to investigate the intra- and inter-rater reliability of PPT in persons with KOA with three raters, rater blinding, and a pause of ≥20 seconds between each measurement. We hypothesized that even physiotherapists with no former experience with the procedure can master it with good reliability after a single 30-min training session.

Methods

This cross-sectional clinical study was approved by the Research Ethics Committee North (reference 2017/2417). All the participants signed an informed consent form before entering the study.

Subjects

The persons enrolled in the study were recruited from the Bergen municipality (Norway) through written and verbal advertisement. They were a convenience sample from an ongoing interventional trial. The inclusion criteria of the trial were women and men aged ≥50 years and KOA according to the American College of Rheumatology criteria, that is, knee pain and at least three of the following: ≥50 years old, ≤30 minutes of morning stiffness, crepitus on active motion, bony tenderness, bony enlargement, and no palpable warmth of synovia. The exclusion criteria were knee alloplastic, total meniscectomy, intra-articular steroid injection and/or oral steroid treatment within the last six months, cancer, rheumatoid arthritis, severe cognitive deficit, neurological deficits in the lower limb, and inability to speak and understand English/Nordic.

PPT assessment procedure

All the knees of the 27 participants (54 knees) were tested for PPT using a digital algometer (Wagner FPX 25) with a round 1 cm^2 rubber tip, starting with the right. However, only the knees with a KOA diagnosis (50 knees) were included in the analyses.

Three physiotherapists, one female (A) and two males (B and C), conducted the measurements using a standardized protocol. The raters practiced the procedure together in a 30-min training session on a person with KOA, before the study started. The rater and participant were seated during the testing. The rater stabilized the participant's knee with one hand. The most tender spot in the joint line of each knee identified by palpation was assessed with PPT algometry three times with ≥20-second intervals by each rater in a single session. The rubber tip was placed perpendicular to the skin. The participants were instructed to give a verbal signal as soon as the sensation of pressure turned into pain, at which time the rater immediately removed the algometer and recorded the score. The rate of pressure application was not fixed, since computerized PPT measurement has shown to be less reliable and sensitive compared to manual PPT measurement (Koo, Guo, and Brown, 2013). The display of the algometer faced the floor during the testing to blind the raters and participants for the levels of force. There was only one rater and participant present during the testing at a time. The pause between each rater was approximately one minute and the rater order changed randomly during the study period. The raters were unaware of each other's test results. Furthermore, the participants were not informed of their results.

Rater A and B had no former experience with PPT assessment of knees, but they had been working as clinicians for 5 and 18 years, respectively. Rater C had only 1 year of experience as a therapist; however, he had practiced the procedure in the ongoing interventional trial.

Statistics

Descriptive statistics were applied using IBM SPSS Statistics 25 and Microsoft Excel 2016. The first measurements of all the knees were excluded from the analysis as it is usually the least reliable in a series of three PPT measurements (Nussbaum and Downes, 1998). Intra-rater reliability was estimated using the second and third measurements and inter-rater reliability was estimated using the mean scores of the second and third measurements. Relative reliability was estimated using Intraclass Correlation Coefficient (ICC) two-way random model 2.1 since the raters were randomly selected from a population of physiotherapists (Koo, Guo, and Brown, 2013).

We interpreted the relative reliability estimates as proposed by Nunnally and Bernstein (1994), that is, ICC values of ≥0.7 and ≥0.9 represent acceptable and good reliability, respectively.

Absolute reliability was calculated using within-subject standard deviation (S_w), sometimes referred to as Standard Error of Measurement (SEM); the difference between a measurement and the true value can be expected to be less than $1.96 \times S_w$ for 95% of observations. The Minimal Detectable Difference (MDD) in pressure that must be exceeded to be 95% confident that a real change has occurred between measurements was estimated using the formula $1.96 \times S_w \times \sqrt{2}$ (Bland and Altman, 1996). The distribution of data was inspected using Bland–Altman plots with means and differences of paired measurements and 95% limits of agreement (Giavarina, 2015).

Results

Characteristics of the participants are described in Table 1. The mean PPTs ranged from 39.94 to 41.81 Newton (N), the intra-rater ICC ranged from 0.909 to 0.956, the s_w ranged from 6.44 to 10.77 N, and the related

MDD ranged from 9.11 to 15.23 N. The three raters achieved an inter-rater ICC of 0.707, a s_w of 17.68 N and a MDD of 25.01 N (Table 2). There was no difference in the mean PPT between the second and third measurements, indicating that no temporal summation occurred. The Bland–Altman plots revealed a neglectable bias in the intra- and inter-rater results (Figure 1–6).

Discussion

In this study, three physiotherapists PPT assessed the knees of KOA patients with good intra-rater and acceptable inter-rater ICCs after a single 30-min training session. Two of the physiotherapists had no former experience with the procedure. It is also important to note that the MDD was twice as large in the inter-rater assessments (25.01 N) compared to in the intra-rater assessments (9.11–15.23 N). The intra- and inter-rater MDD corresponded to ca. 30% and 60% of the mean PPT scores, respectively. Whether the measurement errors are adequate depends on the context in which the measurements are being used, including the analytical goals of the user (Atkinson and Nevill, 1998; Bruton, Conway, and Holgate, 2000). The Bland–Altman plots indicated that there was no association between the size of the scores and variability (heteroscedasticity), which is a prerequisite for estimating absolute reliability by s_w and MDD (Bland and Altman, 1996).

In the reliability study by Jakorinne, Haanpaa, and Arokoski (2018), the PPT values decreased significantly during the sessions; however, this did not occur in our testing, perhaps because we waited longer between each measurement. This may be a reason why Jakorinne, Haanpaa, and Arokoski (2018) did not achieve good intra-rater and acceptable inter-rater ICCs.

Alahmari et al. (2020) reported slightly higher intra-rater reliability and substantially higher absolute reliability than we achieved, however, they did not specify

Table 1. Characteristics of the participants.

Gender	
Women	20 (74%)
Men	7 (26%)
Age (min-max)	65.07 (51–79) years
Height	1.71 meter (SD = 0.085)
Unilateral knee osteoarthritis	4 persons
Bilateral knee osteoarthritis	23 persons
Bony enlargement	20 knees
Most tender spot in joint line	
Medial side	42 knees
Lateral side	8 knees
Duration of knee pain	73.01 months (SD = 99.18)
KOOS pain	56.88 (SD = 19.59)
Use of analgesics in the previous 7 days	10 persons

KOOS = Knee injury and Osteoarthritis Outcome Score; NSAIDs = Non-Steroidal Anti-Inflammatory Drugs; SD = standard deviation. The KOOS pain scores ranges from 0–100 (higher score is better).

Table 2. Intra- and inter-rater reliability results of PPT algometry in persons with KOA.

Rater	ICC (95% CI)	Mean (N)	95% CI of true value (N)	MDD (N)
Intra-rater reliability				
A	0.909 (0.844–0.948)	40.16	±9.79	13.84
B	0.956 (0.924–0.975)	41.81	±6.44	9.11
C	0.914 (0.853–0.950)	39.94	±10.77	15.23
Inter-rater reliability				
ABC	0.707 (0.581–0.809)	40.63	±17.68	25.01
AB	0.707 (0.537–0.822)	41.28	±16.79	23.74
AC	0.718 (0.550–0.830)	40.85	±17.91	25.33
BC	0.695 (0.520–0.815)	41.12	±18.31	25.90

CI = Confidence Interval; ICC = Intraclass Correlation Coefficient; MDD = Minimal Detectable Difference; N = Newton.

Difference against mean for PPT2 and PPT3 by rater A
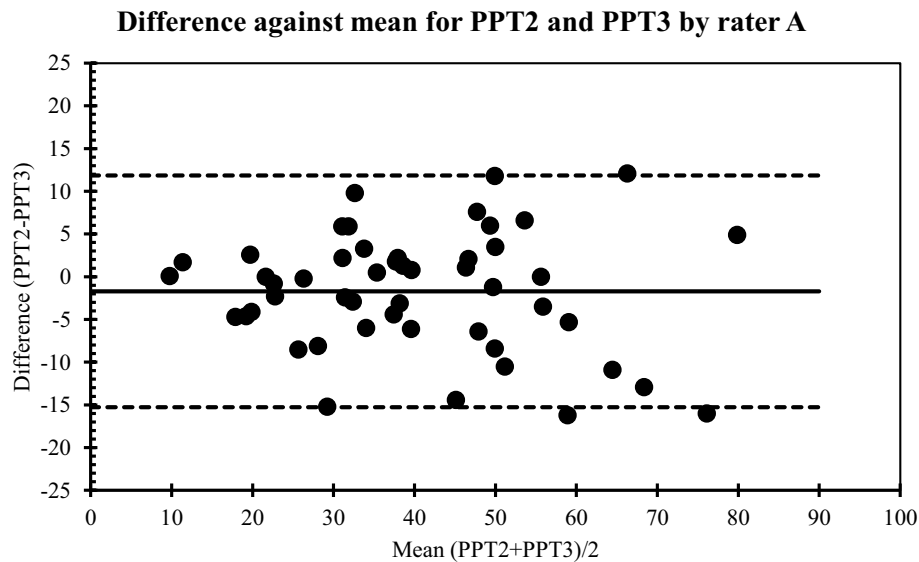
Figure 1. Level of agreement between rater A's 2. and 3. measurements. *Note.* The values are Newton. The thick horizontal solid line represents the mean difference and the dotted horizontal lines represent the 95% limits of agreement; PPT = Pain Pressure Threshold.

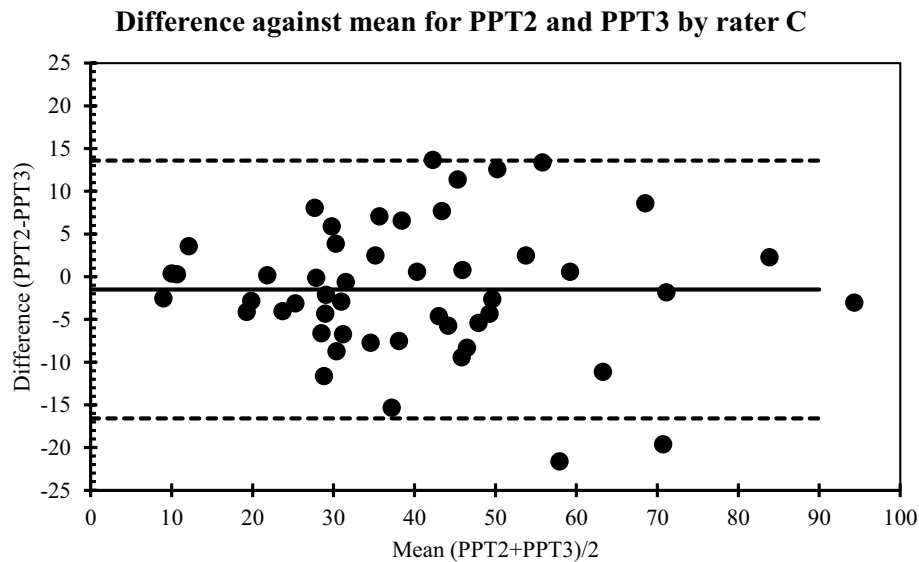## Difference against mean for PPT2 and PPT3 by rater B
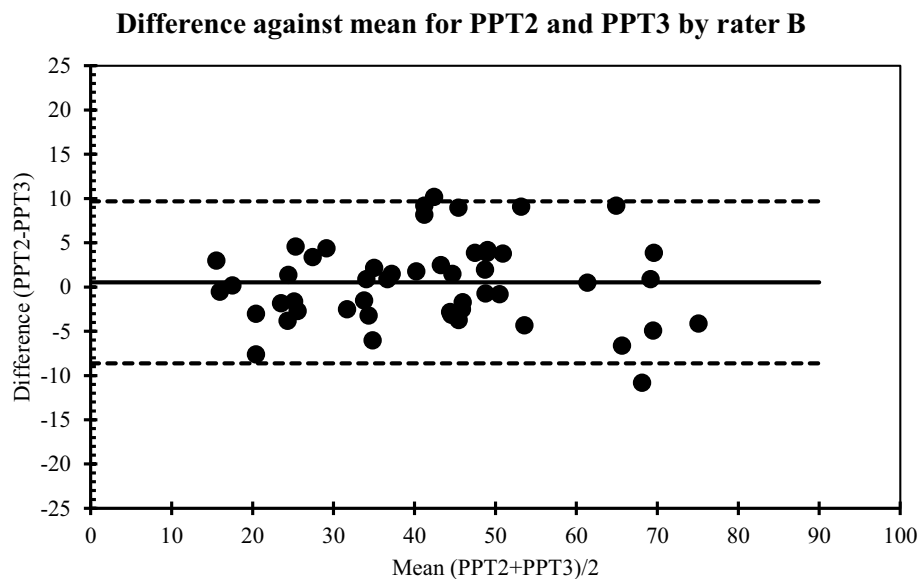
Figure 2. Level of agreement between rater B's 2. and 3. measurements. *Note.* The values are Newton. The thick horizontal solid line represents the mean difference and the dotted horizontal lines represent the 95% limits of agreement; PPT = Pain Pressure Threshold.

the ICC model used, and this is problematic as different ICC models can produce different reliability estimates (Koo, Guo, and Brown, 2013). Moreover, Alahmari et al. (2020) included the ICC values from the unspecified statistical model in the estimation of SEM and MDD.

Our intra- and inter-rater ICCs are similar to those by Osgood et al. (2015). Interestingly, the reliability results by Osgood et al. (2015) were achieved by two raters who practiced the procedure for several months prior to the assessments, whereas the raters in our study only participated in a 30-min PPT training session. Still, we believe that our inter-rater results could have been

improved by additional rater training. We assessed the most tender spot in the joint line of the knee identified by palpation for PPT since the most problematic site of the knee varies between persons with KOA. This is a novel approach.

We opted to assess a total of 50 osteoarthritic knees in a sample of 27 persons with KOA as this would provide a reasonable number of dots in the Bland–Altman plot to estimate the level of agreement (de Vet, Terwee, Mokkink, and Knol, 2011). Furthermore, we assumed that the raters would achieve ICC values of 0.800 and according to the formula provided by Giraudeau and

Difference against mean for PPT2 and PPT3 by rater C

Figure 3. Level of agreement between rater C's 2. and 3. measurements. *Note.* The values are Newton. The thick horizontal solid line represents the mean difference and the dotted horizontal lines represent the 95% limits of agreement; PPT = Pain Pressure Threshold.

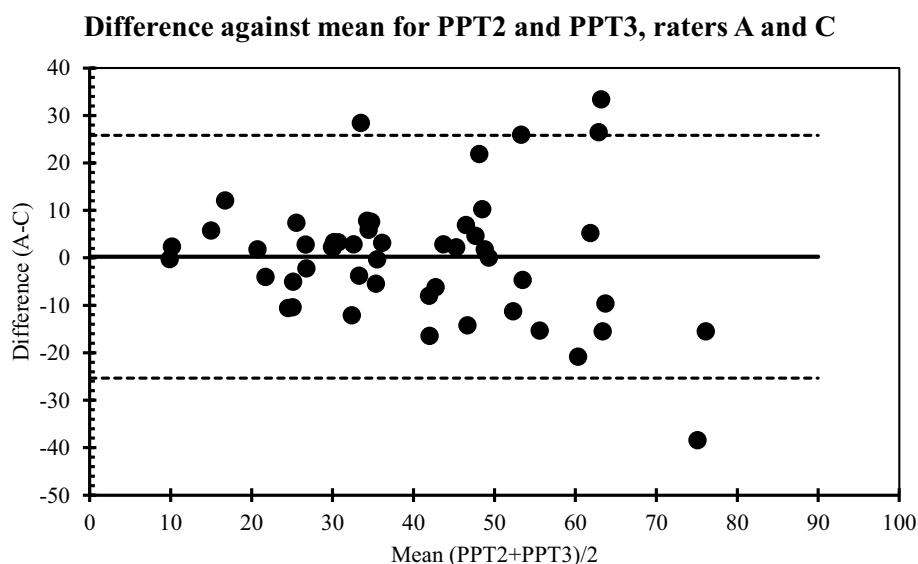## Difference against mean for PPT2 and PPT3, raters A and B
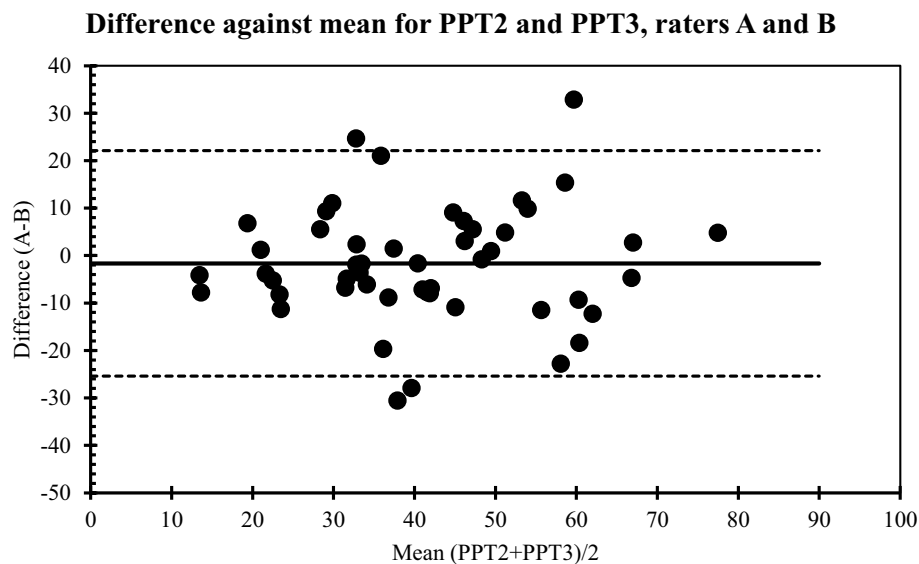
Figure 4. Level of agreement between rater A's and B's 2.-3. measurements. *Note.* The values are Newton. The thick horizontal solid line represents the mean difference and the dotted horizontal lines represent the 95% limits of agreement; PPT = Pain Pressure Threshold.

Mary (2001), 95% confidence interval around ICC point values of 0.800 can be expected to be ±0.1, which is a range from acceptable to good relative reliability. Of note, if the confidence interval was to be halved, it would take four times as many participants (Giraudeau and Mary, 2001).

The assessment by the first rater left a visible pressure mark on the skin, which allowed the other raters to select the same area. This phenomenon has been described as a study limitation, as it could result in relatively higher inter-rater reliability (Sæbø et al., 2019); however, this is merely the equivalent to marking the skin area with a pen for the purpose of reassessments in clinical practice and trials and should therefore not be considered a potential bias.

Somatosensory abnormalities, including inflammatory-mediated pressure hyperalgesia in knees, can be monitored with PPT algometry (Dina, Green, and Levine, 2008; Neogi et al., 2016) with adequate reliability. Further prediction and concurrent validity studies on the topic would provide valuable information regarding the usefulness of the assessment.

Difference against mean for PPT2 and PPT3, raters A and C

Figure 5. Level of agreement between rater A's and C's 2.-3. measurements. *Note.* The values are Newton. The thick horizontal solid line represents the mean difference and the dotted horizontal lines represent the 95% limits of agreement; PPT = Pain Pressure Threshold.

Difference against mean for PPT2 and PPT3, raters B and C
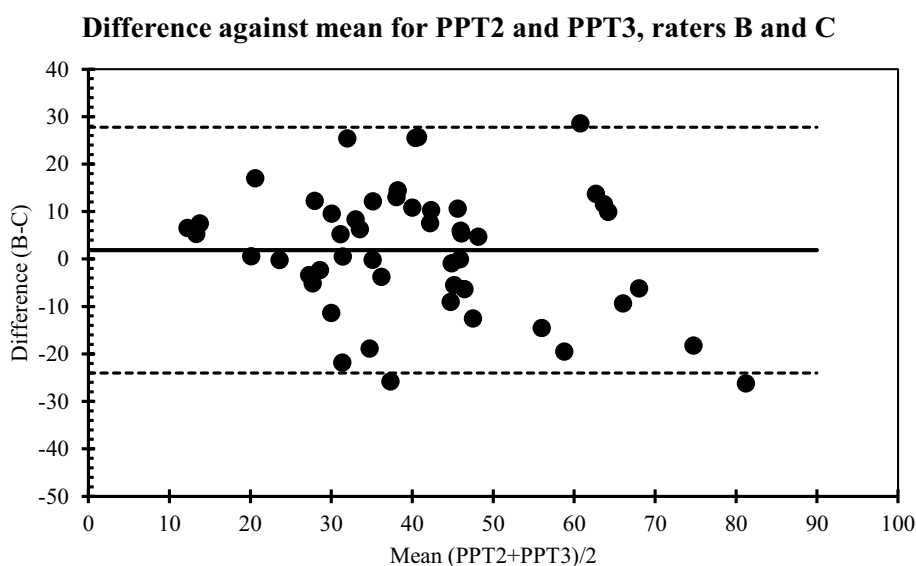
Figure 6. Level of agreement between rater B's and C's 2.-3. measurements. *Note.* The values are Newton. The thick horizontal solid line represents the mean difference and the dotted horizontal lines represent the 95% limits of agreement; PPT = Pain Pressure Threshold.

Limitations of the study

As the study concerned rater reliability, all the assessments were completed in a single session. Therefore, the raters would have a good sense of how much pressure was applied in the first measurement. However, if the assessments were completed on different occasions, a change in symptoms of the participants could possibly have biased the results.

We only evaluated the reliability of PPT measurements in a single spot. It is plausible that PPT assessment of the suprapatellar recess could give additional relevant insight into the inflammatory status of the knee as there are no osteophytes and meniscus in this area.

Practical implications

The participants with KOA tolerated nine consecutive PPT measurements of the most tender spot in the knee joint line well. Physiotherapists with no former experience in the assessment procedure were capable of applying it with good intra-rater and acceptable inter-rater

relative reliability after a 30-minute training session. Interchanging between PPT raters may double the measurement errors.

Conclusions

Our results indicate that PPT algometry is a suitable method for assessment of pain in osteoarthritic knees. After a short session of PPT procedure training, good intra-rater and acceptable inter-rater ICCs were achieved.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

ORCID

Martin Bjørn Stausholm MSc, PT http://orcid.org/0000-0001-9869-0705
Jan Magnus Bjordal PhD, PT http://orcid.org/0000-0002-4804-4366
Rolf Moe-Nilssen PhD, PT http://orcid.org/0000-0002-2167-4974
Ingvill Fjell Naterstad MSc, PT http://orcid.org/0000-0002-3619-4578

References

Alahmari K, Silvian SP, Ahmad I, Reddy RS, Kakaraparthi VN 2020 Subjective and objective evaluation of pain for older adults with knee osteoarthritis in Saudi Arabia: A reliability study. Nigerian Journal of Clinical Pratice 237: 934–943. doi:10.4103/njcp.njcp_270_19

Arendt-Nielsen L, Simonsen O, Laursen M, Roos EM, Rathleff M, Rasmussen S, Skou ST 2018 Pain and sensitization after total knee replacement or nonsurgical treatment in patients with knee osteoarthritis: Identifying potential predictors of outcome at 12 months. European Journal of Pain 22: 1088–1102. doi:10.1002/ejp.1193

Atkinson G, Nevill AM 1998 Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Medicine 264: 217–238. doi:10.2165/00007256-199826040-00002

Bjordal JM, Lopes-Martins RA, Iversen VV 2006 A randomised, placebo controlled trial of low level laser therapy for activated Achilles tendinitis with microdialysis measurement of peri-tendinous prostaglandin E2 concentrations. British Journal of Sports Medicine 40: 76–80. doi:10.1136/bjsm.2005.020842

Bland JM, Altman DG 1996 Statistics notes: Measurement error. British Medical Journal 3137059: 744. doi:10.1136/bmj.313.7059.744

Bruton A, Conway J, Holgate S 2000 Reliability: What is it, and how is it measured? Physiotherapy 862: 94–99. doi:10.1016/S0031-9406(05)61211-4

Chesterton LS, Barlas P, Foster NE, Baxter DG, Wright CC 2003 Gender differences in pressure pain threshold in healthy humans. Pain 1013: 259–266. doi:10.1016/S0304-3959(02)00330-5

Cicuttini FM, Baker J, Hart DJ, Spector TD 1996 Association of pain with radiological changes in different compartments and views of the knee joint. Osteoarthritis and Cartilage 42: 143–147. doi:10.1016/S1063-4584(05)80323-1

Cook JL, Khan KM, Kiss ZS, Purdam CR, Griffiths L 2001 Reproducibility and clinical utility of tendon palpation to detect patellar tendinopathy in young basketball players. Victorian Institute of Sport tendon study group. British Journal of Sports Medicine 351: 65–69. doi:10.1136/bjsm.35.1.65

de Vet HC, Terwee CB, Mokkink LB, Knol DL 2011 Measurement in medicine: A practical guide. Cambridge: Cambridge University Press.

Dina OA, Green PG, Levine JD 2008 Role of interleukin-6 in chronic muscle hyperalgesic priming. Neuroscience 1522: 521–525. doi:10.1016/j.neuroscience.2008.01.006

Giavarina D 2015 Understanding Bland Altman analysis. Biochemia Medica 252: 141–151. doi:10.11613/BM.2015.015

Giraudeau B, Mary JY 2001 Planning a reproducibility study: How many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient. Statistics in Medicine 20: 3205–3214. doi:10.1002/sim.935

Heidari B 2011 Knee osteoarthritis prevalence, risk factors, pathogenesis and features - Part 1. Caspian Journal of Internal Medicine 2: 205–212.

Hunter DJ, Guermazi A, Roemer F, Zhang Y, Neogi T 2013 Structural correlates of pain in joints with osteoarthritis. Osteoarthritis and Cartilage 219: 1170–1178. doi:10.1016/j.joca.2013.05.017

Imamura M, Imamura ST, Kaziyama H, Targino RA, Hsing W, de Souza LP, Cutait MM, Fregni F, Camanho GL 2008 Impact of nervous system hyperalgesia on pain, disability, and quality of life in patients with knee osteoarthritis: A controlled analysis. Arthritis Care and Research 5910: 1424–1431. doi:10.1002/art.24120

Jakorinne P, Haanpaa M, Arokoski J 2018 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 476: 491–500. doi:10.1080/03009742.2018.1433233

Koo TK, Guo JY, Brown C 2013 Test-retest reliability, repeatability, and sensitivity of an automated deformation-controlled indentation on pressure pain threshold measurement. Journal of Manipulative Physiological Therapeutics 362: 84–90. doi:10.1016/j.jmpt.2013.01.001

Kuni B, Wang H, Rickert M, Ewerbeck V, Schiltenwolf M 2015 Pain threshold correlates with functional scores in osteoarthritis patients. Acta Orthopaedica 86: 215–219.

Lautenbacher S, Kunz M, Strate P, Nielsen J, Arendt-Nielsen L 2005 Age effects on pain thresholds, temporal summation and spatial summation of heat and pressure pain. Pain 1153: 410–418. doi:10.1016/j.pain.2005.03.025

Leung YY, Lim Z, Fan Q, Wylde V, Xiong S, Yeo S, Lo N, Chong H, Yeo W, Tan M, et al. 2019 Pre-operative pressure pain thresholds do not meaningfully explain satisfaction or improvement in pain after knee replacement: A cohort study. Osteoarthritis and Cartilage 271: 49–58. doi:10.1016/j.joca.2018.09.003

Maquet D, Croisier JL, Demoulin C, Crielaard JM 2004 Pressure pain thresholds of tender point sites in patients with fibromyalgia and in healthy controls. European Journal of Pain 8: 111–117. doi:10.1016/S1090-3801(03)00082-X

Moe-Nilssen R, Nordin E, Lundin-Olsson L 2008 Criteria for evaluation of measurement properties of clinical balance measures for use in fall prevention studies. Journal of Evaluation in Clinical Practice 142: 236–240. doi:10.1111/j.1365-2753.2007.00839.x

Mutlu E, Ozdincler A 2015 Reliability and responsiveness of algometry for measuring pressure pain threshold in patients with knee osteoarthritis. Journal of Physical Therapy Science 276: 1961–1965. doi:10.1589/jpts.27.1961

Neogi T, Guermazi A, Roemer F, Nevitt MC, Scholz J, Arendt-Nielsen L, Woolf C, Niu J, Bradley LA, Quinn E 2016 Quinn E et al 2016 Association of joint inflammation with pain sensitization in knee osteoarthritis: The Multicenter Osteoarthritis Study. Arthritis and Rheumatology 683: 654–661. doi:10.1002/art.39488

Nunnally JC, Bernstein IH 1994 Psychometric Theory. 3rd, New York: McGraw-Hill.

Nussbaum EL, Downes L 1998 Reliability of clinical pressure-pain algometric measurements obtained on consecutive days. Physical Therapy 782: 160–169. doi:10.1093/ptj/78.2.160

Osgood E, Trudeau J, Eaton T, Jensen MP, Gammaitoni A, Simon LS, Katz N 2015 Development of a bedside pain assessment kit for the classification of patients with osteoarthritis. Rheumatology International 356: 1005–1013. doi:10.1007/s00296-014-3191-z

Ramos LA, Carvalho RT, Garms E, Navarro MS, Abdalla RJ, Cohen M 2009 Prevalence of pain on palpation of the inferior pole of the patella among patients with complaints of knee pain. Clinics 64: 199–202. doi:10.1590/S1807-59322009000300009

Roubille C, Raynauld JP, Abram F, Paiement P, Dorais M, Delorme P, Bessette L, Beaulieu AD, Martel-Pelletier J, Pelletier JP 2014 The presence of meniscal lesions is a strong predictor of neuropathic pain in symptomatic knee osteoarthritis: A cross-sectional pilot study. Arthritis Research and Therapy 16: 507. doi:10.1186/s13075-014-0507-z

Sæbø H, Naterstad IF, Stausholm MB, Bjordal JM, Joensen J 2019 Reliability of pain pressure threshold algometry in persons with conservatively managed wrist fractures. Physiotherapy Research International 25: e1797.

Suokas AK, Walsh DA, McWilliams DF, Condon L, Moreton B, Wylde V, Arendt-Nielsen L, Zhang W 2012 Quantitative sensory testing in painful osteoarthritis: A systematic review and meta-analysis. Osteoarthritis and Cartilage 20: 1075–1085. doi:10.1016/j.joca.2012.06.009

Urban H, Eyles JP, Hunter DJ, Mills K 2018 The relationship between pressure pain thresholds and anxiety in patellofemoral osteoarthritis: Exploratory data. Osteoarthritis and Cartilage 26: S354. doi:10.1016/j.joca.2018.02.701

Wessel J 1995 The reliability and validity of pain threshold measurements in osteoarthritis of the knee. Scandinavian Journal of Rheumatology 24: 238–242. doi:10.3109/03009749509100881

Wylde V, Palmer S, Learmonth ID, Dieppe P 2013 The association between pre-operative pain sensitisation and chronic pain after knee replacement: An exploratory study. Osteoarthritis and Cartilage 219: 1253–1256. doi:10.1016/j.joca.2013.05.008

Yusuf E, Kortekaas MC, Watt I, Huizinga T, Kloppenburg M 2011 Do knee abnormalities visualised on MRI explain knee pain in knee osteoarthritis? A systematic review. Annals of the Rheumatic Diseases 70: 60–67. doi:10.1136/ard.2010.131904