Development of a bedside tool kit for assessing sensitization in patients with chronic osteoarthritis knee pain or chronic knee pain after total knee replacement

Sachau, Juliane; Otto, Jan Carl; Kirchhofer, Viktoria; Larsen, Jesper Bie; Kennes, Lieven Nils; Hüllemann, Philipp; Arendt-Nielsen, Lars; Baron, Ralf

Published in:
Pain

DOI (link to publication from Publisher):
10.1097/j.pain.0000000000002335

Publication date:
2022

Document Version
Accepted author manuscript, peer reviewed version

PAIN Publish Ahead of Print
DOI: 10.1097/j.pain.0000000000002335

Development of a bedside tool kit for assessing sensitization in patients with chronic osteoarthritis knee pain or chronic knee pain after total knee replacement

Juliane Sachau[1*], Jan Carl Otto[1*], Viktoria Kirchhofer[1], Jesper Bie Larsen[2], Lieven Nils Kennes[3], Philipp Hüllemann[1], Lars Arendt-Nielsen[2], Ralf Baron[1]

[1]Division of Neurological Pain Research and Therapy, Department of Neurology, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

[2]Center for Neuroplasticity and Pain & Sport Sciences – Performance and Technology, Department of Health Science and Technology, School of Medicine, Aalborg University, Aalborg, Denmark, [3]Department of Economics and Business Administration, University of Applied Sciences Stralsund, Stralsund, Germany

*contributed equally

Number of pages: 27, figures: 4, tables: 5, Supplement Digital Content: 3

Corresponding author:
Juliane Sachau
Sektion Neurologische Schmerzforschung und -therapie
Klinik für Neurologie, Universitätsklinikum Schleswig-Holstein, Campus Kiel
Arnold-Heller-Straße 3 (Haus D)
24105 Kiel
Tel: +49 431 500 23920
Fax: +49 431 400 23914
E-mail: juliane.sachau@uksh.de

Abstract

Different pathophysiological mechanisms contribute to the pain development in osteoarthritis (OA). Sensitization mechanisms play an important role in the amplification and chronification of pain and may predict the therapeutic outcome. Stratification of patients according to their pain mechanisms could help to target pain therapy. This study aimed at developing an easy-to-use, bedside tool-kit to assess sensitization in patients with chronic painful knee OA or chronic pain after total knee replacement (TKR).

In total, 100 patients were examined at the most affected knee and extra-segmentally by use of four standardized quantitative sensory testing parameters reflecting sensitization (mechanical pain threshold, mechanical pain sensitivity, dynamic mechanical allodynia, pressure pain threshold), a bedside testing battery of equivalent parameters including also temporal summation and conditioned pain modulation, and pain questionnaires. Machine learning techniques were applied to identify an appropriate set of bedside screening tools.

Approximately half of the patients showed signs of sensitization (46%). Based on machine learning techniques a composition of tests consisting of three modalities were developed. The most adequate bedside tools to detect sensitization were pressure pain sensitivity (pain intensity at 4 ml pressure using a 10 ml blunted syringe), mechanical pinprick pain sensitivity (pain intensity of a 0.7 mm nylon-filament) over the most affected knee, and extra-segmental pressure pain sensitivity (pain threshold).

This pilot study presents a first attempt to develop an easy-to-use bedside test to probe sensitization in patients with chronic OA knee pain or chronic pain after TKR. This tool may be used to optimize individualized, mechanism-based pain therapy.

Keywords: osteoarthritis, neuropathic pain, bed-side tool, machine learning, sensitization,quantitative sensory testing

Introduction

Osteoarthritis (OA) is a degenerative joint disease characterized by gradual loss of articular cartilage [20]. OA is considered the most common form of arthritis and is affecting approximately 12% of adults over the age of 60 [34,54]. Hallmark symptoms of OA are pain, tenderness, and variable degrees of local inflammation [27,34]. In Europe around 20% of chronic pain is related to osteoarthritis [35]. Different pathophysiological mechanisms seem to be involved in pain development in OA patients such as peripheral and central sensitization, where the latter may contribute to the manifestation of diffuse radiating referred pain [4,5,32,44]. In some patients, sensory abnormalities can be found in these referred pain areas, i.e. cutaneous mechanical hyperalgesia and allodynia [48], and more recently cold hyperalgesia [30].

Central manifestations of sensitization are likely to be important factors contributing to the chronification and amplification of pain in OA and are responsible for generating extra-segmental widespread sensitization [21,53]. In addition, impaired function of the descending pain modulating pathways can enhance the process of widespread sensitization [5,24]. In a recent meta-analysis sensitization was estimated to be present in about 30% of OA patients [28]. Furthermore, pre-operative sensitization is found to be a predictor for developing chronic postoperative pain after total knee replacements (TKR) [4,46].

Clinically, proxies for centralized sensitization may be detected by e.g. mechanical cutaneous hyperalgesia, cold hyperalgesia, tactile allodynia, deep somatic hyperalgesia, enhanced temporal summation, and impaired conditioned pain

modulation (CPM) [4,21,24]. Subgroups of patients with painful knee OA have already been identified in previous studies by use of different stratification tools such as sensory testing, comorbidity questionnaires and epidemiological data [11,13,15]. However, in order to implement more routine clinical screening and profiling of pain mechanisms in chronic pain patients with e.g. OA there is a need to develop a simple to use and clinical applicable, bedside tool-kit for detecting sensitization and thereby phenotyping patients. Such a diagnostic test may in the future provide better options for stratifying the pain management regime.

Thus, the aim of this pilot study was to develop a clinical applicable, easy-to-use, bedside screening tool-kit to identify sensitization in groups of patients with chronic knee OA pain or chronic postoperative pain after TKR. To achieve this, we applied supervised and unsupervised machine learning techniques in order to identify the most accurate combination of parameters indicating sensitization, i.e. quantitative sensory testing (QST) and bedside items, temporal summation, CPM, pain questionnaires and demographic data.

Methods

This pilot study was approved by the independent local ethics committees of the University Hospital of Kiel (AZ D403/18) and committee of the North Denmark Region (N-20170088).

The conduction of the study was in accordance with the declaration of Helsinki. All patients signed informed written consent prior to initiation of any protocol required procedures.

Study design

The study was conducted in two study centers (Kiel, Germany; Aalborg, Denmark). Patients were recruited through personal contact in department of orthopedics and trauma surgery of the University Hospital of Kiel, through notice sheets in practices of orthopedics and general practitioners, and through telephone calls based on information from medical charts of patients from the Aalborg University Hospital.

After an initial evaluation of eligibility, the patients were invited to the study center. At the beginning of the visit the patients were informed, the in- and exclusion criteria were checked, and demographic data (age, sex, body mass index (BMI), usage of pain medications, and general pain intensities) were collected. The outcome assessment tests were performed as follows

1) Familiarization with the bedside test equipment and the CPM bedside testing procedure

2) Questions about general pain intensities regarding the last week prior to the visit

3) Quantitative sensory testing (4 parameters)

Finally, the patients had to fill out a set of questionnaires.

Study population

To be included in the study, the following requirements had to be fulfilled: 1) Male or female between 40 and 80 years of age, 2) body weight between 40 kg and 150 kg with a BMI between 19-40 kg/m^2, 3) idiopathic osteoarthritic knee pain (index knee) diagnosed in accordance with the American College of Rheumatology modified clinical classification criteria [1] and verified radiologically as Kellgren-Lawrence grade I, II or III at the index knee or chronic knee pain after TKR [23], 4) duration of knee pain >6 months with an average daily pain score of ≥4 on a numeric rating scale (NRS) over the last week prior to visit [17]. Analgesic medications, including

over-the-counter analgesics, non-steroidal anti-inflammatory drug, gabapentin, pregabalin, opioids and antidepressants, were allowed.

The following criteria excluded patients from the study: 1) diagnosed condition suggestive of a secondary cause of knee OA (including but not limited to knee trauma, septic arthritis, inflammatory joint disease, articular fracture, major dysplasia or congenital abnormality, ochronosis, acromegaly, hemochromatosis, Wilson's disease, or primary osteochondromatosis), 2) history of surgery (including arthroscopy) in the index knee within 3 months prior to visit, 3) history of complicated prior injury to the index knee within 12 months prior to visit, 4) history of prior synovial fluid analysis showing a white blood cell count ≥2000mm^3 that is indicative of a diagnosis other than OA at the index knee, 5) use of lower extremity assistive devices other than a knee brace or 'shoe lift', use of a cane in the hand opposite to the index knee was acceptable, 6) presence of any confounding painful or neurological condition that may interfere with assessment of the index knee joint. Knee pain should be the predominant pain, but mild OA of the hands and hips were allowed, 7) skin lesions in the test area, and 8) history of any other musculoskeletal or arthritic condition that may affect the interpretation of clinical efficacy and/or safety data or otherwise contraindicates participation in this clinical study (i.e. currently symptomatic fractures or any concurrent rheumatic disease such as but not limited to fibromyalgia, rheumatoid arthritis, gout, pseudo-gout or Paget's disease and Reiter's syndrome).

Quantitative assessment of sensory function

A comprehensive QST protocol was used to assess the somatosensory signs associated with sensitization. A standardized QST protocol was established by the German Research Network on Neuropathic Pain to precisely assess the function of

the somatosensory innervation in humans [41]. For the present investigation, a sub-set of four QST parameters were selected to detect sensitization, i.e., mechanical pain threshold (MPT), mechanical pain sensitivity (MPS), dynamic mechanical allodynia (DMA) and pressure pain threshold (PPT). The Wind-up Ratio (WUR) was not included as parameter for sensitization as it does not seem to be able to distinguish between different patient groups [7]. In addition, descending pain control was assessed (see below). All tests were performed in the area over the most affected knee (10 cm proximal of the most affected knee on the vastus medialis of the quadriceps femoris), and extra-segmentally at the ipsilateral ventral forearm (superficial flexors). Data evaluation was performed as described elsewhere [41]. Briefly, z-values were calculated for MPT, MPS and PPT in order to compare patient's data with gender- and age-matched healthy controls. Since there are no QST reference values for the thigh and the forearm, the most adjacent reference areas were used for calculation of z-values instead, i.e. the dorsum of the hand and foot. Z-values above +1.96 indicate abnormal gain of function (hyperalgesia) and below -1.96 abnormal loss of function (hypoalgesia) [29]. For DMA, which is normally absent in healthy subjects, raw data were used and values above 0 were defined as abnormal.

Bedside sensory testing battery

All bedside tests (Fig. 1) were performed in the area over the most affected knee, (10 cm proximal of the most affected knee on the vastus medialis of the quadriceps femoris), and extra-segmentally at the ipsilateral ventral forearm (superficial flexors).

Mechanical pinprick pain sensitivity

A single pinprick, using a 0.7 mm CMS-nylon-filament (Chicago Medical Supply, LLC., USA) was applied perpendicularly to the skin (90° angle, until slight bending of the hair, which occurs when a force of 75 gram is applied). The patient had to rate the pain intensity of the filament on an 11-point numerical rating scale (NRS, 0=no pain, 10=worst pain imaginable).

Mechanical temporal summation

The nylon-filament (0.7 mm) was applied perpendicularly to the skin (90 angle, until slight bending of the hair, which occurs when a force of 75 gram is applied). The pain intensity of this single application was compared to a series of 10 repetitive stimuli (1/s applied within an area of 1 cm^2). The patient had to rate the pain intensity of the single stimulus and of the last stimulus of the series on a NRS, directly after each application. Beside the separate rating of both the single and series stimuli, temporal summation was calculated as the difference in pain intensity rating between the last stimulus of the series and the single stimulus. In addition, a modified ratio (series/single stimulus) was calculated with a 0.5 and a 1.0 shift of the NRS. This ratio was used to prevent too many values being lost when dividing by zero, i.e. in case the single stimulus is evaluated as 0.

Dynamic mechanical allodynia

The skin was stroked with a cotton swab for 4 times, i.e. 2 times from each direction of a cross with 90° angles and a velocity of 2-3 cm /sec. The length of each stroke was 3-5 cm. The patient had to rate the evoked pain intensity on a NRS.

Pressure pain sensitivity

A bedside algometer (10 ml syringe with a covered application button) was applied on the skin over a muscle, i.e. vastus medialis or forearm flexor group. First, the air in the syringe was compressed with constant speed (1 ml per second, starting at 10 ml) until the 4 ml mark was reached and 6ml of air had been compressed (see Video Supplement Digital Content 1, which demonstrates the application of the bedside algometer, available at http://links.lww.com/PAIN/B385). After this stimulus, the patient had to rate whether the stimulus was painful or not. If the stimulus was painful, the patient rated the pain intensity on an NRS. Next, the air in the syringe was compressed with constant speed (1 ml per second) until the pressure became painful. The patient had to indicate immediately when the pressure became painful (pressure pain threshold in ml of compressed air within the syringe, with values of 10ml indicate a low threshold, and values of 0 ml high threshold) [40].

Conditioned pain modulation

To investigate the descending pain control system, a newly developed CPM bedside test was used. This bedside tool has been shown to be reliable [25]. The test stimulus was a 6 kg, spring-based pressure algometer, which was applied for 10 sec at the m. tibialis anterior (contralateral to the most affected knee side) followed by a rating of the pain intensity of the test stimulus (continuous 10 cm visual analog scale (VAS), 0=no pain, 10=worst pain imaginable). Then, as conditioning stimulus, a 1.3 kg pressure clip, was applied to the ipsilateral earlobe for 60 sec, followed by a rating of the pain intensity of the conditioning stimulus (VAS). While the tonic earlobe pain stimulation still was ongoing at the end of the 60 sec, the 6 kg pressure algometer was applied again for 10 sec and followed by a rating of pain intensity of the test stimulus (VAS). Conditioned pain modulation effect was calculated as the difference

between test stimulus pain ratings without and with conditioning stimulus. Patients were defined as CPM responders if they perceived a decreased test stimulus pain intensity during conditioning stimulus and as non-responders if they experienced no change or an increased test stimulus pain intensity during conditioning stimulus.

Patient reported outcome measures

Pain intensity ratings

For the index knee, the average daily pain NRS intensity score over the last week prior to the visit was assessed (0=no pain, 10=worst pain imaginable). Furthermore, maximal pain intensity during rest (day and night), stair climbing, and walking was assessed.

Brief Pain Inventory (BPI)

The BPI (Severity and Interference scores) is a self-reported questionnaire that measures the severity of pain and the interference of pain with function [10,49]. The scores range from 0 (no pain) to 10 (pain as severe as you can imagine). There are 4 questions assessing worst pain, least pain, and average pain in the past 24 hours, and the pain right now. The Interference scores range from 0 (does not interfere) to 10 (completely interferes). There are 7 questions assessing the interference of pain in the past 24 hours for general activity, mood, walking ability, normal work, and relations with other people, sleep, and enjoyment of life.

Knee injury and Osteoarthritis Outcome Score (KOOS)

The KOOS is a knee-specific questionnaire which was developed to detect the subjective joint discomfort and the related impairment of the patients [43]. It evaluates short- and long-term knee-conditions and consists of 42 questions. The questions are

grouped into five subscales (pain, symptoms, activities of daily life, sport/recreation and knee related quality of life) and items are assessed on a 5-point Likert-scale (ranging from 0=no symptoms/presence to 4=extreme symptoms/presence). The KOOS scores ranges from 0 (worst) to 100 (best) and are calculated for each subscale [42].

PainDETECT-Questionnaire (PD-Q)

The PD-Q has been developed as a screening tool for assessing neuropathic symptoms in chronic pain disorders. Furthermore, sensitization characteristics in chronic musculoskeletal pain such as chronic low back pain and osteoarthritis can most likely to some degree be captured [16,19].

The questionnaire is comprised of three major components: general pain intensity (current, average, maximum pain), pain course pattern and radiating pain as well as graduation of pain. Pain graduation consists of 7 questions evaluating typical neuropathic symptoms on a 6-point Likert scale (0=never, 5=very strongly). The PD-Q sum score is calculated by addition of the subject's responses to all questions ranging from -1 to 38. Total PD-Q scores of ≤ 12 (negative) indicate that a neuropathic pain component is unlikely, scores of ≥ 19 (positive) indicate that a neuropathic pain component is likely and scores of 13 to 18 are uncertain.

Pain Quality Assessment Scale (PQAS)

The PQAS evaluates both neuropathic pain and non-neuropathic pain components by asking patients to rate 20 pain domains (e.g., intensity, shooting, numb, dull) on an 11-point NRS (0=no pain or [not sensation/item], 10=the most [descriptor] pain sensation imaginable) as the average over the last week [22]. Fifteen items can be categorized in three subgroups, i.e. paroxysmal pain (shooting, sharp, electric, hot,

radiating), surface pain (itchy, cold, numb, sensitive, tingling) and deep/dull pain (aching, heavy, dull, cramping, throbbing) [51].

Self-constructed questions

The patients had to answer three questions concerning sensory qualities, which indicate the presence of sensitization (pinprick hyperalgesia, allodynia, pressure pain) with 'yes' or 'no':

1) Have you experienced pain caused by a pointed object touching your skin (the point of a pencil, for example) on your skin in the area(s) you indicated on the pain drawing during the last week?

2) Have you experienced pain when something brushed lightly against you in the area(s) you indicated on the pain drawing during the last week?

3) Have you recently experienced pain caused by slight pressure on your skin (a finger pushing against you, for example) in the area(s) you indicated on the pain drawing during the last week?

Statistics

Sample size considerations

Sample size was assessed based on the error/precision of the coefficient [12]. Based on observations in previously performed clinical trials and clinical experience/routine, it was assumed that 40% of the study cohort would show signs of sensitization [2,28]. The above-proposed analyses were performed to identify a well performing screening tool. Thus, κ was not expected to be very low and for sample size purposes, κ=0.85 was assumed. The sample size of 100 subjects yielded an acceptable lower bound of the confidence interval of κ of 0.736.

Descriptive statistics

Continuous variables are expressed as mean values ± standard deviation (SD). Categorical data are presented by absolute frequencies and/or percentages. Agreement between two variables was assessed by the kappa () coefficient.

Group allocation

a) QSTGroup method

The patients were sub-grouped into patients with and without sensitization based on two different approaches. First, the four QST parameters were used to assign a sensitization status. A patient was identified as 'sensitized' if at least one of the four investigated QST parameters over the most affected knee indicated hyperalgesia (MPT, MPS, PPT) or DMA.

b) StatGroup method

The second sub-grouping approach was based on a combination of different unsupervised machine learning techniques. Four clustering algorithms were applied to divide the study cohort into several subsets. Hierarchical clustering was performed twice, once for each Ward-fusion-algorithm found in the literature [33] and each of them assessing dissimilarities by an Euclidean distances metric. Additionally, the partitioning clustering algorithm k-means was applied on the data set as well as a principal component analysis (PCA). The latter reduced dimensionality to one dimension allowing the natural ordering on real numbers to create two subsets. A scree-plot for the PCA identified the reduction to one dimension as meaningful. The optimal number of clusters was found by scree-plots for the hierarchical clustering algorithms and, in case of the k-means algorithm, by the average silhouette-method.

The optimal number of clusters was two for all techniques (see Supplement Digital Content 2, available at http://links.lww.com/PAIN/B386, which shows the scree plots for the PCA and both hierarchical clustering analyses and the silhouette plot for the k-means clustering techniques). A patient received the status 'sensitized' if at least three of the four statistical methods clustered the patient accordingly. All four clustering methods have a common advantage compared to the above described QST-method, using the complete information of all assessments as opposed to four single items. As all assessments are at least to some extent clinically sensible to address sensitization, the two clusters which are found by a combination of these methods can indeed be connected to sensitization. Thorough point to point and overall comparisons to QST findings strengthen this conjecture (see results).

Supervised machine learning to identify best performing bedside test

Supervised machine learning algorithms were trained on and applied to the data set. For robustness, again several procedures were applied: Recursive partitioning, random forests, k-nearest neighbors, Naïve Bayes, logistic regression and linear discriminant analyses. First, these algorithms were trained on a defined proportion of the cohort. Next, the algorithm was exerted on the rest of the study cohort, investigating its performance on "unseen", independent data. Due to the low sample size for application of machine learning techniques, the 'Leave-one-out'-method was used: For each patient, the algorithm was trained on the remaining 99 patients and the resulting algorithm utilized to classify the patient, who was left out during the training. This procedure was repeated for all 100 patients. The relative frequency of correct classifications is referred to as the *accuracy*.

Decision tree

To identify potential variables for a new screening tool, recursive partitioning was applied to create several decision trees. Due to the sample size, the full data set was used to train the algorithm. To prevent over-parametrization and enable a better application of the parameter-combination to other patient cohorts, the decision trees underwent a procedure called ‘pruning’. The optimal pruning parameter was identified by a thorough cross validation technique. To identify the most appropriate tool for clinical use, the four time-consuming QST parameters were excluded prior to building the decision trees. The decision trees show different possible combinations of the investigated parameters, which can be used in a defined sequence to characterize a patient as ‘sensitized’ or ‘not sensitized’.

The above supervised machine learning techniques were conducted for both label-variables (QST and the result of the combined clustering technique). Different variables were fed into the supervised machine learning including demographic data like age, weight and height, QST and bedside parameters as well as patient reported outcome measure as the PD-Q, the KOOS and the BPI (see Supplement Digital Content 3, which presents a complete list of variables included in the supervised machine learning algorithms, available at http://links.lww.com/PAIN/B386). Clearly prior to all supervised machine learning techniques on the QST-label, the QST parameters were locally removed to avoid circular reasoning.

Descriptive statistics were performed with IBM SPSS statistics for Windows (version 23.0, NY). The main statistical analyses were conducted using the statistical software R (R Core Team, 2019) [38].

Results

In this study, 100 patients with chronic painful knee OA (n=86) or chronic pain after TKR (n=14) were included. Table 1 and 2 show the descriptive analyses of the questionnaires and bedside test results. As seen in Table 2, the pain intensity of the bedside CMS hair was rated low overall, which may be explained by the fact, that this tool is less sharp compared to the pinpricks used within the DFNS QST protocol.

The sub-set of QST analyses, i.e. frequencies of sensory abnormalities, which were used to allocate the patients into the two groups (sensitized/not sensitized) is displayed in Table 3.

Although no CPM effect was observed when comparing the mean values of all patients, individual analysis showed a CPM effect in 34% of patients (responders), whereas 66% of patients were non-responders, i.e. 36% exhibited an insufficient CPM effect and 30% showed no change in pain intensity (non-responders) (Fig. 2).

Subgrouping: Identifying subgroups of patients with and without sensitization (QST and unsupervised machine learning)

Subgrouping of the patients by the two different approaches, i.e. QST and combined machine learning techniques, identified 46% of the patients as being ‘sensitized’. However, 18 patients received different allocations by the two approaches (Table 4), yielding an agreement of These discordant pairs were revised on a patient level from a medical point of view. Overall, the statistical approach seems to be more comprehensive compared with the standardized QST, as the unsupervised machine learning algorithm includes additional variables, that may have an impact on pain sensitization.

The variable indicating the QST classification was termed *QSTGroup*. The variable indicating the cluster based on the machine learning techniques was named *StatGroup.* As there is no definite answer which of the two allocations actually corresponds to reality, the subsequent analyses were performed for both of these two label-variables.

Classification: Identifying a new screening tool for assessing sensitization (supervised machine learning)

As a first step to identify a new screening tool for assessing sensitization, supervised machine learning techniques were performed for both label-variables, *StatGroup* and *QSTGroup*. The accuracy of the different supervised machine learning techniques for both label-variables are illustrated in Table 5.

As a second step, different decision trees containing varying combinations of the investigated parameters, were generated depending on the label-variable, *StatGroup* or *QSTGroup*. The two most promising decision trees are shown in Figure 2 and Figure 3. The most promising decision tree for the label-variable *StatGroup* consists of three bedside tests, performed as an if/then approach (Fig. 3). The first test assesses pressure pain sensitivity, i.e. pain intensity at 4 ml pressure with the 10 ml bedside syringe in the area over the most affected knee (vastus medialis). A pain intensity below 1.5 on NRS at the 4 ml mark indicates 'not sensitized'. In total, 53 patients rated below 1.5, of whom 85% were correctly defined as 'not sensitized'. These 53 patients are further assessed with the second test of mechanical pinprick pain sensitivity (CMS-nylon) over the most affected knee (vastus medialis). An evoked pain intensity of <2.5 on the NRS indicates 'not sensitized' (n=47). 96% of these patients were correctly defined as 'not sensitized'. Patients with a NRS ≥1.5 within the first step (pressure pain sensitivity, pain intensity at 4 ml pressure), were

further assessed with the third test measuring pressure pain sensitivity, i.e. pain threshold (ml compressed) extra-segmentally (ipsilateral forearm). A pressure pain threshold of <4 ml indicates 'not sensitized' (n=6). The most promising decision tree for the label-variable *QSTGroup* consists of two bedside tests (mechanical temporal summation single stimulus extra-segmentally and pressure pain sensitivity, i.e. painful when compressed to 4 ml at the index knee) and one questionnaire subscale (KOOS knee related quality of life) (Fig. 4).

With regard to the initial sample size calculation (κ=0.85), the most promising decision tree for the label-variable *StatGroup* (Fig. 2) revealed an almost perfect agreement (0.90), while the most promising tree for *QSTGroup* revealed a substantial agreement higher than the lower bound of the sample size determination (0.76).

Discussion

According to both sensitization classification approaches (QST, combined machine learning techniques) 46% of the patients showed signs of sensitization. The classifications of the two approaches coincide for 82 of the 100 patients, showing an agreement of .

Depending on the label-variable, two promising bedside tool-kits were identified that reach an average accuracy of 91.8% for the outcome variable *StatGroup* and 75.8% for the outcome variable *QSTGroup*. Different decision trees were created to discriminate patients with and without sensitization. Two decision trees, one each for the label-variable *StatGroup* and *QSTGroup*, were derived based on optimal statistical properties. These decision trees identify variables for a new, easy-to-use bedside screening tool, revealing a substantial (*QSTGroup*) to almost perfect (*StatGroup*) agreement.

Detection and degree of sensitization

Approximately half of the patients with chronic painful knee OA or chronic pain after TKR demonstrated signs of sensitization. Our finding is higher as compared with previous estimations from a systematic literature review or results of another screening tool where around 30% and 27-38% were found sensitized [2,28]. This is most likely because we included chronic patients (>10 years pain duration) who suffered from a high pain intensity (NRS >5) [5]. A higher degree of pain in OA is related to higher degrees of sensitization [4,14].

It has been reported that 70% of patients with knee OA had at least one sensory QST abnormality [52]. In particular, there was a lower knee pressure pain threshold, which occurred in 32% of the patients and in 20% extra-segmentally on the forearm.

Focusing on parameters that have been described as reflecting sensitization [4,6], we did not examine the full battery of QST parameters. However, cold pain hyperalgesia has recently been shown as an interesting possible proxy for sensitization [31] and hence, the bedside tool kit can be optimized. Still, compared to previous data [52], in the present study almost the same number of patients (34%) showed a reduced pressure pain threshold at the knee, while the pressure pain threshold at the forearm was reduced more frequently (52%). Both studies, the one described above and ours, emphasize the importance of deep somatic hyperalgesia in OA of the knee and the apparent spread of hyperalgesia into extra-segmental areas [26]. The presence of preoperative sensitization is of clinical importance because it is associated with a worse outcome after joint replacement [53]. There have been few studies regarding the mechanical detection threshold, but it has been shown that patients with knee OA also had higher pain intensities to von Frey filament indicating mechanical hyperalgesia [39]. Likewise, we found a lower mechanical pain threshold in 20% of the patients segmentally and 46% extra-segmentally, as well as an increased mechanical pain sensitivity in 11% segmentally and 18% extra-segmentally. A previous study showed that these changes occur even in mild forms of OA and that they are also detectable extra-segmentally on the contralateral leg [39].

Although no averaged CPM effect was observed when comparing the mean values of all patients, the individual CPM values contained CPM-responders and non-responders. Hence, the presentation of the distribution of individual CPM responses provide information of the underlying variation and subgroups, respectively [3]. The mixed distribution of CPM responders and non-responders as previously highlighted [3,45,47] should be considered in the future. In addition, the CPM effect is generally

influenced by the used testing regime applied [36,50]. This variation is not fully understood and future work should address this knowledge gap [37].

Bedside testing

The tool-kit phenotyping of OA and TKR patients with signs of sensitization may help capturing additional facets of the patient's suffering. This is of importance as pain intensity alone cannot mirror the multidimensional pain mechanisms. OA patients experience difficulties to express their impairment because of a lack of fitting descriptors [9]. Also, psychological constraints, like self-imposed stoicism or perseveration of the social-/self-image, complicate the assessment of OA pain [9]. An easy-to-use, clinical applicable, bedside tool-kit may expand the repertoire of assessment options to improve the care and possible stratify management of OA patients. In particular counselling OA patients with signs of sensitization prior to surgery may be an important asset.

Identification of a screening tool based on unsupervised and supervised machine learning

To select the most promising algorithm for the identification of sensitization in OA and TKR chronic pain patients, the method of unsupervised machine learning was used. By this procedure, all assessed parameters could be included into the analysis, accounting for the heterogeneity of sensitization. This does not only include bedside testing or QST, but also all questionnaires and epidemiological data.

Subsequently supervised machine learning techniques were applied to the QST- and the newly determined unsupervised machine learning-labels. The two most powerful screening tools mainly rely on bedside testing except one item of the KOOS (knee related quality of life). A connection between pain intensity and quality of life has

already been described in previous studies [39]. Signs of sensitization assessed via QST were more likely to be found in knee OA patients with higher scores in the modified PD-Q [19]. Thus, the PD-Q could be a useful tool for the identification of sensitization in OA. PD-Q symptoms and QST seem to address different aspects of the pain and sensitization manifestations [18]. In line with this, we found that signs of sensitization were present in 46% of the investigated patients, whereas only 13% showed a likely neuropathic pain component according to the PD-Q score. The majority of patients, however, were characterized by a PD-Q score below 12 (unlikely). Overall, the PD-Q may not be as relevant as other parameters for identification of sensitization in OA.

The most promising algorithm reached an agreement of 90% and consists of three bedside tests.

The first test assesses pressure pain sensitivity with the bedside syringe over the most affected knee. A pain intensity <1.5 (NRS 0-10) during a pressure reaching the 4 ml mark of the syringe indicates ‘not sensitized’. The second test assesses mechanical pinprick pain sensitivity using the stiff nylon-filament (0.7 mm) over the most affected knee. An evoked pain intensity of <2.5 (NRS 0-10) indicates ‘not sensitized’. The third test measures pressure pain sensitivity extra-segmentally. A pressure pain threshold below 4 ml indicates ‘not sensitized’. Thus, three simple bedside tests, two at the affected joint and one at the forearm, allow classifying sensitized OA or TKR patients with high accuracy. It takes about 1 min at most to perform the test battery and hence may form the basis for further tests and applications. Whether these parameters should be used as an if/then approach as shown in the decision tree, or whether it might be easier to perform them sequentially, should be discussed in future studies.

Limitations

Machine learning techniques are generally applied to large data sets. A sample size of 100 automatically restrains performance and performance measures. Especially the split into training and test data set applying supervised machine learning techniques further reduces the respective sample sizes. Results of this pilot study have to be interpreted with caution and should be validated in studies including larger sample sizes. Nevertheless, our sample size calculation using kappa-coefficient yielded an acceptable lower bound of the confidence interval.

Since there are no QST reference values for the thigh and the forearm, the most adjacent reference areas were used for calculation of z-values. This might have had an influence on our results and could potentially explain the higher frequency of mechanical hyperalgesia extra-segmentally compared to the index knee. An important limitation of this study concerns the generalizability of the results. Due to strict in- and exclusion criteria, our results are limited to a specific patient subgroup. Our study population is patients with chronic (>10 years) and moderate-to-high pain intensity (NRS >5). Since sensitization has shown various correlations with pain duration and pain intensity [4], it remains unknown if the bedside tests would exhibit similar findings in pain cohorts with less pain durations and lower pain intensities. In addition, an impact of patients' medication on the presented results could not be fully excluded. Sensitization is complex and can present with different sensory symptoms and signs. For example, sensitization mechanisms reflected by CPM and PPT seem to be different and provide complementary information [8]. Since the bedside tests are based on an a priori definition of sensitization (*QSTGroup*) or a defined selection of variables (*StatGroup*) our results should be considered an attempt to assess patients with specific markers of sensitization.

Conclusion

Using different machine learning techniques enabled us to identify the most accurate parameters out of a variety of questionnaires and clinical items for identification of sensitization in a sample of OA and TKR chronic pain patients. The resulting bedside kit contains three easy-to-use items (pressure pain sensitivity and mechanical pinprick pain sensitivity over the most affected knee, pressure pain sensitivity extra-segmentally). Validation of this tool in a larger cohort is necessary to use it in clinical practice for mechanistically phenotyping OA and TKR pain patients.

Potentially, a validated version of these bedside tools could then assist in the concept and development of individualized, mechanism-based pain therapy.

Acknowledgements

This research was financially supported by Grünenthal GmbH. The final manuscript was offered to Grünenthal GmbH for review and comment. There were no involvements of Grünenthal GmbH during study design, data collection, and composition of the manuscript or decisions regarding the publication process. Center for Neuroplasticity and Pain (CNAP) is supported by the Danish National Research Foundation (DNRF121). Research funding was acknowledged from the Shionogi Science Center, Daiichi Sankyo TaNeDS, and Pfizer The Global Medical Grant.

Conflicts of interest

JS reports consultancy fees from Pfizer Pharma GmbH, speaking fees from Grünenthal GmbH and travel support from Pfizer and Alnylam Pharmaceuticals.

JCO reports research support and speaking fees from Grünenthal GmbH and travel costs from Grünenthal GmbH, Pfizer.

VK reports no conflicts of interest.

JBL reports no conflicts of interest.

LNK reports speaker and consultancy fees from Grünenthal GmbH and AleaNCore.

PH reports research support from the German Federal Ministry of Education and Research (BMBF): Verbundprojekt: Frühdetektion von Schmerzchronifizierung (NoChro) (13GW0338C) and Zambon GmbH.

LAN Received speaker and consultancy fees from: Allergan, Grünenthal, Ono, Abbott, Boehringer-Ingelheim, Pfizer, Bristol-Myers Squibb, Daiichi Sankyo, Shionogi, Ironwood Pharma, Eli Lilly, Mundipharma, Purdue, Pierre Fabre, Sanofi-Aventis, Vertex Pharmaceuticals, UCB. Received unrestricted research grants from: Shionogi, C4Pain, Daiichi Sankyo, Grünenthal, Merck, TeNeDS, and The Global Medical Grant.

RB reports grants and research support from EU Projects: „Europain“ (115007). DOLORisk (633491). IMI Paincare (777500). German Federal Ministry of Education and Research (BMBF): Verbundprojekt: Frühdetektion von Schmerzchronifizierung (NoChro) (13GW0338C). German Research Network on Neuropathic Pain (01EM0903). Pfizer Pharma GmbH, Genzyme GmbH, Grünenthal GmbH, Mundipharma Research GmbH und Co. KG., Novartis Pharma GmbH, Alnylam Pharmaceuticals Inc., Zambon GmbH, Sanofi-Aventis Deutschland GmbH. RB received speaking fees from Pfizer Pharma GmbH, Genzyme GmbH, Grünenthal GmbH, Mundipharma, Sanofi Pasteur, Medtronic Inc. Neuromodulation, Eisai Co.Ltd., Lilly GmbH, Boehringer Ingelheim Pharma GmbH & Co. KG, Astellas Pharma GmbH, Desitin Arzneimittel GmbH, Teva GmbH, Bayer-Schering, MSD

GmbH, Seqirus Australia Pty. Ltd, Novartis Pharma GmbH, TAD Pharma GmbH, Grünenthal SA Portugal, Sanofi-Aventis Deutschland GmbH, Agentur Brigitte Süss, Grünenthal Pharma AG Schweiz, Grünenthal B.V. Niederlande, Evapharma, Takeda Pharmaceuticals Internation AG Schweiz, Ology Medical Education Netherlands. RB reports consultancy fees from Pfizer Pharma GmbH, Genzyme GmbH, Grünenthal GmbH, Mundipharma Research GmbH und Co. KG, Allergan, Sanofi Pasteur, Medtronic, Eisai, Lilly GmbH, Boehringer Ingelheim Pharma GmbH&Co.KG, Astellas Pharma GmbH, Novartis Pharma GmbH, Bristol-Myers Squibb, Biogenidec, AstraZeneca GmbH, Merck, Abbvie, Daiichi Sankyo, Glenmark Pharmaceuticals S.A., Seqirus Australia Pty. Ltd, Teva Pharmaceuticals Europe Niederlande, Teva GmbH, Genentech, Mundipharma International Ltd. UK, Astellas Pharma Ltd. UK, Galapagos NV, Kyowa Kirin GmbH, Vertex Pharmaceuticals Inc., Biotest AG, Celgene GmbH, Desitin Arzneimittel GmbH, Regeneron Pharmaceuticals Inc. USA, Theranexus DSV CEA Frankreich, Abbott Products Operations AG Schweiz, Bayer AG, Grünenthal Pharma AG Schweiz, Mundipharma Research Ltd. UK, Akcea Therapeutics Germany GmbH, Asahi Kasei Pharma Corporation, AbbVie Deutschland GmbH & Co. KG, Air Liquide Sante International Frankreich, Alnylam Germany GmbH, Lateral Pharma Pty Ltd, Hexal AG, Angelini, Janssen, SIMR Biotech Pty Ltd Australien, Confo Therapeutics N. V. Belgium.

References

[1] Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, Christy W, Cooke TD, Greenwald R, Hochberg M. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986;29:1039–1049.

[2] Arendt-Nielsen L, Egsgaard LL, Petersen KK, Eskehave TN, Graven-Nielsen T, Hoeck HC, Simonsen O. A mechanism-based pain sensitivity index to characterize knee osteoarthritis patients with different disease stages and pain levels. Eur J Pain 2015;19:1406–1417.

[3] Arendt-Nielsen L, Larsen JB, Rasmussen S, Krogh M, Borg L, Madeleine P. A novel clinical applicable bed-side tool for assessing conditioning pain modulation: proof-of-concept. Scand J Pain 2020.

[4] Arendt-Nielsen L, Nie H, Laursen MB, Laursen BS, Madeleine P, Simonsen OH, Graven-Nielsen T. Sensitization in patients with painful knee osteoarthritis. Pain 2010;149:573–581.

[5] Arendt-Nielsen L, Skou ST, Nielsen TA, Petersen KK. Altered Central Sensitization and Pain Modulation in the CNS in Chronic Joint Pain. Curr Osteoporos Rep 2015;13:225–234.

[6] Baron R, Hans G, Dickenson AH. Peripheral input and its importance for central sensitization. Ann Neurol 2013;74:630–636.

[7] Baron R, Maier C, Attal N, Binder A, Bouhassira D, Cruccu G, Finnerup NB, Haanpää M, Hansson P, Hüllemann P, Jensen TS, Freynhagen R, Kennedy JD,

Magerl W, Mainka T, Reimer M, Rice ASC, Segerdahl M, Serra J, Sindrup S, Sommer C, Tölle T, Vollert J, Treede R-D. Peripheral neuropathic pain: a mechanism-related organizing principle based on sensory profiles. Pain 2017;158:261–272.

[8] Carlesso LC, Frey Law L, Wang N, Nevitt M, Lewis CE, Neogi T, Multicenter Osteoarthritis Study Group. The Association of Pain Sensitization and Conditioned Pain Modulation to Pain Patterns in Knee Osteoarthritis. Arthritis Care Res (Hoboken) 2020.

[9] Cedraschi C, Delézay S, Marty M, Berenbaum F, Bouhassira D, Henrotin Y, Laroche F, Perrot S. “Let’s talk about OA pain”: a qualitative analysis of the perceptions of people suffering from OA. Towards the development of a specific pain OA-Related questionnaire, the Osteoarthritis Symptom Inventory Scale (OASIS). PLoS ONE 2013;8:e79988.

[10] Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singap 1994;23:129–138.

[11] Cruz-Almeida Y, King CD, Goodin BR, Sibille KT, Glover TL, Riley JL, Sotolongo A, Herbert MS, Schmidt J, Fessler BJ, Redden DT, Staud R, Bradley LA, Fillingim RB. Psychological profiles and pain characteristics of older adults with knee osteoarthritis. Arthritis Care Res (Hoboken) 2013;65:1786–1794.

[12] Donner A, Rotondi MA. Sample size requirements for interval estimation of the kappa statistic for interobserver agreement studies with a binary outcome and multiple raters. Int J Biostat 2010;6:Article 31.

[13] Egsgaard LL, Eskehave TN, Bay-Jensen AC, Hoeck HC, Arendt-Nielsen L. Identifying specific profiles in patients with different degrees of painful knee

osteoarthritis based on serological biochemical and mechanistic pain biomarkers: a diagnostic approach based on cluster analysis. Pain 2015;156:96–107.

[14] Finan PH, Buenaver LF, Bounds SC, Hussain S, Park RJ, Haque UJ, Campbell CM, Haythornthwaite JA, Edwards RR, Smith MT. Discordance between pain and radiographic severity in knee osteoarthritis: findings from quantitative sensory testing of central sensitization. Arthritis Rheum 2013;65:363–372.

[15] Frey-Law LA, Bohr NL, Sluka KA, Herr K, Clark CR, Noiseux NO, Callaghan JJ, Zimmerman MB, Rakel BA. Pain sensitivity profiles in patients with advanced knee osteoarthritis. Pain 2016;157:1988–1999.

[16] Freynhagen R, Baron R, Gockel U, Tölle TR. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Curr Med Res Opin 2006;22:1911–1920.

[17] Gerbershagen HJ, Rothaug J, Kalkman CJ, Meissner W. Determination of moderate-to-severe postoperative pain on the numeric rating scale: a cut-off point analysis applying four different methods. Br J Anaesth 2011;107:619–626.

[18] Gierthmühlen J, Binder A, Förster M, Baron R. Do We Measure What Patients Feel?: An Analysis of Correspondence Between Somatosensory Modalities Upon Quantitative Sensory Testing and Self-reported Pain Experience. Clin J Pain 2018;34:610–617.

[19] Hochman JR, Davis AM, Elkayam J, Gagliese L, Hawker GA. Neuropathic pain symptoms on the modified painDETECT correlate with signs of central sensitization in knee osteoarthritis. Osteoarthritis Cartilage 2013;21:1236–1242.

[20] Hunter DJ, Felson DT. Osteoarthritis. BMJ 2006;332:639–642.

[21] Imamura M, Imamura ST, Kaziyama HHS, Targino RA, Hsing WT, de Souza LPM, Cutait MM, Fregni F, Camanho GL. Impact of nervous system hyperalgesia on pain, disability, and quality of life in patients with knee osteoarthritis: a controlled analysis. Arthritis Rheum 2008;59:1424–1431.

[22] Jensen MP, Gammaitoni AR, Olaleye DO, Oleka N, Nalamachu SR, Galer BS. The pain quality assessment scale: assessment of pain quality in carpal tunnel syndrome. J Pain 2006;7:823–832.

[23] Kohn MD, Sassoon AA, Fernando ND. Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. Clin Orthop Relat Res 2016;474:1886–1893.

[24] Kosek E, Ordeberg G. Lack of pressure pain modulation by heterotopic noxious conditioning stimulation in patients with painful osteoarthritis before, but not following, surgical pain relief. Pain 2000;88:69–78.

[25] Larsen JB, Madeleine P, Arendt-Nielsen L. Development of a new bed-side-test assessing conditioned pain modulation: a test-retest reliability study. Scand J Pain 2019;19:565–574.

[26] Lee YC, Lu B, Bathon JM, Haythornthwaite JA, Smith MT, Page GG, Edwards RR. Pain sensitivity and pain reactivity in osteoarthritis. Arthritis Care Res (Hoboken) 2011;63:320–327.

[27] Lespasio MJ, Piuzzi NS, Husni ME, Muschler GF, Guarino A, Mont MA. Knee Osteoarthritis: A Primer. Perm J 2017;21:16–183.

[28] Lluch E, Torres R, Nijs J, Van Oosterwijck J. Evidence for central sensitization in patients with osteoarthritis pain: a systematic literature review. Eur J Pain

2014;18:1367–1375.

[29] Maier C, Baron R, Tölle TR, Binder A, Birbaumer N, Birklein F, Gierthmühlen J, Flor H, Geber C, Huge V, Krumova EK, Landwehrmeyer GB, Magerl W, Maihöfner C, Richter H, Rolke R, Scherens A, Schwarz A, Sommer C, Tronnier V, Uçeyler N, Valet M, Wasner G, Treede R-D. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): somatosensory abnormalities in 1236 patients with different neuropathic pain syndromes. Pain 2010;150:439–450.

[30] Moss P, Knight E, Wright A. Subjects with Knee Osteoarthritis Exhibit Widespread Hyperalgesia to Pressure and Cold. PLoS ONE 2016;11:e0147526.

[31] Müller M, Bütikofer L, Andersen OK, Heini P, Arendt-Nielsen L, Jüni P, Curatolo M. Cold pain hypersensitivity predicts trajectories of pain and disability after low back surgery: a prospective cohort study. Pain 2020;in press.

[32] Murray GM. Guest Editorial: referred pain. J Appl Oral Sci 2009;17.

[33] Murtagh F, Legendre P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? J Classif 2014;31:274–295.

[34] Neogi T. The epidemiology and impact of pain in osteoarthritis. Osteoarthr Cartil 2013;21:1145–1153.

[35] O'Brien T, Breivik H. The impact of chronic pain-European patients' perspective over 12 months. Scand J Pain 2012;3:23–29.

[36] 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;2:162–169.

[37] Potvin S, Marchand S. Pain facilitation and pain inhibition during conditioned

pain modulation in fibromyalgia and in healthy controls. Pain 2016;157:1704–1710.

[38] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, 2018 p. Available: https://www.R-project.org/.

[39] Rakel B, Vance C, Zimmerman MB, Petsas-Blodgett N, Amendola A, Sluka KA. Mechanical hyperalgesia and reduced quality of life occur in people with mild knee osteoarthritis pain. Clin J Pain 2015;31:315–322.

[40] Reimer M, Forstenpointner J, Hartmann A, Otto JC, Vollert J, Gierthmühlen J, Klein T, Hüllemann P, Baron R. Sensory bedside testing: a simple stratification approach for sensory phenotyping. Pain Reports 2020;5:e820.

[41] Rolke R, Baron R, Maier C, Tolle TR, Treede R-D, Beyer A, Binder A, Birbaumer N, Birklein F, Botefur IC, Braune S, Flor H, Huge V, Klug R, Landwehrmeyer GB, Magerl W, Maihofner C, Rolko C, Schaub C, Scherens A, Sprenger T, Valet M, Wasserka B. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): standardized protocol and reference values. Pain 2006;123:231–243.

[42] Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28:88–96.

[43] Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003;1:17.

[44] Schaible H-G. Mechanisms of chronic pain in osteoarthritis. Curr Rheumatol

Rep 2012;14:549–556.

[45] Schliessbach J, Siegenthaler A, Streitberger K, Eichenberger U, Nüesch E, Jüni P, Arendt-Nielsen L, Curatolo M. The prevalence of widespread central hypersensitivity in chronic pain patients. Eur J Pain 2013;17:1502–1510.

[46] Skou ST, Graven-Nielsen T, Rasmussen S, Simonsen OH, Laursen MB, Arendt-Nielsen L. Facilitation of pain sensitization in knee osteoarthritis and persistent post-operative pain: a cross-sectional study. Eur J Pain 2014;18:1024–1031.

[47] Skovbjerg S, Jørgensen T, Arendt-Nielsen L, Ebstrup JF, Carstensen T, Graven-Nielsen T. Conditioned Pain Modulation and Pressure Pain Sensitivity in the Adult Danish General Population: The DanFunD Study. J Pain 2017;18:274–284.

[48] Suokas AK, Walsh DA, McWilliams DF, Condon L, Moreton B, Wylde V, Arendt-Nielsen L, Zhang W. Quantitative sensory testing in painful osteoarthritis: a systematic review and meta-analysis. Osteoarthr Cartil 2012;20:1075–1085.

[49] Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the Brief Pain Inventory for chronic nonmalignant pain. J Pain 2004;5:133–137.

[50] Vaegter HB, Petersen KK, Mørch CD, Imai Y, Arendt-Nielsen L. Assessment of CPM reliability: quantification of the within-subject reliability of 10 different protocols. Scand J Pain 2018;18:729–737.

[51] Victor TW, Jensen MP, Gammaitoni AR, Gould EM, White RE, Galer BS. The dimensions of pain quality: factor analysis of the Pain Quality Assessment Scale. Clin J Pain 2008;24:550–555.

[52] Wylde V, Palmer S, Learmonth ID, Dieppe P. Somatosensory abnormalities in

knee OA. Rheumatology (Oxford) 2012;51:535–543.

[53] Wylde V, Sayers A, Lenguerrand E, Gooberman-Hill R, Pyke M, Beswick AD, Dieppe P, Blom AW. Preoperative widespread pain sensitization and chronic pain after hip and knee replacement: a cohort analysis. Pain 2015;156:47–54.

[54] Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Clin Geriatr Med 2010;26:355–369.

Legends to the figures

Figure 1: Bedside sensory testing devices.

(1) 0.7 mm CMS-nylon filament, (2) cotton swab, (3) 10 ml syringe with blocked tip, (4) 1.3 kg pressure clip, (5) 6 kg pressure algometer.

Figure 2: Individual conditioned pain modulation (CPM) effect.

Individual CPM effect are ranked and plotted as function of individuals (n=99). The CPM effect was calculated as change between pain ratings without conditioning stimulus and pain ratings with conditioning stimulus. Patients with negative VAS scores were defined as CPM responders (n=34) and patients with no change or positive VAS scores were defined as CPM non-responders (n=65). VAS, visual analog scale (0=no pain, 10=worst pain imaginable)

Figure 3: Most promising decision tree for the label-variable *StatGroup*.

'Pruned' decision tree with all PCA-variables, factor loadings >0,5; green = not sensitized (NS), red = sensitized (S).

Figure 4: Most promising decision tree for the label-variable *QSTGroup*.

'Pruned' decision tree based on all variables (green = not sensitized (NS), red = sensitized (S).

Table 1: Descriptive statistics of the study cohort.

Descriptive statistics	**n = 100**
Age, years	62.9 ± 9.6
Females (%)	66 (66%)
BMI (kg/m^2)	28.4 ± 5.0
Pain duration in index knee (years)	10.9 ± 10.2
Average daily pain intensity in index knee, last 7 days (NRS)	5.3 ± 1.6
Max daily pain intensity at rest (NRS)	4.0 ± 2.6
Max nightly pain intensity at rest (NRS)	4.1 ± 3.1
Pain walking (NRS)	4.9 ± 2.1
Pain climbing stairs (NRS)	5.8 ± 2.1
Brief Pain Inventory (0-10 each)	
Pain severity score	3.7 ± 1.8
Pain interference score	3.4 ± 2.3
Knee injury and Osteoarthritis Outcome Score (0-100 each)	
Symptoms	50.0 ±17.9
Pain	51.3 ± 15.3
Activities of daily living	49.6 ± 20.5
Sport and recreation	55.8 ± 32.3
Knee-related quality of life	57.1 ± 20.3
PainDETECT Questionnaire score (0-38)	
Total score	11.0 ± 6.3
Neuropathic component:	
Unlikely (%)	66%
Uncertain (%)	21%
Likely (%)	13%
Pain Quality Assessment Scale (0-10 each)	
Paroxysmal pain	3.5 ± 2.0
Surface pain	1.6 ± 1.5
Deep pain	3.3 ± 2.0
Self-constructed questions	
1) Pain caused by sharp object in the last week (%)	16%
2) Pain caused by light touch in the last week (%)	20%
3) Pain caused by light pressure in the last week (%)	44%
Values are given as mean ± standard deviation or by absolute frequencies and/or percentages NRS, numeric rating scale (0=no pain, 10=worst pain imaginable)	

Table 2: Descriptive data of bed-side tests.

Bed-side tests	**Index knee (most affected knee)**	**Extra-segmental (ipsilateral ventral forearm)**
Mechanical pinprick pain sensitivity (NRS)	1.3 ± 4.6	1.2 ± 1.4
Mechanical temporal summation (NRS)		
single stimulus	1.04 ± 1.2	1.3 ± 1.4
series stimuli	2.6 ± 1.9	2.6 ± 2.0
difference (series-single stimulus)	1.6 ± 1.4	1.3 ± 1.1
ratio (series/single stimulus)[#]	2.2 ± 1.2	1.9 ± 0.8
ratio (series/single stimulus) with 0.5 shift	2.7 ± 1.9	2.1 ± 1.3
ratio (series/single stimulus) with 1.0 shift	2.0 ± 1.0	1.7 ± 0.7
Dynamic mechanical allodynia (NRS)	0.02 ± 0.2	0
Pressure pain sensitivity		
Painful when compressed to 4 ml? (Yes %)	59%	67%
Pain intensity at 4 ml pressure (NRS)	2.0 ± 2.2	2.3 ± 2.4
Pain threshold (ml compressed)	4.9 ± 1.9	5.0 ± 1.8
Conditioned pain modulation (VAS)		
Test stimulus without conditioning stimulus pain rating*	5.6 ± 2.7	
Test stimulus with conditioning stimulus pain rating*	5.6 ± 3.0	
Conditioning stimulus pain rating*	6.2 ± 2.6	

Values are given as mean ± standard deviation or by absolute frequencies and/or percentages

[#] 44 missing values for the index knee and 37 extra-segmental (single stimulus was rated as '0')

*one missing value due to unbearable pain during the test, which led to termination of test (n=99)

NRS, numeric rating scale (0=no pain, 10=worst pain imaginable); VAS, visual analog scale (0=no pain, 10=worst pain imaginable)

Table 3: Quantitative sensory testing results for the most affected knee (index knee) and the ipsilateral ventral forearm (extra-segmental).

QST parameter	Hypoalgesia, %	Normal, %	Hyperalgesia, %
Mechanical pain threshold			
Index knee	9	71	20
Extra-segmental	6	48	46
Mechanical pain sensation			
Index knee	6	83	11
Extra-segmental	6	76	18
Pressure pain threshold			
Index knee	1	65	34
Extra-segmental	2	46	52
Dynamic mechanical allodynia			
Index knee		96	4
Extra-segmental		97	3

Table 4: Distribution of sensitized patients classified with *QSTGroup* and *StatGroup*.

StatGroup / ***QSTGroup***	Not sensitized (%)	Sensitized (%)
Not sensitized (%)	45	9
Sensitized (%)	9	37

Table 5: Results for label-variable *StatGroup* and *QSTGroup* using the 'Leave-one-out' method.

Method	***StatGroup* Accuracy (%)**	***QSTGroup* Accuracy (%)**
Recursive Partitioning (cp=0.05)	94.0	80.0
Random Forest	96.0	80.0
k-nearest neighbors	94.0	73.0
Naïve Bayes	81.0	73.0
Logistic regression		75.0
Linear discriminant analysis	94.0	74.0
Average	91.8	75.8
Comparison QST complete data set	82.0	
Logistic regression algorithm did not converge for *StatGroup* due to a perfect separation