

RESEARCH

Open Access

Cross-cultural adaptation and validation of the German Central Sensitization Inventory (CSI-GE)



Michel Klute^{1*}, Marjan Laekeman², Katrin Kuss³, Frank Petzke¹, Angela Dieterich⁴, Andreas Leha⁵, Randy Neblett⁶, Steffen Ehrhardt⁷, Joachim Ulma⁸ and Axel Schäfer⁹

Abstract

Background: The Central Sensitization Inventory (CSI) is a screening tool designed to detect symptoms related to Central Sensitization (CS) and Central Sensitivity Syndromes (CSS) by measuring the degree of related phenomena. The objective of this study was to create a German, culturally-adapted version of the CSI and to test its psychometric properties.

Methods: A German version of the CSI (CSI-GE) was developed, culturally-adapted, and pretested for comprehensibility. The psychometric properties of the resulting version were validated in a clinical study with chronic pain and pain-free control subjects. To assess retest reliability, the CSI-GE was administered twice to a subgroup of patients. Structural validity was tested using factor analyses. To investigate construct validity a hypotheses testing approach was used, including (1) correlations between the CSI-GE and several other well-established questionnaires as well as (2) an investigation of the CSI-GE discriminative power between different subgroups of participants believed to have different degrees of CS.

Results: The CSI-GE showed excellent reliability, including high test-retest characteristics. Factor analyses confirmed a bi-factor dimensionality as has been determined previously. Analysing construct validity 6 out of 11 hypotheses (55%) were met. CSI-GE scores differentiated between subgroups according to expectations. Correlations between CSI-GE scores and other questionnaires suggested that none of the correlated constructs was identical, but there was overlap with other questionnaires based on symptom load. Several correlations did not fit with our current understanding of CS.

Conclusion: The CSI-GE appears to be a reliable tool for measuring CS/CSS-related symptomatology. Whether this implies that the CSI-GE measures the degree of CS within an individual subject remains unknown. The resulting score should be interpreted cautiously until further clarification of the construct.

Keywords: Central sensitization inventory, Central sensitization, Central sensitivity syndromes, CSI, CSI-GE, Chronic pain, Cross-cultural adaptation, Psychometric validation

* Correspondence: michel.klute@stud.uni-goettingen.de

¹Pain Clinic, Department of Anaesthesiology, University Medical Center, Georg August University of Göttingen, Robert-Koch-Str. 40, 37075 Göttingen, Germany

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Chronic Pain is often related to a multitude of underlying factors, which can trigger, contribute to, and maintain it [1, 2]. Recently, a new classification for chronic pain was added to the ICD-11 (International Classification of Diseases). It introduced a whole chapter on chronic pain conditions that are now understood as primary health problems in themselves [3, 4]. Central Sensitization (CS) appears to be an important feature for the development and maintenance of many of these chronic pain conditions, irrespective of other etiological aspects [5–7]. The International Association for the Study of Pain [8] defines CS as “*increased responsiveness of nociceptive neurons in the central nervous system to their normal or subthreshold afferent input*”. No gold standard for the diagnosis of CS exists, so it is difficult to assess its presence and magnitude [5, 9, 10]. There have been many different attempts to objectively quantify CS [11], including Quantitative Sensory Testing (QST) [12] and imaging techniques [11]. However, these tools are complex, time-consuming and expensive [12, 13].

Yunus [14] postulated that CS is a common feature in a number of insufficiently understood syndromes, often called MUS (medically unexplained symptoms) due to the lack of structural pathology. He suggested that these disorders be renamed Central Sensitivity Syndromes (CSS) and introduced the idea that CS may be a common feature causing similar and overlapping symptoms in these syndromes. In addition to a lack of structural pathology, most CSSs objectively share a lowered pain threshold and heightened pain sensitivity [15] which is a main feature of CS [16].

Mayer et al. [17] developed a patient-reported screening instrument called the Central Sensitization Inventory (CSI) to help identify and quantify CS/CSS-related symptomology. The concept of the CSI is based on Yunus’s [14] model of CSSs, in which different conditions with different phenotypes share overlapping symptoms related to CS. These symptoms were extracted from the CSS conglomerate via literature search and condensed in one questionnaire. The instrument has attracted growing attention and has been translated, culturally adapted and validated in different languages [18]. Some inconsistent results have been found regarding the dimensionality of the CSI in these initial validation studies (Supplement 1). To help settle the question of CSI dimensionality, a multi-country study with over 2000 subjects determined a bi-factor model, with one general factor of “CS-related symptoms” and four latent factors [19]. The validation of the underlying construct in different countries with different translations has been done in different ways or not addressed at all.

In this project, the CSI was first translated into the German language, pretested for comprehensibility, and culturally adapted. Then its psychometric properties were tested, including internal consistency, dimensionality, construct validity, and discriminative ability in a group of subjects with a broad spectrum of chronic musculoskeletal pain disorders and a separate group of pain-free control subjects.

Methods

Translation, cultural adaptation and pre-test

This study consisted of two parts. First, the original version of the CSI was forward (English to German) and backward (German translation back into English) translated, and cross-culturally adapted into German by an expert translation committee following the multistep approach recommended by the American Association of Orthopaedic Surgeons Outcomes Committee [20, 21]. A pretest was performed with 15 patients with chronic pain recruited at the pain clinic of the University Medical Center Göttingen. Based on the pretest results, the expert committee adapted two items, resulting in the final German version of the CSI (CSI-GE). The translation and cross-cultural adaptation of the CSI-GE has previously been presented elsewhere [22]. A detailed description of the steps involved in the translation is presented in Supplement 2. The German version of the CSI (CSI-GE) is presented in Supplement 3.

Clinical study for the psychometric validation

Secondly, a multicentre clinical study was conducted to assess the psychometric properties of the CSI-GE with four recruiting partners: the multidisciplinary pain clinic at the University Medical Center Göttingen; the pain clinic at Rotes Kreuz Krankenhaus Bremen; the MVZ Endokrinologikum (an outpatient rheumatology practice) based in Göttingen; and an outpatient practice for pain medicine, also located in Göttingen. The multicentre study design followed the COSMIN recommendations [23, 24] and was chosen to ensure the needed number of patients with a broad spectrum of pain-related diagnoses.

Ethical approval

For each involved institution ethical approval through the responsible ethics committee was provided: Ethics Committee of the University Medical Center Göttingen (15/09/2017) (including the pretests), Ethics Committee of the Ärztekammer Bremen (Antrags-Nr. 666–11/04/2019), and the Ethics Committee of the Ärztekammer Niedersachsen (Grae/055/2019).

The validation study was registered at the German Clinical Trials Register (DRKS-ID: DRKS00015252). All participants signed informed consent.

Participants

Patients with diverse musculoskeletal pain disorders were included in this study. Inclusion criteria were chronic pain for at least 3 months, sufficient physical and cognitive ability to participate, sufficient knowledge of the German language, and age older than 18 years. Exclusion criteria were psychiatric disease with pain as the main symptom (somatoform disorders, severe depression), initial or unstable phase of a rheumatological disease, or a primarily neuropathic pain component. All patients with chronic pain seeking routine care during the recruitment period were considered for participation and screened for eligibility by the responsible physicians of the individual case and at least one member of the study team. Eligible patients were informed about the study and asked for consent to participate.

In addition, a healthy control group (HC) was recruited from the general population via local contacts. For that purpose, members of a local choir were asked to participate in the study. To broaden the age range for the control group, students from the University Göttingen and acquaintances of the study team were asked to participate as well. The primary inclusion criteria were that they reported neither acute nor chronic pain. This was determined by an additional questionnaire asking for the presence of other than every day kind of pain (initial question of the brief pain inventory [25]) or pain-related conditions and was also explicitly assessed by a member of the study team during the recruitment. All other inclusion and exclusion criteria for the control participants were the same as for the clinical group.

To determine the appropriate sample size, we followed the recommendations for a factor analysis (FA), suggesting 4 to 10 participants per item [24] and more than 100 participants overall [26]. Interpreting this rule of thumb conservatively, 250 participants were needed because the CSI includes 25 items. We aimed to recruit 250 chronic pain patients (CPP) and 50 healthy control (HC) subjects.

Subgrouping of the chronic pain patients (CPP)

The chronic pain group was classified into five subgroups as described below. The subgrouping was based on clinical reasoning and on the main pain-related diagnosis and distribution of pain as used for classification in KEDOQ-Schmerz, a German pain register project for chronic pain based on the data provided by the widely-used German pain questionnaire [27, 28]. The German pain questionnaire includes multiple well-validated instruments. It serves for initial clinical assessment as well as progress assessment of pain patients. The results were available for all patients prior to inclusion into the present study. Diagnoses were made based on the results of the German pain questionnaire as well as an in-

person interdisciplinary assessment. Complicated cases with unclear classifications were discussed and resolved with the responsible physicians assigned to the individual cases and two of the authors (FP, MK), who were mainly responsible for the subgrouping. The following subgroups were presumed to represent different levels of CS [7] on a continuum, with Fibromyalgia representing the highest levels and regional chronic pain the lowest levels:

- The fibromyalgia (FMS) subgroup included patients with fibromyalgia as the primary diagnosis. Fibromyalgia was determined by the preliminary 2011 ACR criteria, including a full clinical assessment [29].
- Multisite chronic pain (MCP) included patients with a diverse spectrum of chronic pain disorders, with pain in at least three body regions.
- Regional chronic pain (RCP) included patients with a localized pain disorder in one well-defined body region (e.g., pain in the hand or foot only).
- Chronic back and/or neck pain (CBNP) included patients with a chronic pain syndrome in the back and/or neck region.
- In addition, patients with Rheumatoid arthritis in remission (RAR) were included if their rheumatologist clinically determined they were in disease remission at the time of their consultation. The level of CS in this subgroup was presumed to be lower than in the multisite chronic pain and fibromyalgia subgroups. All patients within this (RAR) subgroup were recruited at the outpatient rheumatological practice and evaluated by both the rheumatologist and one member of the study team (MK).

Questionnaires

Central sensitization inventory (CSI)

The CSI consists of two parts (A and B). Part A includes 25 items which assess typical symptoms associated with CS/CSS. Patients rate the degree of these symptoms on a five-point Likert scale ranging from never to always (never = 0, rarely = 1, sometimes = 2, often = 3, always = 4). The summation of all single item scores results in a total score ranging from 0 to 100. Part B inquires about 10 previously-diagnosed disorders in the patient's medical history, including seven common CSSs and three other conditions linked to CS/CSS. Part B of the CSI only aims to provide additional information and is not scored [17, 18]. We decided to quantify a sum score for part B by adding one point for each positively answered question resulting in a score ranging from 0 to 10. This score was used to correlate CSI parts A and B.

Pain sensitivity questionnaire minor (PSQm)

The PSQm is a condensed version of the PSQ which assesses individual subjective sensitivity to pain. The instrument includes 7 questions rated between 0 and 10. The mean rating of the 7 questions is used as the resulting score. The PSQm has previously demonstrated high correlations with sensitivity to experimental pain in healthy controls [30] and chronic pain patients [31].

Depression anxiety stress scale (DASS)

The DASS includes three scales which measure symptoms of depression, anxiety, and stress. The instrument contains 21 items (7 on each scale). All items can be rated from 0 to 3, higher scores represent more severe symptoms. The German version was validated in 2015 and has demonstrated acceptable psychometric properties for screening for depression, anxiety, and stress in patients with chronic pain [32].

Patient health questionnaire 15 (PHQ15)

The PHQ15 is a self-administered instrument that contains 15 items which assess the severity of somatic symptoms, indicating the individual degree of somatization. It has been translated into the German language and validated [33]. We used 13 items of the PHQ15 in order to have the same total score for women and men and to skip one problematic item that has often not been answered. Therefore, the questions asking for “menstrual cramps or other problems with your periods” and “pain or problems during sexual intercourse” were excluded. This resulted in a maximal score of 26 points for the adapted version.

painDETECT questionnaire

The painDETECT is a screening tool to detect a neuropathic pain component in chronic pain patients [34]. It has been developed and validated in collaboration with the German Research Network on Neuropathic Pain. As previously described, we used only one subscale of the painDETECT that includes 7 questions assessing neuropathic pain symptoms [35]. Questions are rated between never (0) and very strongly (5) resulting in a total score ranging from 0 to 35.

Pain Catastrophizing scale (PCS)

The PCS focuses on the quantification of catastrophizing attitudes and thoughts towards pain [36]. It contains 13 items which are rated on a 5-point Likert scale, resulting in an overall score ranging from 0 to 52. It has previously been translated and validated in the German language [37].

Fibromyalgia survey questionnaire (FSQ)

The FSQ is a self-administered tool to identify FMS in survey research without a physical examination. It has been validated for the German population [38] and consists of two subscales. One is the Widespread Pain Index (WPI), which assesses pain or tenderness at 19 different body parts, resulting in a total score between 0 and 19. The other subscale is the Somatic Severity Score (SSS). It captures the somatic symptom burden by inquiring about fatigue, trouble thinking, tiredness after waking up, pain in the lower abdomen, depression, and headache. SSS total scores range from 0 to 12.

Marburger questionnaire on habitual well-being (MFHW)

The MFHW is a short Questionnaire capturing perceived general wellbeing by addressing positive thoughts in 7 questions. The total score ranges from 0 to 35. Higher scores indicate a higher degree of habitual well-being [39].

Veterans Rand 12 (VR12)

The VR 12 [40] measures health-related quality of life and is very similar to the SF-12 (12-Item Short Form Health Survey) [41]. It results in two scores measuring the physical (physical composite summary - PCS) and the mental (mental composite summary - MCS) status separately. Higher scores indicate higher quality of physical or mental health-related quality of life. The VR12 has previously been translated and validated in the German language [42].

Graded chronic pain scale (VonKorff scale)

Von Korff et al. [43] introduced an instrument to grade the severity of chronic pain on a scale of 0 to 4 based on pain intensity and pain-related disability. The instrument also includes a numeric rating scale (0–10), which asks for the mean pain intensity within the last 4 weeks. The score of that subscale was used to indicate each subject’s perceived pain intensity.

Mainz pain staging system (MPSS)

The MPSS by Gerbershagen distinguishes three increasing degrees of pain chronicity by its associated features (like medication use, distribution and variability of pain, prior treatments) [1]. Throughout the past decades, it has been validated several times for different patient groups [44, 45].

In addition to the questionnaires, we included a question about the duration (in months) of each patient’s chronic pain status.

Data collection

Study participants had the choice to complete the questionnaires during a hospital admission, outpatient visit,

or at home (postage-prepaid envelope provided). All returned questionnaires were checked for missing items and participants were contacted to provide the missing answers. Participants were instructed in a standardized manner to complete missing items including the option not to answer.

Test-retest reliability

To analyse test-retest reliability, a time-interval of 2 weeks between two measurement occasions and a sample size of $n = 50$ was determined adequate, assuming an ICC (intraclass-correlation-coefficient) of 0.8 with a 95% CI of ± 0.1 [24]. A subgroup of pain patients ($n = 56$) received a second questionnaire for completion of the CSI-GE 2 weeks after the initial study visit. They were provided with a prepaid envelope and reminded after 3 weeks if the envelope had not been received by then. The test-retest time interval of 2 weeks, which has previously been recommended by de Vet et al. [24], seemed long enough to avoid patients remembering previous answers and short enough to avoid changes in health status affecting answers.

Statistical analysis

To assess the psychometric properties of the CSI-GE part A, both reliability and validity were investigated. Reliability was tested by analysing internal consistency and test-retest reliability. For assessment of the structural validity, both exploratory and confirmatory factor analyses (FA) were used to analyse the dimensionality of the questionnaire. To assess the construct validity, the relationship with other well-established clinical variables and questionnaires was analysed. In addition, to assess the discriminative validity, the ability of the CSI-GE to differentiate among different patient subgroups, believed to have different levels of CS (as described above) was investigated. The normality and data distributions were assessed by the skewness, kurtosis, histograms, Q-Q Plots, and Kolmogorov-Smirnov as well as Shapiro-Wilk tests. Demographic variables are described by means and standard deviation. Overall sex differences within the different subgroups were analysed using Fisher's-Exact-Test. Age differences were analysed using a Kruskal-Wallis-Test followed post hoc analysis for every specific subgroup comparison.

Floor and ceiling effects were determined by examining the prevalence of participants scoring the lowest and highest possible score on the CSI-GE sum score. As proposed by McHorney and Tarlov [46] the effects were considered relevant and problematic if observed in more than 15% of the participants.

Reliability

Internal consistency was evaluated using Cronbach's α . To evaluate test-retest reliability, the intraclass-correlation-coefficient_{2,1} (ICC_{2,1} – two-way random, absolute agreement, single measures) [47] was calculated for the CSI-GE part A overall sum score. Following KOO and Li [48] ICC values < 0.5 were considered poor, 0.5 to 0.75 moderate, 0.75 to 0.9 good, and > 0.9 excellent. The limits of agreement were analysed using a Bland Altman Plot. Therefore, the differences were plotted against the averages of every patient measurement pair included in the test-retest analysis. Standard Error of Measurement (SEM) and Smallest Detectable Change (SDC) were computed, using formulas proposed by de Vet et al. [24]. SEM was calculated by dividing the standard deviation of the difference between timepoint one and timepoint two, by the square root of two ($SEM = SD_{\text{difference}} / \sqrt{2}$). The SDC was calculated following the formula $SDC = \pm 1.96 * SEM * \sqrt{2}$.

Validity

Structural validity

Structural validity was analysed using FA. Exploratory factor analysis (EFA) using Maximum-Likelihood-Extraction with Promax-Rotation was carried out with data from the chronic pain sample to assess whether there would be a different factor structure than described in previous validation studies. Confirmatory factor analyses (CFA) were performed to assess the fit of three previously published models to our dataset, including both subject groups (CPP and HC). We assessed the fit to the original model by Mayer et al. [17] with four latent factors: “physical symptoms” (items 1, 2, 5, 6, 8, 9, 12, 14, 17, 18, 22), “emotional distress” (items 3, 13, 15, 16, 23, 24), “headache” (items 4, 7, 10, 19, 20) and “urological symptoms” (items 11, 21, 25); the fit to the 1-factor model by Cuesta-Vargas et al. [49] with one “general factor” that included all 25 items; and the fit to the bifactor model by Cuesta-Vargas et al. [19] that included both one general factor “CS related symptoms” (all 25 items) and 4 latent factors based on the four factors described by Mayer et al. [17]. The factor structure of all models can be found in Table 3. Diagonally weighted least squares estimation was used with listwise deletion of cases with missing information. Latent factors were standardized, allowing free estimation of all factor loadings. The Tucker-Lewis Index (TLI) and the root mean square error of approximation (RMSEA) with its 90% confidence interval were reported for each confirmatory model fit. Following Browne and Cudeck [50] a model fit was judged excellent for $RMSEA < 0.05$, good for $RMSEA < 0.08$, mediocre for $RMSEA < 0.1$, and poor for $RMSEA > 0.1$. Following Schermelleh-Engel et al. [51], a

model fit was judged acceptable for $TLI > 0.95$ and good for $TLI \geq 0.97$. Additionally, the factor loadings and (if applicable) factor correlations were analysed, each with the p -values from the respective significance tests. An ANOVA-type χ^2 difference test between the nested models was performed.

Construct validity – hypotheses testing

To assess construct validity 11 different hypotheses (Table 6) were formulated, and potential outcomes were postulated. This hypotheses testing used two approaches: (1) predefined differences between relevant subgroups based on differences in CSI-GE total score and (2) predefined correlations of the CSI-GE score with questionnaires that were selected based on potential clinical characteristics of CS as an overall construct. As proposed by Prinsen et al. [52] construct validity was considered satisfactory if $\geq 75\%$ of the hypotheses were met as predefined. In addition to the 11 hypotheses, further group comparisons and correlations with other instruments were analysed but not included in the hypotheses testing approach as a clear prediction of the outcome and relationship to the CSI-GE was not possible prior to our analysis. Therefore, the purpose of these additional measures was explorative aiming to further characterize the CSI-GE construct.

The discriminative power of the CSI-GE was assessed by examining differences among the six subgroups (FMS, MCP, CBNP, RAR, RCP, HC) using a Kruskal-Wallis test and Dunn-Bonferroni post-hoc-tests for non-parametric data. Bonferroni-Correction for multiple comparisons was used. A conservative non-parametric statistical approach was chosen for this analysis due to the violation of normal distribution and heterogeneity of variance between the different subgroups.

Pairwise correlations between total scores on each questionnaire and on the CSI-GE were calculated using Kendall's τ as a non-parametric correlation coefficient. Only data from the CPP were used, and pairwise deletion of missing cases was applied. Tests against the null hypothesis of no correlation were performed. Correction

for multiple testing was done using Holm's procedure. Following Cohen [53], correlations were considered small for Pearson's correlation coefficient $r > 0.1$, medium for $r > 0.3$, and large for $r > 0.5$. To use these categories for the differently scaled Kendall's τ , we translated these to the scale of Kendall's τ , using the formula given by Kendall [54]. Consequently, correlations with $\tau > 0.16$ were considered small, correlations with $\tau > 0.48$ were considered medium and correlations with $\tau > 0.82$ were considered large.

The significance level was set to $\alpha = 5\%$ for all statistical tests. The pairwise correlations with other questionnaires and the CFA were performed with the statistic software R version 3.6.1 [55] using the R-package lavaan version 0.6.5 [56] for the confirmatory factor analysis. Analyses for the demographics, the reliability, group comparisons, and the EFA were performed using IBM SPSS Statistics for Macintosh, Version 26.0, Released 2019 (Armonk, NY, USA: IBM Corp.).

Results

Demographic data and distribution

At the end of the recruitment phase, 346 individuals had been recruited to the study. Thirty-six datasets could not be included for different reasons listed in Table 1, resulting in 310 valid datasets for analysis. All exclusions were discussed and confirmed prior to analyses in a data validation meeting by the authors. The 310 analysed datasets consisted of 247 CPP and 63 HC. The retest was given to 56 CPP. However, only 45 valid retests were available for analyses. Reasons for the loss of 11 datasets are also listed in Table 1. The allocation of the 310 datasets to subgroups and the corresponding demographics are shown in Table 2. The allocation shows only a small group with regional pain; most patients had low back or neck pain, or some form of multisite pain.

As shown in Table 2, the mean age was 54.7 years for all individuals, with the HC group showing the lowest (49.4) and the RAR group showing the highest (59.8) mean age. The respective Kruskal-Wallis-Test showed the presence of an overall significance age difference

Table 1 Exclusions of datasets

	CPP	HC	Total CPP + HC	Retest
Agreed to participate, but did not send back the questionnaire	19	5		6
Excluded because:				
Diagnosis was not certain or in contrast to the recruitment criteria	7			
Missing items in CSI part A	4			1
Relevant Pain condition		1		
Significant health status changes within the retest time interval				4
Total of excluded datasets	30	6	36	11

Abbreviations: CPP chronic pain patients, HC healthy controls

Table 2 Subgroups & demographic figures

Different grouping	n	Age			Sex %female	CSI-GE part A			CSI-GE part B		
		mean	SD	Min/Max		mean	SD	Min/Max	mean	SD	Min/Max
All individuals	310	54.7	13.1	22/81	70.0	38.5	17.3	4/77	1.6	1.75	0/9
HC	63	49.4	15.6	22/78	58.7	18.4	9.3	4/41	0.3	0.8	0/3
CPP											
All CPP	247	56.1	12.0	22/81	72.9	43.6	15.0	7/77	1.9	1.8	0/9
FMS	37	55.0	12.3	27/77	94.6	54.9	11.7	22/77	3.6	2.0	0/9
MCP	63	56.8	11.4	37/81	74.6	49.8	11.8	11/74	2.5	1.6	0/8
CBNP	83	54.8	12.5	25/79	66.3	40.9	13.5	13/75	1.6	1.5	0/6
RAR	47	59.8	10.4	31/77	66.0	35.9	14.4	9/63	0.66	1.0	0/4
RCP	17	51.7	13.7	22/74	70.6	31.4	15.5	7/58	1.0	1.1	0/4

Abbreviations: HC healthy controls, CPP chronic pain patients, FMS fibromyalgia syndrome, MCP multisite chronic pain, CBNP chronic back and neck pain, RAR rheumatoid arthritis in status of remission, RCP regional chronic pain, n number of cases, SD standard derivation, Min/Max minimum/maximum

($p = 0.01$). However, post hoc comparisons showed that this age difference was only significant between the HC and RAR group. Seventy percent of the individuals were female, with 50.8% in the HC group and 94.6% in the FMS Group. Fisher’s exact test ($p = 0.002$) indicated that overall, sex was not distributed similarly within all subgroups. Correlating the CSI-GE sum score with age ($\tau = 0.19$; $p = 0,62$) and sex ($\tau = 0.22$; $p < 0,001$) showed only a weak correlation with age ($p = n.s.$) and weak correlation with sex. The mean CSI-GE total score in the total CPP sample was 43.6 and in the HC sample was 18.4. The FMS group had the highest mean CSI-GE score (54.9) and the HC group had the lowest. The FMS group also reported the highest number of summed responses (3.6) on CSI-GE part B, and the HC group reported the lowest number of responses (0.3).

No floor (sum score = 0) or ceiling effects (sum score = 100) occurred with total CSI-GE scores. The lowest score was 4 points by two participants (0,65%), and the highest score was 77 points by one participant (0,32%).

Reliability

Reliability analysis of the CSI-GE yielded a Cronbach’s α of 0.928, which can be considered high [24]. Excellent test-retest reliability was demonstrated by an ICC_{2,1} of 0.917 (95% KI: 0.855; 0.954) for the overall sum score of the CSI-GE, with a mean test-retest time interval of 18.42 (min. 14/max. 32) days. The limits of agreement can be observed in Fig. 1. The SEM amounted to 4.144 and the SDC to ± 11.486 .

Validity

Exploratory factor analysis

The Kaiser-Meyer-Olkin criterion (0.88) and Bartlett’s test ($p < 0.001$) showed an overall good [57] suitability of the data for the EFA. The EFA found five possible factors with an Eigenvalue > 1. The Eigenvalue decreased

strongly from the first factor (7.67) to the second factor (1.76) (scree plot Fig. 2), which indicated a 1-factor model. After rerunning the analysis and extraction of one factor, the 1-factor model was able to explain 27.94% of the variance. As demonstrated in Table 3, four items (4 = 0.39; 10 = 0.35; 11 = 0.38; 24 = 0.27) did not load above 0.4 on that single factor.

Confirmatory factor analysis

- A). Original 4-factor model proposed by Mayer et al. [17]: The model fit was good, with a TLI of 0.99, a RMSEA of 0.06 (90%-CI: [0.05; 0.07]), and $\chi^2(269) = 553.09$, $p < 0.001$. As expected, all items showed significant positive factor loadings, with standardized coefficients ranging from 0.438 to 0.861. As demonstrated in Table 4, there were also significant positive correlations among all four factors (physical symptoms, emotional distress, headache, urological symptoms), indicating that individuals who showed high scores in one dimension were also likely to demonstrate high scores in the other dimensions.
- B). The 1-factor model proposed by Cuesta-Vargas et al. [49]: The model fit was good, with a TLI of 0.98, a RMSEA of 0.08 (90%-CI: [0.07; 0.08]), and $\chi^2(275) = 756.39$, $p < 0.001$. All items showed significant positive factor loadings, with standardized coefficients ranging from 0.407 to 0.848.
- C). The Bifactor model proposed by Cuesta-Vargas et al. [19]: The model fit was excellent, with a TLI of 0.99, a RMSEA of 0.05 (90%-CI: [0.04; 0.06]), and $\chi^2(250) = 430.6$, $p < 0.001$. All items showed significant positive factor loadings for the general factor,

Table 3 Factor analysis - summary of different factor models

CSI-GE item	Factor analysis														
	EFA (n = 247)	CFA (n = 310)	1-factor model	4 factor model											
			1-factor model TL1 0.98 RMSEA 0.08 (90%-CI: [0.07; 0.08]) $\chi^2(275) = 756.39, p < 0.001$	4 factor model TL1 0.99 RMSEA 0.06 (90%-CI: [0.05; 0.07]) $\chi^2(269) = 553.09, p < 0.001$											
			1-factor model TL1 0.99 RMSEA 0.05 (90%-CI: [0.04; 0.06]) $\chi^2(250) = 430.6, p < 0.001$	bifactor model TL1 0.99 RMSEA 0.05 (90%-CI: [0.04; 0.06]) $\chi^2(250) = 430.6, p < 0.001$											
			General factor Cron. α 0.928 Cron. α 0.928	General factor Cron. α 0.928 Cron. α 0.928											
			Physical symptoms Cron. α 0.893 Cron. α 0.893	Physical symptoms Cron. α 0.893 Cron. α 0.893											
			Emotional distress Cron. α 0.785 Cron. α 0.785	Emotional distress Cron. α 0.785 Cron. α 0.785											
			Headache Cron. α 0.734 Cron. α 0.734	Headache Cron. α 0.734 Cron. α 0.734											
			Urological symptoms Cron. α 0.637 Cron. α 0.637	Urological symptoms Cron. α 0.637 Cron. α 0.637											
			General factor Cron. α 0.928 Cron. α 0.928	General factor Cron. α 0.928 Cron. α 0.928											
			Physical symptoms Cron. α 0.893 Cron. α 0.893	Physical symptoms Cron. α 0.893 Cron. α 0.893											
			Emotional distress Cron. α 0.785 Cron. α 0.785	Emotional distress Cron. α 0.785 Cron. α 0.785											
			Headache Cron. α 0.734 Cron. α 0.734	Headache Cron. α 0.734 Cron. α 0.734											
			Urological symptoms Cron. α 0.637 Cron. α 0.637	Urological symptoms Cron. α 0.637 Cron. α 0.637											
1. I feel unrefreshed when I wake up in the morning	0.669	0.796	0.786	0.777	0.163	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
2. My muscles feel stiff and achy	0.619	0.791	0.78	0.736	0.364	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
3. I have anxiety attacks	0.460	0.684	0.636	0.626	0.219	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
4. I grind or clench my teeth	0.391	0.608	0.493	0.436	0.585	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
5. I have problems with diarrhea and/or constipation	0.428	0.549	0.54	0.591	-0.251	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
6. I need help in performing my daily activities	0.477	0.666	0.655	0.618	0.315	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
7. I am sensitive to bright lights	0.446	0.688	0.563	0.553	0.238	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
8. I get tired very easily when I am physically active	0.737	0.859	0.848	0.805	0.35	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
9. I feel pain all over my body	0.634	0.796	0.786	0.737	0.375	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
10. I have headaches	0.353	0.624	0.513	0.494	0.308	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
11. I feel discomfort in my bladder and/or burning when I urinate	0.383	0.647	0.541	0.526	0.389 ^a	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
12. I do not sleep well	0.565	0.695	0.684	0.685	0.102 ^a	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
13. I have difficulty concentrating	0.728	0.861	0.795	0.772	0.487	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
14. I have skin problems such as dryness, itchiness, or rashes	0.429	0.504	0.496	0.54	-0.229	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
15. Stress makes my physical symptoms get worse	0.655	0.799	0.741	0.762	0.00822 ^a	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
16. I feel sad or depressed	0.563	0.735	0.684	0.668	0.301	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
17. I have low energy	0.735	0.852	0.839	0.816	0.241	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
18. I have muscle tension in my neck and shoulders	0.568	0.722	0.712	0.72	0.0584 ^a	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637
19. I have pain in my jaw	0.434	0.741	0.605	0.552	0.656	0.928	0.893	0.785	0.734	0.637	0.928	0.893	0.785	0.734	0.637

Table 3 Factor analysis - summary of different factor models (Continued)

CSI-GE item	Factor analysis	
	EFA (n = 247)	CFA (n = 310)
	1-factor model	4 factor model
	Cron. α 0.928	Cron. α 0.893
	TLI 0.99	TLI 0.99
	RMSEA 0.06 (90%-CI: [0.05; 0.07])	RMSEA 0.06 (90%-CI: [0.05; 0.07])
	$\chi^2(269) = 553.09, p < 0.001$	$\chi^2(269) = 553.09, p < 0.001$
	1-factor model	1-factor model
	Cron. α 0.928	Cron. α 0.928
	TLI 0.98	TLI 0.99
	RMSEA 0.08 (90%-CI: [0.07; 0.08])	RMSEA 0.05 (90%-CI: [0.04; 0.06])
	$\chi^2(275) = 756.39, p < 0.001$	$\chi^2(250) = 430.6, p < 0.001$
	General factor	General factor
	Cron. α 0.928	Cron. α 0.928
	Physical symptoms	Physical symptoms
	Cron. α 0.893	Cron. α 0.893
	Emotional distress	Emotional distress
	Cron. α 0.785	Cron. α 0.785
	Headache	Headache
	Cron. α 0.734	Cron. α 0.734
	Urological symptoms	Urological symptoms
	Cron. α 0.637	Cron. α 0.637
	General factor	General factor
	Cron. α 0.928	Cron. α 0.928
	Physical symptoms	Physical symptoms
	Cron. α 0.893	Cron. α 0.893
	Emotional distress	Emotional distress
	Cron. α 0.785	Cron. α 0.785
	Headache	Headache
	Cron. α 0.734	Cron. α 0.734
	Urological symptoms	Urological symptoms
	Cron. α 0.637	Cron. α 0.637
	General factor	General factor
	Cron. α 0.928	Cron. α 0.928
	Physical symptoms	Physical symptoms
	Cron. α 0.893	Cron. α 0.893
	Emotional distress	Emotional distress
	Cron. α 0.785	Cron. α 0.785
	Headache	Headache
	Cron. α 0.734	Cron. α 0.734
	Urological symptoms	Urological symptoms
	Cron. α 0.637	Cron. α 0.637
20. Certain smells, such as perfumes, make me feel dizzy and nauseated	0.418	0.647
21. I have to urinate frequently	0.403	0.661
22. My legs feel uncomfortable and restless when I am trying to go to sleep at night	0.472	0.605
23. I have difficulty remembering things	0.529	0.576
24. I suffered trauma as a child	0.268	0.438
25. I have pain in my pelvic area	0.470	0.767
	0.418	0.647
	0.403	0.661
	0.472	0.605
	0.529	0.576
	0.268	0.438
	0.470	0.767
	0.523	0.255
	0.544	0.796 ^a
	0.586	0.142 ^a
	0.482	0.549
	0.406	0.114 ^a
	0.652	0.106 ^a

Abbreviations: EFA exploratory factor analysis, CFA confirmatory factor analysis, n number of cases, Cron. α Cronbach's alpha, TLI Tucker-Lewis Index, RMSEA root mean square error of approximation, CI confidence interval

Acceptable values for the factor model indicators:

Following Browne and Cudeck (1992) model fit was judged excellent for RMSEA < 0.05, good for RMSEA < 0.08, mediocre for RMSEA < 0.1, and poor for RMSEA > 0.1

Following Schemmle-Engel et al. (2003) a model fit was judged acceptable for TLI > 0.95 and good for TLI ≥ 0.97

^aThe marked factor loadings were not significant

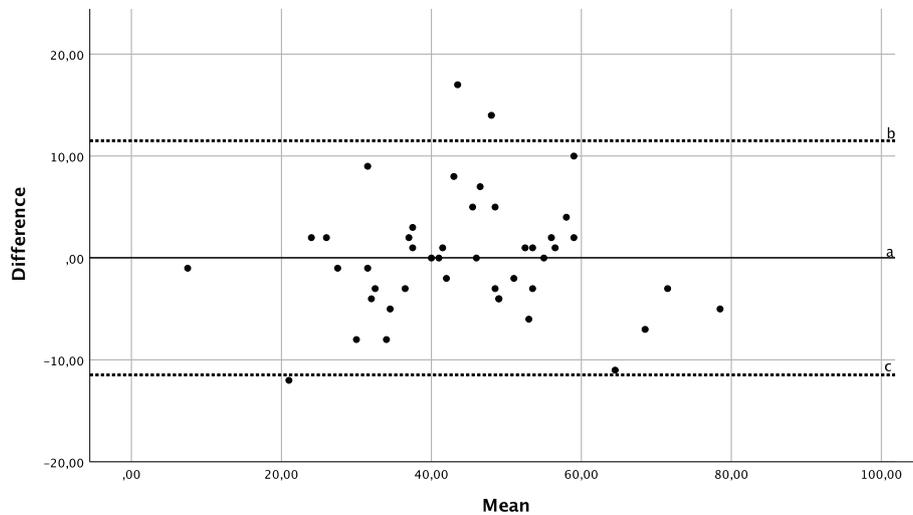


Fig. 1 Limits of agreement – Bland Altman Plot: The difference (y-axis) = CSI-GE sum score time point 1 – CSI-GE sum score time point 2 was plotted against the mean (x-axis) = (CSI-GE sum score time point 1 + CSI-GE sum score time point 2)/2 for every patient. Line a (=0.022) represents the mean systematic difference between the two time points. Line b represents a + 1.96*standard deviation_{difference} (=11.508). Line c represents a -1.96*standard deviation_{difference} (= - 11.464). Therefore, lines b and c show the limits of agreement enclosing 95% of the patients in between

but not all factor loadings were significant or positive for the other four latent factors, with standardized coefficients ranging from - 0.251 to 0.816.

Comparing the different models used in the CFA, the bifactor model fit the data significantly better ($\chi^2(19) = 115.1, p < .001$) than the original 4-factor model, while the original 4-factor model fit the data significantly better than the 1-factor model ($\chi^2(6) = 136.9, p < .001$).

Construct validity – discriminative power

The Kruskal-Wallis-Test indicated the presence of significant differences ($p < 0.001$) in CSI-GE total scores among the subject subgroups. Post-hoc analyses found that the control group scored lower than all the other groups except patients with only regional pain. Patients with multisite or FMS-related pain were not significantly different from each other but scored higher than the other groups (Table 5).

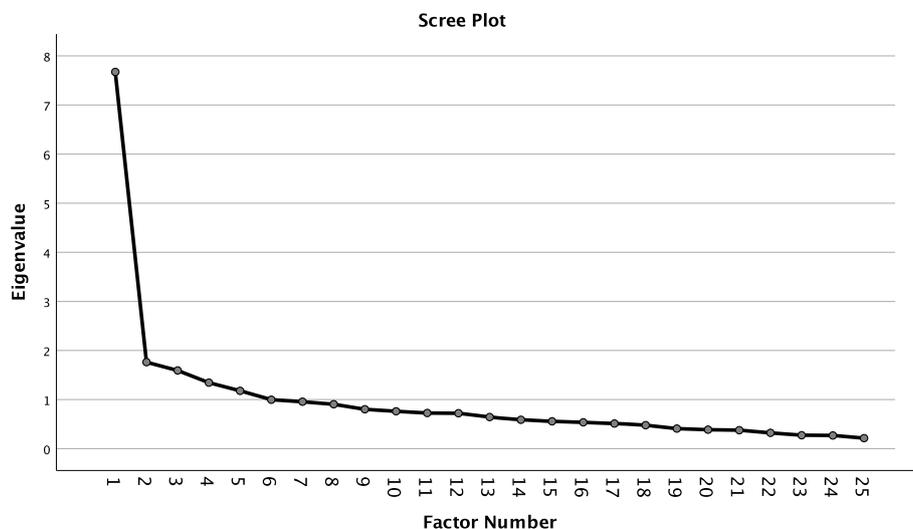


Fig. 2 Scree plot EFA

Table 4 Inter-Factor correlations of the CFA of the 4-factor model

Factor 1	Factor 2	Correlation
physical symptoms	emotional distress	0.87
physical symptoms	headache	0.83
physical symptoms	urological symptoms	0.84
emotional distress	headache	0.74
emotional distress	urological symptoms	0.69
headache	urological symptoms	0.61

Significance for all listed inter-factor correlation was ≤ 0.001

Construct validity - correlations

Using the adapted figures for Kendall’s τ for the classification, the CSI-GE demonstrated medium correlations with the PHQ-15 ($\tau = 0.57$) and FSQ-SSS ($\tau = 0.56$), and low correlations with the FSQ-WPI ($\tau = 0.47$), CSI-GE-Part-B ($\tau = 0.45$), DASS-Anxiety ($\tau = 0.45$), DASS-Stress ($\tau = 0.43$), PainDETECT-subscore ($\tau = 0.43$), DASS-Depression ($\tau = 0.41$), VonKorff ($\tau = 0.36$), VR12-MCS ($\tau = -0.35$), MPSS ($\tau = 0.32$), general well-being ($\tau = -0.28$), PCS ($\tau = 0.28$), pain intensity ($\tau = 0.27$), PSQM ($\tau = 0.23$) and negligible or no significant correlations with pain duration and VR12-PCS (Supplement 4).

Table 5 Results of the Dunn-Bonferroni post-hoc test – pairwise comparisons of the groups

Sample 1 – Sample2	Std. Test Statistic	Adjusted Significance
Significant pairwise comparisons		
HC - RAR	-5.13	< 0.001
HC - CBNP	-7.49	< 0.001
HC - MCP	-10.14	< 0.001
HC - FMS	-10.2	< 0.001
RCP - FMS	4.69	< 0.001
RCP - MCP	3.9	0.001
RAR - MCP	4.24	< 0.001
RAR - FMS	5.11	< 0.001
CBNP - FMS	4.36	< 0.001
CBNP - MCP	3.32	0.013
Not significant pairwise comparisons		
HC - RCP	-2.71	0.101
RCP - CBNP	-1.92	0.827
RCP - RAR	-0.88	1.0
RAR - CBNP	1.43	1.0
MCP - FMS	1.48	1.0

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is .05. Significance values have been adjusted by the Bonferroni correction for multiple tests (adjusted significance) *Abbreviations: HC* healthy controls, *FMS* fibromyalgia syndrome, *MCP* multistate chronic pain, *CBNP* chronic back and neck pain, *RAR* rheumatoid arthritis in status of remission, *RCP* regional chronic pain

Construct validity – hypotheses testing

The results of the hypotheses testing approach can be observed within Table 6.

Discussion

The objective of this study was to create a culturally-adapted German version of the CSI and to test its psychometric properties. Internal consistency of different CSI translations has been examined in most international studies (Supplement 1) using Cronbach’s α , which has ranged from 0.87 [58, 59] to 0.993 [60], and is in agreement with our result of a Cronbach’s α of 0.928. Test-retest reliability has been examined using Pearson’s correlation, with $r_p = 0.817$ [17] and intra-class correlation, with results ranging from 0.85 [61] to 0.991 [60], which was in line with the ICC_{2,1} of 0.917 in our analysis. SEM and SDC were computed only by a few of the initial validation studies. The standard error of measurement (SEM) ranged from 0.31 [59] to 3.16 [62], whereas the smallest detectable change (SDC), which is also called minimal detectable change (MDC), ranged from 0.86 [59] to 8.12 [62]. Therefore, the values of SEM = 4.144 and SDC = ± 11.486 in our study seem high. However, the time interval with a mean of 18.42 days between the two measurement time points was longer than in the other studies. This longer interval may have reduced memory effects but may also represent relevant fluctuations in symptoms over the longer (baseline) observation interval since all patients reporting relevant changes in health prior to the retest were excluded.

A 4-factor structure of the CSI was originally determined by Mayer et al. [17] and supported by further international studies [10, 62, 63]. Other studies have demonstrated a 1-factor [49, 58] or 5-factor structure [61]. Due to the diverse reports of the factor structure of different CSI translations, Cuesta-Vargas et al. [19] performed a large FA with pooled data from multiple countries and multiple language-versions of the CSI. They demonstrated that the best fit was a bifactor model, with one general factor of “CS-related symptoms” and four latent factors. Considering that the bifactor model provided the best fit in our CFA and the EFA yielded a 1-factor solution, it seems

Table 6 Construct validity – Hypotheses testing

Hypothesis- Group Comparison	Expected Outcome	Results	Was the Hypothesis met?
CSI score of the HC group	HC group scoring the lowest	HC group showed lowest mean CSI-GE score, significantly different from all but the RCP group	Yes
CSI score of the FMS group	FMS group scoring the highest of all pain groups	highest mean CSI-GE score, significantly different from all but the MCP group	Yes
CSI score of the MCP group	MCP group scoring high in comparison to other pain groups but probably lower than the FMS group	Second highest mean CSI-GE score of all 5 pain groups	Yes
CSI score of the RCP group	RCP group scoring low in comparison to other pain groups	Lowest mean CSI-GE score of all 5 pain groups but higher than HC group (although not significant)	Yes
CSI Score of the RAR group	RAR group Scoring lower in comparison to FMS and MCP	Lower mean CSI-GE score than in FMS and in MCP groups	Yes
Questionnaire- Correlations with CSI-GE Score	Expected Outcome	Kendal's τ , the associated Holm adjusted p value and classification	Was the Hypothesis met?
CS is related to subjective pain sensitivity	High correlation between CSI-GE and PSQminor	$\tau = 0.23$; $p < 0.001$; low	No
CS is related to somatization or general symptom load	At least medium correlation between CSI-GE and (a) PHQ15 and (b) FSQ subscale SSS	a: $\tau = 0.57$; $p < 0.001$; medium b: $\tau = 0.56$; $p < 0.001$; medium	Yes
CS is related to spread of pain	At least medium correlation between CSI-GE and FSQ subscale WPI	$\tau = 0.47$; $p < 0.001$; low	No
CS is related to neuropathic pain characteristics	At least medium correlation between CSI-GE and PainDETECT	$\tau = 0.43$; $p < 0.001$; low	No
CS is related to chronification of pain	At least medium correlation between CSI-GE and MPSS	$\tau = 0.32$; $p < 0.001$; low	No
CS is related to pain catastrophizing	At least medium correlation between CSI-GE and PCS	$\tau = 0.28$; $p < 0.001$; low	No
How many Hypotheses were met?			6/11 (55%)

Abbreviations: CS central sensitization, HC healthy control, FMS Fibromyalgia, MCP multisite chronic pain, RCP regional chronic pain, RAR Rheumatoid arthritis in remission, PHQ15 Patient Health Questionnaire 15, FSQ-SSS Fibromyalgia Survey Questionnaire-Somatic Severity Score, FSQ-WPI Fibromyalgia Survey Questionnaire-Widespread Pain Index, DASS Depression Anxiety Stress Scale, PCS Pain Catastrophizing Scale, MPSS Mainz Pain Staging System, PSQm Pain Sensitivity Questionnaire minor

justified to only compute an overall sum score for the CSI-GE, representing one universal general construct underlying all items. The four remaining latent factors within the bifactor-model suggest an underlying structure of specific factors that enclose these specific features. Our analysis does not support the use of subscales with the CSI-GE due to the non-significant or low loadings of the four latent factors (Table 3) and the questionable additional benefit of subscales. These findings are in line with Cuesta-Vargas et al. [19] who recommended that only total CSI scores be used and reported.

Assessing construct validity with the hypotheses testing approach only 55% of the hypotheses were met, which can not be considered an entirely persuasive and satisfactory result. However, the construct of the CSI-GE showed convincing evidence regarding the group comparisons and overall symptom load (PHQ, SSS), but demonstrated limited construct validity with respect to hypotheses more closely related to the construct of central sensitisation itself. Nevertheless, only the PHQ and

SSS can be considered comparator instruments, the other questionnaires measured potentially related yet different constructs than the CSI. Subsequently, individual hypotheses, as well as the explorative correlations are further discussed.

The CSI-GE construct was supported by its ability to differentiate between subgroups believed to have different degrees of CS. The FMS and MCP groups scored significantly higher than all other groups except each other. This was expected as those groups were believed most likely to have the highest degree of CS. Also, the HC group differed significantly from all pain groups except the RCP group, which was believed to most likely have lower levels of CS. These finding are supported by Knezevic et al. [64], who found similar results in FMS, multiple pain sites, and localized pain subgroups.

However, correlations with other instruments were not as unequivocal. The highest correlations with somatic symptom severity may partly be explained by

overlapping items between instruments. FMS patients, known to score highly on these instruments and to embody features of CS [65], were the highest scoring group in our study and indicate that symptom load is a key aspect of the CSI-GE construct. This symptom load may otherwise simply reflect the degree of polysymptomatic distress as suggested by Wolfe et al. for FMS [66] and RA [67].

All three scales of the DASS showed a low, positive correlation with the CSI-GE, demonstrating an increasing negative affective/emotional burden with increasing scores in the CSI-GE. Chiarotto et al. [58] showed positive correlations with depression and anxiety measures with the Italian version of the CSI as well.

The correlation between the CSI-GE and the painDETECT showed a small association with neuropathic pain components. Rehm et al. [35] found that the painDETECT indicated possible neuropathic pain in more than 50% of FMS patients. It is unclear whether this hints at neuropathic pain components in FMS patients, the presence of CS, or simply increased somatic symptoms. It is unlikely that neuropathic pain was a primary symptom in our patient sample since we attempted to exclude patients with a primary neuropathic pain component. The correlation between the painDETECT and the CSI-GE may suggest that either CS-related symptoms overlap with neuropathic pain features, or that the CSI truly represents symptoms related to CS and that the painDETECT results in false-positive detection of neuropathic pain in the presence of CS. This second assumption has been supported by a qualitative study [68].

The degree of chronification (MPSS) and severity (Korff) of chronic pain both correlated poorly with the CSI-GE. The slightly higher correlation with the Von Korff scale is possibly explained with pain-related disability, which is considered within the score. This is supported by findings of Kregel et al. [69] who showed positive correlations between the CSI and impairment caused by pain in daily life.

The PCS correlated to a small degree with the CSI-GE. This was unexpected as we believed that the concept of catastrophizing is an amplifying factor for CS, and previous CSI studies have reported higher correlations with the PCS [60, 64]. Correlations with pain sensitivity and intensity were low. The latter was also found by Knezevic et al. [64]. This appears somewhat counter-intuitive, as CS is believed to trigger higher pain intensity and decreased thresholds to painful stimulation by increasing the sensitivity of the somatosensory system. Therefore, we expected a higher correlation with the PSQm, which captures individual sensitivity towards pain. In line with the low correlation with the PSQm, Kregel et al. [69] reported weak correlations between

CSI total scores and pressure pain thresholds (PPT). Also Hendriks et al. [70] found that the CSI did not correlate significantly with PPT and concluded that the CSI captures the psychopathology associated with CS and not neurobiological alterations.

The CSI-GE demonstrated no association with the duration of pain. This finding is in line with other CSI studies [59, 61]. One might argue that with a longer chronic pain state, CS is expected to increase due to a higher degree of afferent input sensitizing the nervous system. On the other hand, CS could be understood as a self-maintaining process once initiated and not influenced by duration.

Our results support the negative correlations between mental quality of life and the CSI scores that have been found in previous studies [61, 64, 69]. Surprisingly in our study, the physical quality of life did not correlate significantly with the CSI-GE. This is in contrast to other validation studies showing significant negative correlations with the physical quality of life [58, 64, 69]. Considering that the CSI-GE correlated highest with instruments measuring the degree of somatic symptom load, a negative correlation with physical quality of life had been expected. However, Knezevic et al. [64] found lower correlations with SF-36-PCS than with SF-36-MCS supporting the finding that the mental component seems more prominent.

After the responses on CSI part B were summed, the FMS group reported the highest number of CSS-related diagnoses (3.6) and the HC group reported the lowest number (0.3). These results are in line with previous studies that have assessed CSI part B [17, 63]. Though we found a positive correlation between total scores on part A and the summed responses of part B of the CSI-GE, the correlation was low, considering that part B lists typical CSSs and related disorders, and part A includes symptoms associated with CSSs. However, these results should be interpreted with caution, as patients reported difficulties answering the questions in part B, which asks for previous diagnoses made by a physician. Patients may have marked diagnoses based on their subjective opinion and understanding of the respective label. Also, the comparability of the criteria a diagnosis was based on is unclear. Nevertheless, as noted previously, CSI part B is not designed to be scored when used clinically. It only provides additional information to help identify when a patient's symptom presentation may be related to CS or be indicative of a CSS.

Limitations

As with all studies of this kind, the results are based on a limited sample of subjects within two regions of Germany, so the findings may not generalize to other populations. More women than men were recruited as

in previous CSI validation studies [10, 58], which is in accordance with more women being affected by musculoskeletal pain syndromes [71]. Slight age and sex differences between the subgroups may have influenced the group comparisons. However, our analysis demonstrated a very weak relationship between CSI-GE scores and age. A significant age difference was observed in only in one subgroup comparison (HC – RAR).

As there was no formal a priori definition of the hypothesis, the selected cut-off points (like high or medium correlation) remain somewhat arbitrary and of limited validity. The possibility of treatment effects, prior to data collection, on CSI scores needs to be considered. This highlights a potential limitation in ours, as well as other CSI validation studies, as they include little information about previous treatments such as psychotherapy, interdisciplinary multimodal pain therapy, or drug therapies that could affect questionnaire responses [72, 73]. One previous study has demonstrated that the CSI can be responsive to treatment interventions, as CSI scores improved in a group of chronic spinal pain patients who completed a functional restoration program [74]. The comparison of our results with previous studies must be interpreted cautiously because different correlation coefficients have been used and Kendall's τ tends to be smaller in magnitude than Spearman rho correlation coefficient [75]. Our study did not include measures like QST to quantify the patients' sensitivity and pain status. This should be explored in future studies. Although the calculation of overall sum scores for the CSI-GE, other CSI versions, and other patient-reported outcome measures is well established, it should be acknowledged that the summation of ordinal measured items must be viewed critically and is a point of controversy in the scientific literature [76, 77].

Conclusion

The CSI-GE demonstrated robust psychometric properties as well as solid reliability. Based on the results of our factor analysis, it thus seems justified to compute one overall sum score. Our construct validation assessment suggests that the questionnaire reflects a dimension that no other tool that we compared has captured in the same way, although high symptom load was a prominent overlapping feature. Some of the correlations were unexpected within the current understanding of CS. It remains uncertain to which degree the CSI-GE captures CS and quantifies its symptoms. Combining our findings with previously published research regarding the CSI, we can conclude that interpretation of the total CSI score is made more difficult because definitions of CS are diverse; symptoms are broad and overlapping in a variety of conditions and may indicate

polysymptomatic distress; new concepts such as nociplastic or chronic primary pain may be insufficiently considered; and no gold standard exists for the clinical or experimental quantification of CS. Experimental studies and studies examining the responsiveness to interventions in well-characterized patient groups may help to better define the scope of the instrument.

In conclusion, we recommend using the CSI-GE in clinical practice only with caution and primarily as a screening for symptoms that may be related to CS. The concept of CS requires further clarification within a research context. There is currently no established clinical diagnostic or treatment pathway in case of a positive screening result using the CSI-GE.

Abbreviations

CBNP: Chronic back and/or neck pain; CFA: Confirmatory factor analyses; CPP: Chronic pain patients; CSI: Central Sensitization Inventory; CSI-GE: German version of the Central Sensitization Inventory; CS: Central Sensitization; CSS: Central Sensitivity Syndromes; DASS: Depression Anxiety Stress Scale; EFA: Exploratory factor analysis; FA: Factor analysis; FMS: Fibromyalgia; FSQ: Fibromyalgia Survey Questionnaire; FSQ-WPI: Fibromyalgia Survey Questionnaire - Widespread Pain Index; FSQ-SSS: Fibromyalgia Survey Questionnaire - Somatic Severity Score; HC: Healthy control group; ICC: Intraclass-correlation-coefficient; MCP: Multisite chronic pain; MFHW: Marburger questionnaire on habitual well-being; MPSS: Mainz Pain Staging System; PCS: Pain Catastrophizing Scale; PHQ15: Patient Health Questionnaire 15; PPT: Pressure pain thresholds; PSQm: Pain Sensitivity Questionnaire minor; QST: Quantitative Sensory Testing; RAR: Rheumatoid arthritis in remission; RCP: Regional chronic pain; SDC: Smallest Detectable Change; SEM: Standard Error of Measurement; VonKorff Scale: Graded Chronic Pain Scale by VonKorff; VR12: Veterans Rand 12; VR12-PCS: Veterans Rand 12-physical composite summary; VR12-MCS: Veterans Rand 12-mental composite summary

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12891-021-04481-5>.

Additional file 1: Supplement 1. Different international CSI validation studies in different languages.

Additional file 2: Supplement 2. The German translation and cross-cultural adaptation of the Central Sensitization Inventory (CSI-GE).

Additional file 3: Supplement 3. CENTRAL SENSITIZATION INVENTORY German version (CSI-GE)_V_19.09.2020.

Additional file 4: Supplement 4. Visualized pairwise correlations between the CSI-GE part A sum score and each questionnaire.

Acknowledgements

We thank Dr. Heppner C., Hagenguth-Görs A. and the staff in all four involved institutions for their support with the data collection as well as Hogan D., Laporte Uribe F., Mazolek U., Paul U., & Seeger D for their cooperation in the translation team.

Authors' contributions

All authors discussed the results and commented on the manuscript. Klute, Michel: Institution: Pain Clinic, Department of Anaesthesiology, University Medical Center Göttingen, Germany. Contributions: Conception and design, Data acquisition, Data analysis, Writing- original draft preparation. Laekeman, Marjan: Institution: Physiological Psychology, Otto-Friedrich- University of Bamberg, Germany. Contributions: Conception and design, Data analysis, Writing- review and editing. Kuss, Katrin: Institution: Department of General Practice/Family Medicine, Philipps University Marburg, Germany. Contributions: Conception and design, Data analysis, Writing- review and editing.

Petzke, Frank: Institution: Pain Clinic, Department of Anaesthesiology, University Medical Center Göttingen, Germany. Contributions: Conception and design, Data acquisition, Data analysis, Writing- review and editing. Dieterich, Angela: Institution: Physiotherapy, Faculty of Health, Safety, Society, Furtwangen University, Germany. Contributions: Conception and design, Data analysis, Writing- review and editing. Leha, Andreas: Institution: Department of Medical Statistics, University Medical Center Göttingen, Germany. Contributions: Data analysis, Writing- review and editing. Neblett, Randy: Institution: PRIDE Research Foundation, Dallas, Texas, USA. Contributions: Conception and design, Writing- review and editing. Ehrhardt, Steffen: Institution: Faculty of Social Sciences, City University of Applied Sciences, Bremen, Germany. Contributions: Conception and design, Writing- review and editing. Ulma, Joachim: Institution: Clinic for Pain Medicine Bremen, Rotes-Kreuz-Krankenhaus Bremen, Germany. Contributions: Data acquisition, Writing- review and editing. Schäfer, Axel: Institution: Faculty of Social Work and Health, University of Applied Science and Art, Hildesheim, Germany. Contributions: Conception and design, Data analysis, Writing- review and editing. The authors read and approved the final manuscript.

Funding

This project was financially supported by the German Association for Manual Therapy (DVMT). The DVMT did not influence the content of this project in any way. Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available as they are currently also used on a dissertation. After completion, they will be publicly available. Nevertheless, the datasets can also be preliminary available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

All methods were performed in accordance with the relevant guidelines and regulations.

They were approved by the responsible ethics committees: Ethics Committee of the University Medical Center Göttingen (15/09/2017) (including the pretests); Ethics Committee of the Ärztekammer Bremen (Antrags-Nr. 666–11/04/2019); Ethics Committee of the Ärztekammer Niedersachsen (Grae/055/2019).

The validation study was registered at the German Clinical Trials Register (DRKS-ID: DRKS00015252).

All participants signed informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Pain Clinic, Department of Anaesthesiology, University Medical Center, Georg August University of Göttingen, Robert-Koch-Str. 40, 37075 Göttingen, Germany. ²Physiological Psychology, Otto-Friedrich- University of Bamberg, Bamberg, Germany. ³Department of General Practice/Family Medicine, Philipps University Marburg, Marburg, Germany. ⁴Physiotherapy, Faculty of Health, Safety, Society, Furtwangen University, Furtwangen, Germany. ⁵Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany. ⁶PRIDE Research Foundation, Dallas, TX, USA. ⁷Faculty of Social Sciences, City University of Applied Sciences, Bremen, Germany. ⁸Clinic for Pain Medicine Bremen, Rotes-Kreuz-Krankenhaus Bremen, Bremen, Germany. ⁹Faculty of Social Work and Health, University of Applied Science and Art, Hildesheim, Germany.

Received: 11 February 2021 Accepted: 14 June 2021

Published online: 18 August 2021

References

- Nilges P, Nagel B. Was ist chronischer Schmerz? Dtsch Med Wochenschr. 2007;132(41):2133–8. <https://doi.org/10.1055/s-2007-985653>.

- Pak DJ, Yong RJ, Kaye AD, Urman RD. Chronification of pain: mechanisms, current understanding, and clinical implications. *Curr Pain Headache Rep*. 2018;22(2):9. <https://doi.org/10.1007/s11916-018-0666-8>.
- Treede R-D, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. A classification of chronic pain for ICD-11. *Pain*. 2015;156(6):1003–7. <https://doi.org/10.1097/j.pain.000000000000160>.
- Treede R-D, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. Chronic pain as a symptom or a disease: the IASP classification of chronic pain for the international classification of diseases (ICD-11). *Pain*. 2019;160(1):19–27. <https://doi.org/10.1097/j.pain.0000000000001384>.
- Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. *Pain*. 2011;152(3 Supplement):S2–15. <https://doi.org/10.1016/j.pain.2010.09.030>.
- Gatchel RJ, Neblett R. Central sensitization: a brief overview. *J Appl Biobehav Res*. 2018;23(2):e12138. <https://doi.org/10.1111/jabr.12138>.
- Nijs J, George SZ, Clauw DJ, Fernández-de-las-Peñas C, Kosek E, Ickmans K, et al. Central sensitisation in chronic pain conditions: latest discoveries and their potential for precision medicine. *Lancet Rheumatol*. 2021;3(5):e383–92. [https://doi.org/10.1016/S2665-9913\(21\)00032-1](https://doi.org/10.1016/S2665-9913(21)00032-1).
- International Association for the Study of Pain (IASP). IASP Terminology. 2017. <https://www.iasp-pain.org/terminology>. Accessed 18 Sept 2020.
- Bid DD, Soni NC, Rathod PV, Ramalingam AT. Content validity and test-retest reliability of the Gujarati version of the central sensitization inventory. *Natl J Integr Res Med*. 2016;7:18–24.
- Kregel J, Vuijk PJ, Descheemaeker F, Keizer D, van der Noord R, Nijs J, et al. The Dutch central sensitization inventory (CSI): factor analysis, discriminative power, and test-retest reliability. *Clin J Pain*. 2016;32(7):624–30. <https://doi.org/10.1097/AJP.0000000000000306>.
- den Boer C, Dries L, Terluin B, van der Wouden JC, Blankenstein AH, van Wilgen CP, et al. Central sensitization in chronic pain and medically unexplained symptom research: a systematic review of definitions, operationalizations and measurement instruments. *J Psychosom Res*. 2019;117:32–40. <https://doi.org/10.1016/j.jpsychores.2018.12.010>.
- Nijs J, Van Houdenhove B, Oostendorp RAB. Recognition of central sensitization in patients with musculoskeletal pain: application of pain neurophysiology in manual therapy practice. *Man Ther*. 2010;15(2):135–41. <https://doi.org/10.1016/j.math.2009.12.001>.
- Akinci A, Shaker MA, Chang MH, Cheung CW, Danilov A, Dueñas HJ, et al. Predictive factors and clinical biomarkers for treatment in patients with chronic pain caused by osteoarthritis with a central sensitisation component. *Int J Clin Pract*. 2016;70(1):31–44. <https://doi.org/10.1111/ijcp.12749>.
- Yunus MB. Fibromyalgia and overlapping disorders: the unifying concept of central sensitivity syndromes. *Semin Arthritis Rheum*. 2007;36(6):339–56. <https://doi.org/10.1016/j.semarthrit.2006.12.009>.
- Aaron LA, Buchwald D. A review of the evidence for overlap among unexplained clinical conditions. *Ann Intern Med*. 2001;134(9_Part_2):868–81.
- Latreoliere A, Woolf CJ. Central sensitization: a generator of pain hypersensitivity by central neural plasticity. *J Pain*. 2009;10(9):895–926. <https://doi.org/10.1016/j.jpain.2009.06.012>.
- Mayer TG, Neblett R, Cohen H, Howard KJ, Choi YH, Williams MJ, et al. The development and psychometric validation of the central sensitization inventory. *Pain Pract*. 2012;12(4):276–85. <https://doi.org/10.1111/j.1533-2500.2011.00493.x>.
- Neblett R. The central sensitization inventory: a user's manual. *J Appl Biobehav Res*. 2018;23(2):e12123. <https://doi.org/10.1111/jabr.12123>.
- Cuesta-Vargas AI, Neblett R, Chiarotto A, Kregel J, Nijs J, van Wilgen CP, et al. Dimensionality and reliability of the central sensitization inventory in a pooled multicountry sample. *J Pain*. 2018;19(3):317–29. <https://doi.org/10.1016/j.jpain.2017.11.006>.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25:3186–91.
- Laekeman M, Kuss K, Seeger D, Schäfer A. Zentrale Sensibilisierung erkennen. In: *Der Central Sensitization Inventory wird ins Deutsche übersetzt und validiert*. München: pt Zeitschrift für Physiotherapeuten; 2017. p. 71–3.
- Laekeman M, Ehrhardt S, Kuss K, Petzke F, Dieterich A, Neblett R, et al. Expert and Patient perspectives on the cross-cultural translation and adaptation of the Central Sensitization Inventory into German; 2019. <https://doi.org/10.26226/MORRESSIER.5D4434D306F5C40FC448474D>.

23. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol*. 2010;10(1):22. <https://doi.org/10.1186/1471-2288-10-22>.
24. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in medicine: a practical guide*. Cambridge: Cambridge University Press; 2011. <https://doi.org/10.1017/CBO9780511996214>.
25. Radbruch L, Loick G, Kiencke P, Lindena G, Sabatowski R, Grond S, et al. Validation of the German version of the brief pain inventory. *J Pain Symptom Manag*. 1999;18(3):180–7. [https://doi.org/10.1016/S0885-3924\(99\)00064-0](https://doi.org/10.1016/S0885-3924(99)00064-0).
26. Kline P. *The handbook of psychological testing*. 2nd ed. London: Routledge; 2000.
27. Hüppe M, Kükenhöner S, Bosse F, Casser HR, Kohlmann T, Lindena G, et al. Schmerztherapeutische Versorgung in Deutschland – was unterscheidet ambulante und stationäre Patienten zu Behandlungsbeginn?: Eine Auswertung auf Basis des KEDOO-Schmerz-Datensatzes. *Schmerz*. 2017; 31(6):559–67. <https://doi.org/10.1007/s00482-017-0240-z>.
28. Hüppe M, Kükenhöner S, Böhme K, Bosse F, Casser H-R, Kohlmann T, et al. Schmerztherapeutische Versorgung in Deutschland – unterscheiden sich teilstationär versorgte Patienten von den ambulant oder stationär versorgten bei Behandlungsbeginn?: Eine weitere Auswertung auf Basis des KEDOO-Schmerz-Datensatzes. *Schmerz*. 2020;34(5):421–30. <https://doi.org/10.1007/s00482-020-00480-y>.
29. Wolfe F, Clauw DJ, Fitzcharles M-A, Goldenberg DL, Häuser W, Katz RS, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR preliminary diagnostic criteria for fibromyalgia. *J Rheumatol*. 2011;38(6):1113–22. <https://doi.org/10.3899/jrheum.100594>.
30. Ruscheweyh R, Marziniak M, Stumpfenhorst F, Reinholz J, Knecht S. Pain sensitivity can be assessed by self-rating: Development and validation of the Pain Sensitivity Questionnaire. *Pain*. 2009;146(1-2):65–74. <https://doi.org/10.1016/j.pain.2009.06.020>.
31. Ruscheweyh R, Verneuer B, Dany K, Marziniak M, Wolowski A, Colak-Ekici R, et al. Validation of the pain sensitivity questionnaire in chronic pain patients. *Pain*. 2012;153(6):1210–8. <https://doi.org/10.1016/j.pain.2012.02.025>.
32. Nilges P, Essau C. Die Depressions-Angst-Stress-Skalen: Der DASS – ein Screeningverfahren nicht nur für Schmerzpatienten. *Schmerz*. 2015;29(6): 649–57. <https://doi.org/10.1007/s00482-015-0019-z>.
33. Löwe B, Spitzer RL, Zipfel S, Herzog W. *Gesundheitsfragebogen für Patienten (PHQ-D)*. Manual und Testunterlagen. 2nd ed. Karlsruhe: Pfizer; 2002. https://www.klinikum.uni-heidelberg.de/fileadmin/Psychosomatische_Klinik/download/PHQ_Manual1.pdf
34. Freynhagen R, Baron R, Gockel U, Tölle TR. Pain detect: a new screening questionnaire to identify neuropathic components in patients with back pain. *Curr Med Res Opin*. 2006;22(10):1911–20. <https://doi.org/10.1185/030079906X132488>.
35. Rehm SE, Koroschetz J, Gockel U, Brosz M, Freynhagen R, Tolle TR, et al. A cross-sectional survey of 3035 patients with fibromyalgia: subgroups of patients with typical comorbidities and sensory symptom profiles. *Rheumatology (Oxford)*. 2010;49(6):1146–52. <https://doi.org/10.1093/rheumatology/keq066>.
36. Sullivan MJL, Bishop SR, Pivik J. The pain Catastrophizing scale: development and validation. *Psychol Assess*. 1995;7(4):524–32. <https://doi.org/10.1037/1040-3590.7.4.524>.
37. Meyer K, Sprott H, Mannion AF. Cross-cultural adaptation, reliability, and validity of the German version of the pain Catastrophizing scale. *J Psychosom Res*. 2008;64(5):469–78. <https://doi.org/10.1016/j.jpsychores.2007.12.004>.
38. Häuser W, Jung E, Erbslöh-Möller B, Gesmann M, Kühn-Becker H, Petermann F, et al. Validation of the fibromyalgia survey questionnaire within a cross-sectional survey. *PLoS One*. 2012;7(5):e37504. <https://doi.org/10.1371/journal.pone.0037504>.
39. Basler H-D. Marburger Fragebogen zum habituellen Wohlbefinden. *Schmerz*. 1999;13(6):385–91. <https://doi.org/10.1007/s004820050216>.
40. Kazis LE, Selim A, Rogers W, Ren XS, Lee A, Miller DR. Dissemination of methods and results from the veterans health study: final comments and implications for future monitoring strategies within and outside the veterans healthcare system. *J Ambul Care Manag*. 2006;29(4):310–9. <https://doi.org/10.1097/00004479-200610000-00007>.
41. Ware JE, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220–33. <https://doi.org/10.1097/00005650-199603000-00003>.
42. Petzke F, Hüppe M, Kohlmann T, Kükenhöner S, Lindena G, Pflingsten M, et al. *Handbuch Deutscher Schmerz-Fragebogen*. 2020. https://www.schmerzgesellschaft.de/fileadmin/pdf/DSF_Handbuch_2020.pdf.
43. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. *Pain*. 1992;50(2):133–49. [https://doi.org/10.1016/0304-3959\(92\)90154-4](https://doi.org/10.1016/0304-3959(92)90154-4).
44. Frettlöh J, Maier C, Gockel H, Hüppe M. Validität des Mainzer Stadienmodells der Schmerzchronifizierung bei unterschiedlichen Schmerzdiagnosen. *Schmerz*. 2003;17(4):240–51. <https://doi.org/10.1007/s00482-003-0227-9>.
45. Schuler M, Schwarzmann G. Das Mainzer Stadienmodell der Schmerzchronifizierung ist auch bei stationären geriatrischen Patienten zur Graduierung chronischer Schmerzen geeignet. *Schmerz*. 2020;34(4):332–42. <https://doi.org/10.1007/s00482-020-00455-z>.
46. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res*. 1995;4(4):293–307. <https://doi.org/10.1007/BF01593882>.
47. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420–8. <https://doi.org/10.1037/0033-2909.86.2.420>.
48. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–63. <https://doi.org/10.1016/j.jcmm.2016.02.012>.
49. Cuesta-Vargas AI, Roldan-Jimenez C, Neblett R, Gatchel RJ. Cross-cultural adaptation and validity of the Spanish central sensitization inventory. *Springerplus*. 2016;5(1):1837. <https://doi.org/10.1186/s40064-016-3515-4>.
50. Browne MW, Cudeck R. Alternative ways of assessing model fit. *Sociol Methods Res*. 1992;21(2):230–58. <https://doi.org/10.1177/0049124192021002005>.
51. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res Online*. 2003;8:23–74.
52. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a “Core outcome set” – a practical guideline. *Trials*. 2016;17(1):449. <https://doi.org/10.1186/s13063-016-1555-2>.
53. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale: L. Erlbaum Associates; 1988.
54. Kendall MG. *Rank correlation methods*. London: Griffin; 1975.
55. R Core Team. *R: a language and environment for statistical computing*. Vienna: R Foundation for statistical Computing; 2018. <https://www.R-project.org/>
56. Rosseel Y. *Lavaan: an R package for structural equation modeling*. *J Stat Softw*. 2012;48:1–36.
57. Bühner M. *Exploratorische Faktorenanalyse*. In: *Einführung in die Test- und Fragebogenkonstruktion*. 3rd ed. München: Pearson; 2011. p. 295–378.
58. Chiarotto A, Viti C, Sulli A, Cutolo M, Testa M, Piscitelli D. Cross-cultural adaptation and validity of the Italian version of the central sensitization inventory. *Musculoskelet Sci Pract*. 2018;37:20–8. <https://doi.org/10.1016/j.msksp.2018.06.005>.
59. Sharma S, Jha J, Pathak A, Neblett R. Translation, cross-cultural adaptation, and measurement properties of the Nepali version of the central sensitization inventory (CSI). *BMC Neurol*. 2020;20(1):286. <https://doi.org/10.1186/s12883-020-01867-1>.
60. Bilika P, Neblett R, Georgoudis G, Dimitriadis Z, Fandridis E, Strimpakos N, et al. Cross-cultural adaptation and psychometric properties of the Greek version of the central sensitization inventory. *Pain Pract*. 2020;20(2):188–96. <https://doi.org/10.1111/papr.12843>.
61. Tanaka K, Nishigami T, Mibu A, Manfuku M, Yono S, Shinohara Y, et al. Validation of the Japanese version of the central sensitization inventory in patients with musculoskeletal disorders. *PLoS One*. 2017;12(12):e0188719. <https://doi.org/10.1371/journal.pone.0188719>.
62. Knezevic A, Neblett R, Jeremic-Knezevic M, Tomasevic-Todorovic S, Boskovic K, Colovic P, et al. Cross-cultural adaptation and psychometric validation of the Serbian version of the central sensitization inventory. *Pain Pract*. 2018; 18(4):463–72. <https://doi.org/10.1111/papr.12618>.
63. Caumo W, Antunes LC, Elkfury JL, Herbstrith EG, Busanello Sipmann R, Souza A, et al. The central sensitization inventory validated and adapted for a Brazilian population: psychometric properties and its relationship with brain-

- derived neurotrophic factor. *J Pain Res.* 2017;10:2109–22. <https://doi.org/10.2147/JPR.S131479>.
64. Knezevic A, Neblett R, Colovic P, Jeremic-Knezevic M, Bugarski-Ignjatovic V, Klasnja A, et al. Convergent and discriminant validity of the Serbian version of the central sensitization inventory. *Pain Pract.* 2020;20(7):724–36. <https://doi.org/10.1111/papr.12900>.
 65. Petzke F, Clauw DJ, Ambrose K, Khine A, Gracely RH. Increased pain sensitivity in fibromyalgia: effects of stimulus type and mode of presentation. *Pain.* 2003; 105(3):403–13. [https://doi.org/10.1016/S0304-3959\(03\)00204-5](https://doi.org/10.1016/S0304-3959(03)00204-5).
 66. Wolfe F, Walitt B, Rasker JJ, Häuser W. Primary and secondary fibromyalgia are the same: the universality of Polysymptomatic distress. *J Rheumatol.* 2018;46:204–12.
 67. Wolfe F, Michaud K, Busch RE, Katz RS, Rasker JJ, Shahouri SH, et al. Polysymptomatic distress in patients with rheumatoid arthritis: understanding disproportionate response and its Spectrum: disproportionate patient response in RA. *Arthritis Care Res (Hoboken).* 2014; 66(10):1465–71. <https://doi.org/10.1002/acr.22300>.
 68. Schäfer AGM, Joos LJ, Roggemann K, Waldvogel-Röcker K, Pfungsten M, Petzke F. Pain experiences of patients with musculoskeletal pain + central sensitization: a comparative group Delphi study. *PLoS One.* 2017;12(8): e0182207. <https://doi.org/10.1371/journal.pone.0182207>.
 69. Kregel J, Schumacher C, Dolphens M, Malfliet A, Goubert D, Lenoir D, et al. Convergent validity of the Dutch central sensitization inventory: associations with psychophysical pain measures, quality of life, disability, and pain cognitions in patients with chronic spinal pain. *Pain Pract.* 2018;18(6):777–87. <https://doi.org/10.1111/papr.12672>.
 70. Hendriks E, Voogt L, Lenoir D, Coppieters I, Ickmans K. Convergent validity of the central sensitization inventory in chronic whiplash-associated disorders; associations with quantitative sensory testing, pain intensity, fatigue, and psychosocial factors. *Pain Med.* 2020;21:2401–3412.
 71. Wijnhoven HAH, de Vet HCW, Picavet HSJ. Prevalence of musculoskeletal disorders is systematically higher in women than in men. *Clin J Pain.* 2006; 22(8):717–24. <https://doi.org/10.1097/01.aip.0000210912.95664.53>.
 72. Craner JR, Sperry JA, Evans MM. The relationship between pain Catastrophizing and outcomes of a 3-week comprehensive pain rehabilitation program. *Pain Med.* 2016;17(11):2026–35. <https://doi.org/10.1093/pm/pnw070>.
 73. Donath C, Geiß C, Schön C. Validation of a core patient-reported-outcome measure set for operationalizing success in multimodal pain therapy: useful for depicting long-term success? *BMC Health Serv Res.* 2018;18(1):117. <https://doi.org/10.1186/s12913-018-2911-6>.
 74. Neblett R, Hartzell MM, Williams M, Bevers KR, Mayer TG, Gatchel RJ. Use of the central sensitization inventory (CSI) as a treatment outcome measure for patients with chronic spinal pain disorder in a functional restoration program. *Spine J.* 2017;17(12):1819–29. <https://doi.org/10.1016/j.spinee.2017.06.008>.
 75. Arndt S, Turvey C, Andreasen NC. Correlating and predicting psychiatric symptom ratings: Spearman's r versus Kendall's tau correlation. *J Psychiatr Res.* 1999;33(2):97–104. [https://doi.org/10.1016/S0022-3956\(98\)90046-2](https://doi.org/10.1016/S0022-3956(98)90046-2).
 76. Piscitelli D, Pellicciari L. Responsiveness: is it time to move beyond ordinal scores and approach interval measurements? *Clin Rehabil.* 2018;32(10): 1426–7. <https://doi.org/10.1177/0269215518794069>.
 77. Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med.* 2012;44(2):97–8. <https://doi.org/10.2340/16501977-0938>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

