UCLA
UCLA Previously Published Works

Title
A Scoping Review of Chronic Low Back Pain Classification Schemes Based on Patient-Reported Outcomes.

Permalink
https://escholarship.org/uc/item/0891f15c

Journal
Pain physician, 25(6)

ISSN
1533-3159

Authors
Herman, Patricia M
Coulter, Ian D
Hays, Ron D
et al.

Publication Date
2022-09-01

Peer reviewed

Scoping Review

A Scoping Review of Chronic Low Back Pain Classification Schemes Based on Patient-Reported Outcomes

Patricia M. Herman, ND, PhD[1], Ian D. Coulter, PhD[1], Ron D. Hays, PhD[1,2], Anthony Rodriguez, PhD[3], and Maria O. Edelen, PhD[3,4]

From: [1]RAND Corporation, Santa Monica, CA; [2]UCLA Department of Medicine, Division of General Internal Medicine & Health Services Research, Los Angeles, CA; [3]RAND Corporation, Boston, MA; [4]Patient Reported Outcomes, Value and Experience (PROVE) Center, Department of Surgery, Brigham and Women's Hospital, Boston

Address Correspondence:
Patricia M. Herman, ND, PhD
RAND Corporation
1776 Main Street
PO Box 2138
Santa Monica, CA 90407-2138
E-mail: pherman@rand.org

Background: In 2014, the National Institutes of Health Pain Consortium Research Task Force recommended that patients with chronic low back pain (CLBP) be stratified by its impact on their lives. They proposed the Impact Stratification Score (ISS) to help guide therapy and facilitate study comparability. The ISS has been evaluated as a continuous measure, but not for use as a stratification or classification scheme.

Objectives: Identify the characteristics of successful schemes to inform the use of the ISS for stratification or classification.

Study Design: Scoping review of the peer-reviewed literature.

Methods: Search of PubMed, CINAHL, and APA PsycInfo to identify patient self-report-based classification schemes applicable to CLBP. Data were captured on the methods used for each scheme's development, the domains covered, their scoring criteria and what the classification has successfully measured. The study was reviewed and approved by the RAND Human Subjects Protection Committee (2019-0651-AM02).

Results: The search identified 87 published articles about the development and testing of 5 classification schemes: 1) The Subgroups for Targeted Treatment (STarT) Back Screening Tool, 2) Multiaxial Assessment of Pain, 3) Graded Chronic Pain Scale, 4) Back Pain Classification Scale, and 5) Chronic Pain Risk Score. All have been shown to be predictive of future outcomes and the STarT Back has been found useful in identifying effective classification-specific treatment. Each scheme had a different classification scoring structure, was developed using different methods, and 3 included domains not found in the ISS.

Limitations: Expanding the search to other databases may have identified more classification schemes. Our minimum number of publications inclusion criterion eliminated dozens of cluster analyses, some of which may have eventually been replicated.

Conclusions: The methods used to develop these successful classification schemes, especially those that use straightforward scoring schemes, should be considered for use in the development of a scheme based on the ISS.

Key words: Back pain, chronic pain, stratification, classification, grading, subgrouping, patient-reported outcome measures, Impact Stratification Score

Pain Physician 2022: 25:471-482

n 2014, the National Institutes of Health Pain Consortium Research Task Force (RTF) on research standards for chronic low back pain (CLBP) recommended that patients with CLBP be stratified by its impact on their lives (1). The RTF felt that stratification could have "descriptive and prognostic value and could

supplement any pathophysiologic description (1)," and improved "prognostic stratification of patients with CLBP is important clinically to help guide the nature and intensity of therapy, and important for researchers to adjust for confounding and to improve comparability among studies (1)."

The Institute of Medicine 2011 report, Relieving Pain in America, noted that "No simple clinical test can assess a person's subjective experience of pain. Seriousness depends on self-report . . . [of] pain's impact on a person's activities of daily living, ability to work, and quality of life (2)." The National Pain Strategy (NPS) went on to define high-impact chronic pain, in 2015, as that "associated with substantial restriction of participation in work, social, and self-care activities for six months or more (3)." The NPS further stated that in order to lower the burden of pain and better target effective interventions: "It is important to differentiate people with high-impact chronic pain from those who maintain normal activities although experiencing chronic pain (3)."

The RTF proposed the Impact Stratification Score (ISS) as a measure of CLBP impact. The ISS is calculated as the sum of the raw scores from 9 Patient-Reported Outcomes Measurement Information System-29 (PROMIS-29) items covering physical function, pain interference, and pain intensity with a possible range from a low of 8 (least impact) to 50 (greatest impact). The ISS has been evaluated as a continuous measure (1,4,5), but it has not yet been evaluated for stratification or classification. The RTF offered cutoff scores for classifying patients as having CLBP of mild (ISS 8-27), moderate (ISS 28-34), and severe impact (ISS > 35), but noted that these cutoffs were "relatively arbitrary (1)."

Useful classification schemes have been identified for many diseases—e.g., breast cancer (6), hip or knee osteoarthritis (7), heart failure (8), and chronic and musculoskeletal pain (4,9-12). These schemes use information from a variety of sources, including patient history, physical exam, lab tests, imaging, and patient-reported outcome measures (PROMs). Our focus in this paper is to review classification schemes that have been used for CLBP that, like the ISS, depend only on PROMs.

The general goal of all classification schemes is to segment large diverse patient populations (e.g., patients with CLBP) into relatively homogeneous subgroups. Homogeneity can be defined in at least 3 ways. The subgroups could be similar in their level of current severity and concomitant effects—e.g., subgroups with similar levels of chronic pain impact according to several measures have been shown to have similar health care costs, unemployment, and absenteeism (13-16). The subgroups could also be defined by having similar future outcomes or recovery (i.e., prognosis), regardless of treatment (10,17,18). Additionally, the subgroups could have similar response to specific treatments—i.e., vary by factors that are treatment modifiers (10,17,18). An implicit goal (the "Holy Grail") (17) of defining more homogeneous groups is to guide treatment. However, only the last definition of homogeneity (treatment modification) identifies the best treatments for each subgroup. The second (prognostic stratification) could also provide a more limited guide to treatment—e.g., by avoiding unnecessary treatment for those who were going to improve on their own. Prognostic stratification would also be useful for designing studies that minimize the heterogeneity of treatment effects. The first (severity) guides treatment only in the sense of identifying those most in need. There is no guarantee that the same classification scheme can generate subgroups with all 3 types of homogeneity (10,17).

As a first step in evaluating the ISS as a classification scheme for CLBP, we reviewed other PROM-based schemes to determine how they were developed, the domains they measure, the way their classification categories are determined, and whether they have been shown useful in creating categories with similar severity, prognosis, and/or benefit of particular treatment. This information will be used to inform future research on the use of the ISS as a classification system.

METHODS

This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews protocol (19) and the checklist is included as Appendix A. To identify existing PROM-based classification schemes for CLBP, we searched the abstracts of articles in Medline (Ovid), CINAHL, and APA PsycInfo from their inception through September 7, 2021. The full search for Medline is shown in Appendix B, but in general, we looked for articles whose abstracts included either back pain or chronic pain and variations on stratification, classification, categorization, grading, subgrouping, or clustering. We restricted the search to human studies published in English.

We chose for consideration studies that described, used, and/or evaluated classification schemes (i.e., methods by which patients are classified into mutually exclusive homogenous groups) that were used for adults with CLBP; required only information from

PROMs for classification; and were the topic of at least 3 publications—i.e., the classification scheme was of enough interest to warrant more than one other article. Schemes that required information from physical exam, lab tests, or imaging studies were excluded as were those with no more than 2 publications. Inclusion/exclusion criteria and whether individual studies met these criteria were collectively agreed upon by the first 3 authors. For each scheme, the first author extracted the number of items used, the domains included, the method(s) used to develop the scheme, the formula used to classify, the variables the scheme's results could discriminate, and the situations in which the scheme was shown to be useful for prognosis or as a guide for treatment.

RESULTS

The database search resulted in 7,550 articles to consider (Fig. 1) (20). After removal of duplicates and articles excluded based on reading the title and abstract, we reviewed 161 full-text articles and identified 87 articles describing 5 classification schemes that met our inclusion criteria. These are included in our narrative review.

The domains and scoring rules used in the ISS and the 5 classification schemes are shown in Table 1. Details

Fig. 1. *Flow of articles into the review.*
BPCS, Back Pain Classification Scale; CLBP, chronic low back pain; CPRS, Chronic Pain Risk Score; GCPS, Graded Chronic Pain Scale; MAP, Multiaxial Assessment of Pain; PROM, Patient Reported Outcome Measure; SBST, STarT Back Screening Tool.
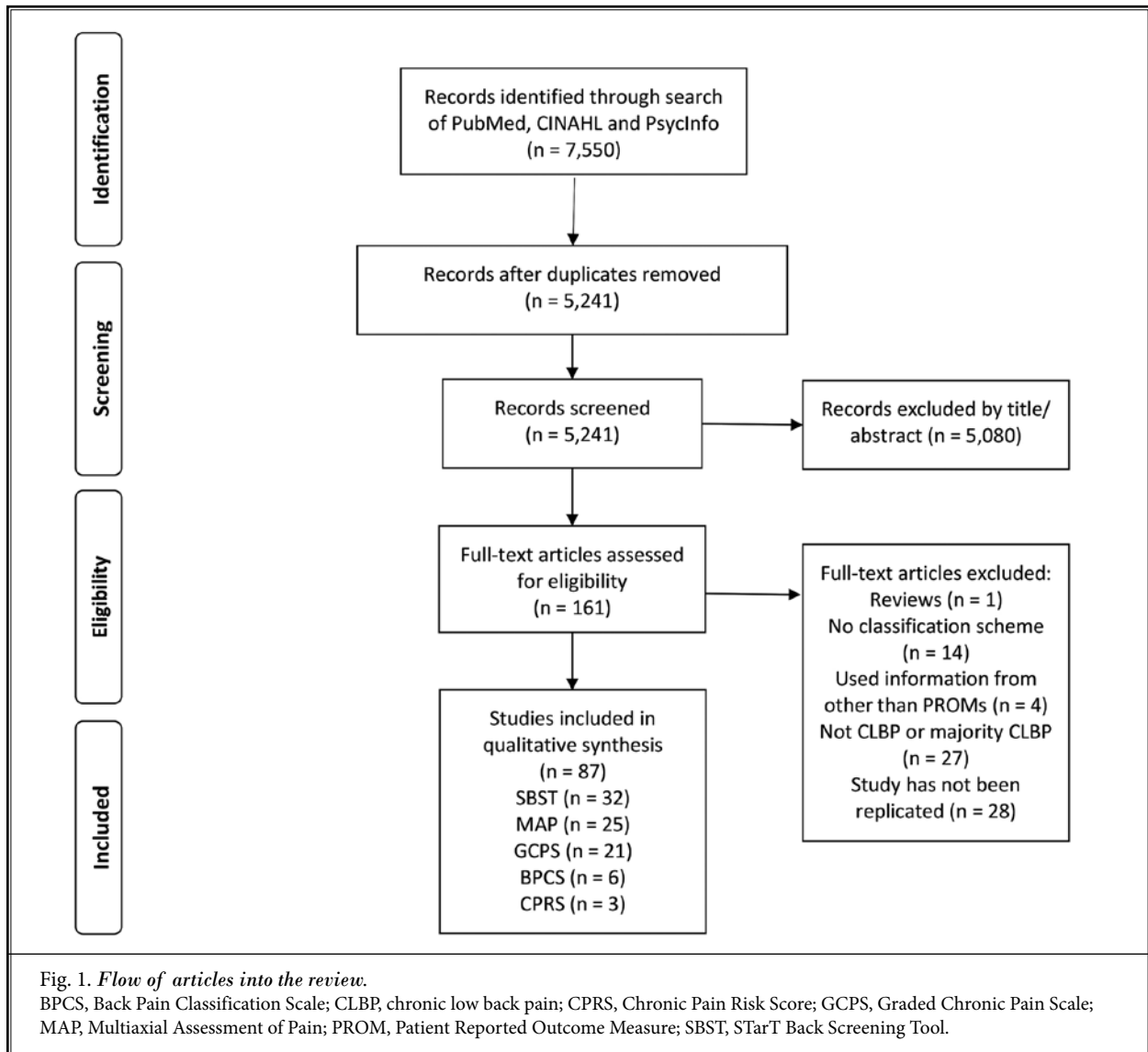
Table 1. Items, domains, scoring, and demonstrated usefulness of the Impact Stratification Score and 5 existing classification schemes.

	ISS	[Keele] SBST	MAP	GCPS	BPCS	CPRS
How Developed	Selected from PROMIS	Clinical advisory panel review of statistically promising items; ROC curves for cutoffs	Cluster analysis of responses to Multidimensional Pain Inventory	Mokken analysis to develop Guttman scale	Stepwise discriminant analysis	Latent transition regression analysis
Number of Items	9	9	52	7	13	8 items + the # used for depression scale
Pain Intensity	X	Bothersome	X	X	Pain descriptor words	X
Pain Interference	X		X	X		X
Physical Function	X	X		X		X
Pain Diffusion, Frequency & Duration		X				X
Beliefs About Pain		X	X			
Emotional Well-Being/Distress		X	X			X
Social Support			X			
Scoring	Sum scores 8-27 = mild 28-34 = moderate >= 35 = severe	Sum score < 3 = low risk; Sum score 4+ and subscore < 3 = moderate risk; subscore 4+ = high risk	Group means from the original study or a computer program to allocate individuals into groups	Disability points 3-4 = Grade III; Disability points 5-6 = Grade IV; Disability points 0-2 and pain intensity >50 = Grade II; Rest = Grade I	The sum of the weighted values given to each checked word is compared to mean scores for each classification	Low risk = 0-7; Intermediate risk = 8-15; Possible chronic pain = 16-21; Probable chronic pain = 22+
Severity	X	X	X	X	X	
Prognostic		X	X	X	X	X
Specify Treatment		X				

Abbreviations: ISS, Impact Stratification Score; SBST, STarT Back Screening Tool; MAP, Multiaxial Assessment of Pain; GCPS, Graded Chronic Pain Scale; BPCS, Back Pain Classification Scale; CPRS, Chronic Pain Risk Score; ROC, receiver operating characteristic.

on each scheme's background, development approach, scoring, ability to differentiate baseline characteristics, and use for predicting outcomes and guiding treatment are below.

The Subgroups for Targeted Treatment Back Screening Tool

The Subgroups for Targeted Treatment back screening tool (SBST) (21) was developed in the United Kingdom for use in primary care with adults experiencing the spectrum of nonspecific back pain. The goal was to develop and evaluate a brief and easy-to-score tool that used treatment-modifiable indicators to allocate primary care patients into 1 of 3 a priori initial treatment options based on their risk: 1) low-risk group suitable for primary care management (e.g., analgesia, advice, and education); 2) medium-risk group with high levels of physical indicators, appropriate for physiotherapy; and 3) high-risk group with consistently high levels of psychosocial indicators, appropriate for a combination of physical and cognitive-behavioral management.

The set of prognostic constructs included in the tool were identified through analysis of existing datasets using forward stepwise binary logistic regression to predict the reference standards listed below and a review of the literature. A clinical advisory panel reviewed the list of identified constructs, excluded those considered rare or non-

modifiable in primary care, and helped choose the final constructs based on strength, independence, consistency of association with outcomes, and perceived face validity. Items were selected for each construct based on receiver operating characteristic (ROC) curves to identify optimal items (i.e., items that identify patients above the median scores seen using the full questionnaires) from multi-item constructs and input from the expert panel. Items to include in a psychosocial subscale were also identified.

In addition to estimating internal consistency and test-retest reliability of the tool, discriminant validity was assessed using area under the curve from ROC curves for the overall tool scores and the score of the psychosocial subscale against the following dichotomized reference standards: back pain disability (Roland-Morris Disability Questionnaire [RMDQ] score [22] >= 7), whether there was referred leg pain, very or extremely bothersome back pain, catastrophizing (Pain Catastrophizing Scale score [23] >= 20), fear avoidance (Tampa Scale of Kinesiophobia [24] >= 41), and depression (Patient Health Questionnaire-2 [25] score >= 2).

To identify cutoff scores for each risk subgroup, ROC curves were examined. First, the optimal overall score (highest average sensitivity and specificity) that most consistently discriminated between reference standard cases and noncases in terms of patients pre-defined as being suitable for standard primary care management (low disability, no leg pain, low bothersomeness) was determined. Then, the psychosocial subscale score that best discriminated between the medium- and high-risk groups was identified using pre-defined psychosocial reference standards (catastrophizing, fear avoidance, depression). Emphasis was given to maximizing specificity for the psychosocial subscale because it was believed that physiotherapy could help lower distress and that there could be negative impacts from cognitive-behavioral approaches in those without distress. The validity of the tool in terms of predicting 6-month disability (RMDQ >= 7) was then evaluated using standard contingency table indices. The result was that individuals were classified as "low risk" if their total SBST score (out of 9 possible) was 3 or less, and "high risk" if their psychosocial subscale score (out of 5 possible) was 4 or above. The rest were considered "medium risk."

Although the SBST was developed to be useful for patients who present to primary care with all types of nonspecific LBP, and despite studies showing that it can be a better predictor of future outcomes in patients with longer pain duration (26), it has also been considered by some to be a screening tool to predict whether acute or subacute LBP would become chronic (27). Since our review specifically focused on classification schemes used for adults with CLBP, we included and focus here only on studies where the SBST has been used in CLBP (or majority CLBP) samples.

Across reviewed studies, the subgroups at baseline consisted of patients with different levels of pain intensity, activity limitations, disability (RMDQ), trunk motion, medication use, and a number of psychological measures (28-33). The SBST classification has also been shown to predict future Oswestry Disability Index (ODI) (34-36) and RMDQ (22,26,28,30,32,37-40) scores, pain intensity (28,35,36,41,42), fear of movement (35), work ability (43), preference-based health-related quality of life (EuroQol-5D [EQ-5D]) (32,36,44,45), 6-week ODI scores from an exercise program (46), and 2-year Graded Chronic Pain Scale grade and 12-item Short-Form Health Survey physical and mental health composite scores from a comprehensive health program (47).

Treatment assigned based on the SBST has also been shown to improve RMDQ scores (31,48,49) and preference-based health-related quality of life (EQ-5D) (45,49,50), reduce time off work (45,48), and is likely cost-effective (45,48-51). However, one trial of the SBST in a large health system where clinicians were trained on the tool and it was incorporated into the electronic health record system found no significant effect 2 or 6 months later on patients' back-related physical function, pain severity, or health care utilization (52). It also had limited effect on clinician behavior; clinicians used it to assess risk in only about half of their patients and the treatments they recommended did not change.

Multiaxial Assessment of Pain

The goal of the original Multiaxial Assessment of Pain (MAP) study was to see whether psychosocial and behavioral measures could be used to derive a reliable and valid classification system for patients with chronic pain (53). The authors conducted cluster analyses of the 9 scale scores of the West Haven-Yale Multidimensional Pain Inventory (MPI) (54) in 2 samples of patients with chronic pain referred to an outpatient pain clinic. Three profiles were identified from the cluster analysis: dysfunctional, interpersonally distressed, and adaptive copers. Relative to those in the other 2 groups, those in the dysfunctional profile reported higher pain severity and interference

and psychological distress, and lower general activity levels and ability to control their lives. Those in the interpersonally distressed profile were more likely to report that their families and significant others were not very supportive of them and their pain, and those in the adaptive copers profile reported lower levels of pain severity, interference and psychological distress, and higher levels of daily activity and control of their lives relative to the other 2 groups.

Two methods can be used to classify patients into these profiles. There is a computer program developed by Rudy (55) that assigns those who have completed the MPI to 1 of the 3 groups or to an "other" group. A second more ad hoc approach is to classify individuals based on how their MPI scores compare to the group means from the original study (53).

The stability of the 3-cluster solution was confirmed by using cluster analysis on 2 applications of the MPI in the same sample (56) and replication in samples with different subgroups of chronic pain patients (57,58), including those with different chronic pain syndromes (i.e., LBP, headache, and temporomandibular disorders) (59). The clusters were also evaluated by third-party reports (60). One study used the Comprehensive Pain Evaluation Questionnaire, a shorter measure modelled after the MPI, and generated clusters very similar to the 3 patient profiles identified using the MPI (61). Another study (63) proposed that a fourth cluster (defensive repressors) was needed (62) and increased the applicability of the MAP groups.

In reviewed studies (58,64-73), the clusters at baseline were able to discriminate between patients with different levels of pain intensity, disability, affective distress, anxiety, depression, pain behaviors, fear avoidance, endurance coping, catastrophizing, functional self-efficacy, personality types, psychopathology, and medication use.

MAP chronic pain profile status has been found to predict future sickness absence (74,75), cost of lost productivity (74), reductions in pain intensity and interference, and improvement of mental health and coping in response to various pain management programs (58,66), whether someone completes treatment (i.e., a functional restoration program) (64), absence from work, general health status, and use of health care resources following a vocational rehabilitation program (76), and outcomes from Interdisciplinary Multimodal Pain Rehabilitation Programs (77). However, one study (78) found that targeting specific treatments to each profile was not more effective than standard care.

Graded Chronic Pain Scale

The Graded Chronic Pain Scale (GCPS) was developed to offer a classification of chronic pain based on global (across-domain) measures of its severity (16). The authors used the Mokken analysis (Table 1) to test whether a set of pain-related items form a Guttman Scale. From previous work, they hypothesized that the lower range of pain severity would be measured by pain intensity and persistence and that the upper range would be measured by pain-related disability.

The authors found that 3 variables formed a Guttman Scale: pain intensity measured as the mean of present, and worst and average pain in past 6 months; number of days in past 6 months kept from usual activities because of pain; and disability measured as pain interference with daily activities, changes in ability to take part in recreational, social, and family activities, and changes in ability to work, including housework all in the past 6 months. Scoring of the GCPS yields 4 chronic pain grades. Grades I and II are defined as those with fewer than 3 disability points (determined by the number of disability days and the disability score, with a maximum of 6 points), and either pain intensity < 50 on a 0-100 scale (Grade I - low disability-low intensity) or pain intensity > 50 (Grade II - low disability-high intensity). Grades III and IV are defined by disability, regardless of pain intensity: 3 or 4 disability points define Grade III and 5 or 6 disability points define Grade IV.

At baseline, the grades were associated with significant and monotonic increases in the proportion of patients with depression, fair-poor self-rated health, frequent opioid use, frequent pain visits, unemployment and high-pain impact (defined as 8 or more "yes" answers to a list of 16 pain-related functional limitation items) (16). Baseline chronic pain grades were also significantly associated with pain duration and physical and psychosomatic comorbidity (79), health-related quality of life (single summary score of the 8-item Short-Form Health Survey [80]) and somatization (81), job change (82), and with back pain advice and misconceptions (83), days of sick leave, doctor visits, nights in hospital and unemployment (84).

Chronic pain grade at baseline has also been found to predict one-year pain grade, depression, fair-poor health status, frequent opioid use, frequent pain visits, high-pain impact, and unemployment (16); and pain grade and high-pain impact at 3 years (16). Baseline chronic pain grade predicted 6-month functional capacity, pain and the 36-item Short-Form Health Survey Physical Component Summary score (79); one-year

health care costs, number of visits and admissions, number of radiologic procedures and pain medication fills (13); and health care costs and future chronic pain grade at 2 years after a back exercise program (85).

A revised version of the GCPS was published in 2020 that categorizes those with chronic pain into mild (Grade 1), moderate (Grade 2), and high-impact (Grade 3) chronic pain (86). Based on work for the US National Pain Strategy (87) and by the National Center for Health Statistics cognitive library (88), those with high-impact chronic pain were identified based on responses of "most days" or "every day" to an item asking how often pain limited life or work activities. A summary score of 12 or greater on the Pain, Enjoyment, and General Activity Scale (89) identified those with moderate chronic pain and those with lower scores had mild pain. The 12 or greater cutoff was chosen to represent a mean of 4 or higher across the scale's three 0-10 items. A 4 on a 0-10 pain scale has been shown by others (90-92) to be the lower bound in identifying those with moderate pain. At baseline, the Revised GCPS grades were associated with coping beliefs, reported health status, depression/anxiety, activity limitations, and pain medication, including long-term opioid use (86).

Back Pain Classification Scale

The Back Pain Classification Scale (BPCS) was developed to provide an easy-to-administer indicator of whether a patient had functional (psychological) or organic (physiological) CLBP (93,95,96). The measure was developed using CLBP patients referred to neurosurgeons and orthopedic surgeons in the United Kingdom with probable intervertebral disc disease. These patients' clinical and laboratory findings were reviewed by board-certified surgeons and assigned to 1 of the 2 groups. Patients were shown a list of 71 pain descriptor words from the Low Back Pain Questionnaire (94) and asked to choose the words that best describe how their pain typically feels. The authors then used stepwise discriminant analysis to identify the best combination of pain words that would distinguish between the functional and organic groups. The resulting set of 13 pain words were able to correctly classify patients as organic or functional with an overall 94% accuracy. Applying the discriminant scores to a second validation sample resulted in an accuracy rate of 83% overall. Another study team (95) using a different sample found 80% accuracy for patients with chronic, intractable back pain.

Patients classified as functional at baseline were found to have a higher incidence of neurotic disorders than those classified as organic (96), and they were especially higher on the Minnesota Multiphasic Personality Inventory hypochondriasis scale. Classification at baseline using the BPCS was also associated significantly with medication need, patients' rating of improvement and change in pain over a 12-month period (97).

Chronic Pain Risk Score

The goal of the Chronic Pain Risk Score (CPRS) was to "discard the notion that 'chronic' means unlikely to change" and shift to predicting "the likelihood that clinically significant back pain will continue and, by extension, to [shift the focus to] steps that might be taken to reduce future risks of significant pain and dysfunction (98)." The CPRS was developed using a sample of patients in the United States with a history of primary care back pain visits.

Latent transition regression analysis (99) was used to empirically identify 4 pain severity classes (no pain, mild pain, moderate pain and limitation, and severe, limiting pain), and then estimate the probabilities of transitioning between these pain severity classes from one year to the next. Pain severity class was estimated using pain intensity, disability days, pain interference, pain impact score, unable to work for any health reason, and kept from full-time work due to back pain. The first 3 of these were elements of the GCPS. Transition probabilities were based on 3 prognostic variables (depression, pain duration, and diffuse pain), which were chosen because they "have been consistently found to have prognostic value in predicting pain outcomes in longitudinal outcome studies (98)." Clinically significant back pain was defined as having Chronic Pain Grade II, III, or IV—i.e., intense back pain accompanied by mild-to-severe dysfunction (16).

The resulting CPRS (possible range 0-28) is calculated as the sum of items from the GCPS (16) and prognostic variables. Days of activity limitation due to back pain from the GCPS was coded 0-4 and the pain intensity and pain interference items were each recoded from 0-10 to 0-2. The prognostic variables included the Symptom Check-List-90-R (SCL-90-R) depression score (recoded to 0-4), number of other pains (0-4), and the number of days with back pain in the prior 6 months (recoded to 0-4). The item scores were summed and the risk subgroups for the CPRS formed based on these cutoffs: > 22 = probable chronic pain (> 80% probability of future clinically significant back pain); 16-21 = possible chronic pain (> 50% probability of future clinically significant back pain); 8-15 = intermediate risk of chronic

pain (> 20% probability of future clinically significant back pain); and 0-7 = low risk of chronic pain (< 20% probability of future clinically significant back pain).

The chronic pain classification based on the CPRS at baseline was designed to determine risk of clinically significant back pain in the future (Chronic Pain Grade II, III, or IV) (98), but it has also been found to predict unemployment at 6 months and long-term opioid use (100).

One study (101) in the United Kingdom used the 14-item Hospital Anxiety and Depression Scale (HADS) instead of the SCL-90-R (90 items) as their measure of depression and found slightly different cutoff points to be optimal in patients with LBP (102). The probable chronic pain cutoff in the LBP population was the same, but the cutoff for possible chronic pain was increased and the cutoffs for intermediate or low-risk chronic pain were slightly reduced.

Another study (103) used a sample of US patients initiating primary care for back pain to examine whether the original CPRS could be improved by adding additional variables. Their results are not directly comparable because they used a more stringent definition of a negative outcome (Graded Chronic Pain Grade III or IV), but the success of their models indicates that adding other variables (e.g., college graduate, recovery expectations) may enhance prediction.

DISCUSSION

There are compelling reasons to classify CLBP patients into homogeneous subgroups. From a research perspective, this subgrouping could reduce patient heterogeneity and enhance trial efficiency, be used to report on the heterogeneity of treatment effect, and would allow adjustment for baseline sample differences for standardized outcome comparisons. For providers and patients, classification would contribute directly to both diagnosis and prognosis and could help guide treatment.

The authors' search identified 5 established PROM-based classification schemes for CLBP. All have been shown to be useful for predicting future outcomes for patients, but only one (SBST) has shown benefit in being used to guide class-specific treatment. It has been noted that it was designed to do this (104); whereas, one other (CPRS) was specifically designed for prognosis (98).

Each scheme was developed using different methods. The SBST identified constructs through the analysis of existing datasets and a review of the literature and then used a clinical advisory panel and ROC curves to identify the items to use and the cutoff scores for each risk subgroup (21). The MAP used cluster analysis on the 9 MPI scale scores to identify 3 profiles (21). These clusters proved remarkably replicable, especially in contrast to the dozens of other cluster analyses found that were never replicated more than once. Mokken analysis was used to test whether GCPS items covering pain intensity and pain-related disability formed a hierarchical severity scale (16). They then used inflection points on the relationships seen between pain intensity and disability to identify their cutoff points for grades. The BPCS used stepwise discriminant analysis to identify the best combination of pain descriptor words that would identify those with functional and organic CLBP (93). The CPRS used latent transition regression analysis to identify 4 pain severity classes and to estimate the probabilities of transitioning between these pain severity classes from one year to the next (99). In summary, a variety of analytic approaches have been used to develop and evaluate classification schemes. One or more of these approaches might be useful in evaluating the ISS. The common element is to identify meaningful subgroups that are associated with differences in CLBP impact and/or can predict future outcomes.

At present, no published studies have shown the ISS to be capable of prognostic stratification or treatment modification. All 5 identified classification schemes have been shown to be good for prognosis and most included domains not included in the ISS—e.g., at least 3 of the 5 schemes included measures of emotional well-being/distress. Two included measures of pain diffusion, frequency, and duration, and 2 included measures of pain beliefs. Therefore, it may be worth adding one or more of these domains to the ISS. It should be noted that to enhance prognosis the CPRS added measures of depression, duration of pain, and number of pain sites to the GCPS (98).

The schemes also vary widely in the number of items required, and thus, in patient burden. The GCPS has 7 items, the SBST has 9, the BPCS has 13, and the MAP has 52 items. The number of items in the CPRS depends on the instrument used to measure depression. Different studies used the SCL-90-R (98,100), HADS (102,105,106), or the Patient Health Questionnaire-8 (103) resulting in 98, 24, or 16 items, respectively. The ISS has 9 items.

Three of the five classification schemes used empirically derived cutoff scores to identify their subgroups. The MAP was developed using cluster analysis so that

subgroup classification requires calculation of the pattern of scores across subscales. Classification based on the BPCS, developed using stepwise discriminant analysis, requires the summation of weights applied to each selected pain descriptor words. These schemes are more difficult to apply clinically than those that apply upper- and/or lower-bound cutoffs to simple total scores. Of the 3 schemes using cutoff scores, one (CPRS) applies cutoffs to a total score to define subgroups, but the 2 others also apply cutoffs to subscores. The SBST applies one cutoff to the total score to identify the lowest risk subgroup, and then applies a different cutoff to the score of a subset of items to identify the moderate- and high-risk groups. The GCPS applies 2 cutoffs to a disability score—one that identifies Grade IV and one that separates Grade III from Grades I and II. Another cutoff is applied to pain intensity to differentiate between Grades I and II. Different scoring systems could be developed for the ISS in the future.

It should also be noted that refinement of the ISS-based classification scheme will require a different study design depending on the type of homogeneity desired (18). If one only needs subgroups to be homogeneous in terms of the current severity of CLBP, cross-sectional data will be sufficient. If, instead, we want the subgroups to be homogeneous in how their members change over time (i.e., prognosis; where the members of some subgroups do better no matter the treatment), then longitudinal data and analysis will be needed. However, if it is important to identify the best treatment for each subgroup (i.e., whether subgroup membership modifies the effect of a treatment), prospective randomized controlled trials are required. These designs can either take the shape of having those in each subgroup randomized to treatment or control or randomizing all patients to receive either usual care or treatment matched to their subgroup. The SBST (49,50) and the MAP (78) were both tested using this last design.

The authors identified PROM-based classification schemes through a detailed review of the literature. However, it is not without limitations. The search only included PubMed, CINAHL, and APA PsycInfo. It is possible that the inclusion of other databases may have identified more classification schemes. We did not include classification schemes that had not been described, utilized, and/or evaluated in at least 3 published studies. This exclusion criterion mainly eliminated cluster analyses on various instruments and groups of instruments. Finally, since we were only seeking to identify previously developed classification schemes, we did not perform a critical appraisal of the quality or validity of the identified studies.

Conclusions

The ISS is made up of 9 items from the PROMIS-29 and was proposed by the RTF to stratify CLBP patients by the level of impact their condition has on their lives. The goals of this stratification were prognosis, to help guide therapy, and to aid researchers in identifying more homogeneous samples for trials and in comparing across studies. Nevertheless, to date, the ISS has only been evaluated as a continuous measure. The authors present a review of the literature that identified 5 classification schemes that have been developed for CLBP and that have achieved one or more of these goals. The results of this search identified the methods used to develop these classification schemes and differences between the ISS and these schemes, which can inform the use of the ISS for classification. Methods that result in classification according to empirically derived cutoff scores are favored for clinical ease of application. Like the Mokken scale analysis used in the development of the GCPS, item response theory may be useful in identifying levels of the ISS that represent clinically important differences in pain impact. Regression analyses, including latent transition analysis, may be useful in evaluating how well ISS subgroups predict future outcomes. It also may be worthwhile to supplement the ISS items to include domains included in the other schemes. Further work is needed to achieve the goals originally put forth for impact stratification using the ISS.

Appedices available at www.painphysicianjournal.com

REFERENCES

1. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on research standards for chronic low back pain. *Pain Med* 2014; 15:1249-1267.

2. Institute of Medicine. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*, The National Academies Press, Washington, DC 2011.

3. National Pain Strategy Task Force. *National Pain Strategy: A Comprehensive Population Health-Level Strategy for Pain*. Interagency Pain Research Coordinating Committee (IPRCC), National Institutes of Health (NIH), Bethesda, MD, 2015, pp 1-83.

4. Deyo RA, Ramsey K, Buckley DI, et al. Performance of a patient reported outcomes measurement information system (PROMIS) short form in older adults with chronic musculoskeletal pain. *Pain Medicine* 2016; 17:314-324.

5. Dutmer AL, Reneman MF, Preuper HRS, et al. The NIH minimal dataset for chronic low back pain: Responsiveness and minimal clinically important change. *Spine* 2019; 44:E1211.

6. Viale G. The current state of breast cancer classification. *Annals of Oncology* 2012; 23:x207-x210.

7. Altman RD. The classification of osteoarthritis. *The Journal of Rheumatology Supplement* 1995; 43:42-43.

8. Brahmbhatt DH, Cowie MR. Heart failure: Classification and pathophysiology. *Medicine* 2018; 46:587-593.

9. Riddle DL. Classification and low back pain: A review of the literature and critical analysis of selected systems. *Physical Therapy* 1998; 78:708-737.

10. Fairbank J, Gwilym SE, France JC, et al. The role of classification of chronic low back pain. *Spine* 2011; 36:S19-S42.

11. Ford JJ, Hahne AJ. Pathoanatomy and classification of low back disorders. *Manual Therapy* 2013; 18:165-168.

12. McCarthy CJ, Arnall FA, Strimpakos N, et al. The biopsychosocial classification of non-specific low back pain: A systematic review. *Physical Therapy Reviews* 2004; 9:17-30.

13. Engel CC, von Korff M, Katon WJ. Back pain in primary care: Predictors of high health-care costs. *Pain* 1996; 65:197-204.

14. Herman PM, Broten N, Lavelle TA, et al. Exploring the prevalence and construct validity of high-impact chronic pain across chronic low-back pain study samples. *The Spine Journal* 2019; 19:1369-1377.

15. Herman PM, Broten N, Lavelle TA, et al. Health care costs and opioid use associated with high-impact chronic spinal pain in the United States. *Spine* 2019; 44:1154-1161.

16. Von Korff M, Ormel J, Keefe FJ, et al. Grading the severity of chronic pain. *Pain* 1992; 50:133-149.

17. Foster NE, Hill JC, Hay EM. Subgrouping patients with low back pain in primary care: Are we getting any better at it? *Manual Therapy* 2011; 16:3-8.

18. Kamper SJ, Maher CG, Hancock MJ, et al. Treatment-Based subgroups of low back pain: A guide to appraisal of research studies and a summary of current evidence. *Best Practice & Research Clinical Rheumatology* 2010; 24:181-191.

19. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann Intern Med* 2018; 169:467-473.

20. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine* 2009; 6:e1000097.

21. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: Identifying patient subgroups for initial treatment. *Arthritis and Rheumatism* 2008; 59:632-641.

22. Roland M, Morris R. A study of the natural history of back pain: Part I: Development of a reliable and sensitive measure of disability in low-back pain. *Spine* 1983; 8: 141-144.

23. Sullivan MJ, Bishop SR, Pivik J. The Pain Catastrophizing Scale: Development and validation. *Psychol Assess* 1995; 7:524.

24. Kori S. Kinisophobia: A new view of chronic pain behavior. *Pain Manage* 1990: 35-43.

25. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Med Care* 2003:1284-1292.

26. Morso L, Kongsted A, Hestbaek L, et al. The prognostic ability of the STarT Back Tool was affected by episode duration. *European Spine Journal* 2016; 25:936-944.

27. Pauli J, Starkweather A, Robins JL. Screening tools to predict the development of cronic low back pain: An integrative review of the literature. *Pain Medicine (Malden, Mass)* 2019; 20:1651-1677.

28. Kendell M, Beales D, O'Sullivan P, et al. The predictive ability of the STarT Back Tool was limited in people with chronic low back pain: A prospective cohort study. *Journal of Physiotherapy* 2018; 64:107-113.

29. Robinson HS, Dagfinrud H. Reliability and screening ability of the StarT Back Screening Tool in patients with low back pain in physiotherapy practice, a cohort study. *BMC Musculoskeletal Disorders* 2017; 18:232.

30. Medeiros FC, Salomão EC, Costa LOP, et al. Use of the STarT Back Screening Tool in patients with chronic low back pain receiving physical therapy interventions. *Brazilian Journal of Physical Therapy* 2021; 25:286-295.

31. Murphy SE, Blake C, Power CK, et al. Comparison of a stratified group intervention (STarT Back) with usual group care in patients with low back pain: A nonrandomized controlled trial. *Spine* 2016; 41:645-652.

32. Matsudaira K, Oka H, Kikuchi N, et al. The Japanese version of the STarT Back Tool predicts 6-month clinical outcomes of low back pain. *Journal of Orthopaedic Science* 2017; 22:224-229.

33. Davoudi M, Shokouhyan SM, Abedi M, et al. A practical sensor-based methodology for the quantitative assessment and classification of chronic non specific low back patients (NSLBP) in clinical settings. *Sensors (Basel, Switzerland)* 2020; 20:2902.

34. Fairbank J, Couper J, Davies J, et al. The Oswestry Low Back Pain Disability Questionnaire. *Physiotherapy* 1980; 66:271-273.

35. Page I, Abboud J, O Shaughnessy J, et al. Chronic low back pain clinical outcomes present higher associations with the STarT Back Screening Tool than with physiologic measures: A 12-month cohort study. *BMC Musculoskeletal Disorders* 2015; 16:201.

36. Unsgaard-Tondel M, Vasseljen O, Nilsen TIL, et al. Prognostic ability of STarT Back Screening Tool combined with work-related factors in patients with low back pain in primary care: A prospective study. *BMJ Open* 2021; 11:e046446.

37. Kneeman J, Battalio SL, Korpak A, et al. Predicting persistent disabling low back pain in veterans affairs primary care using the STarT Back Tool. *PM & R* 2021; 13:241-249.

38. Morso L, Kent P, Manniche C, et al. The predictive ability of the STarT Back Screening Tool in a Danish secondary care setting. *European Spine Journal* 2014; 23:120-128.

39. Suri P, Delaney K, Rundell SD, et al. Predictive validity of the STarT Back Tool for risk of persistent disabling back pain in a U.S. primary care setting. *Archives of Physical Medicine and Rehabilitation* 2018; 99:1533-1539.e1532.

40. Bier JD, Ostelo RWJG, van Hooff ML, et al. Validity and reproducibility of the STarT Back Tool (Dutch version) in patients with low back pain in primary care settings. *Physical Therapy* 2017; 97:561-570.

41. Fritz JM, Beneciuk JM, George SZ. Relationship between categorization with the STarT Back Screening Tool and prognosis for people receiving physical therapy for low back pain. *Physical Therapy* 2011; 91:722-732.

42. Medeiros FC, Costa LOP, Added MAN, et al. Longitudinal monitoring of patients with chronic low back pain during physical therapy treatment using the STarT Back Screening Tool. *J Orthop Sports Phys Ther* 2017; 47:314-323.

43. Forsbrand MH, Grahn B, Hill JC, et al. Can the STarT Back Tool predict health-related quality of life and work ability after an acute/subacute episode with back or neck pain? A psychometric validation study in primary care. *BMJ Open* 2018; 8:e021748.

44. Brooks R, Group E. EuroQol: The current state of play. *Health Policy* 1996; 37:53-72.

45. Whitehurst DGT, Bryan S, Lewis M, et al. Implementing stratified primary care management for low back pain: Cost-Utility analysis alongside a prospective, population-based, sequential comparison study. *Spine* 2015; 40:405-414.

46. Majeed AS, Ts A, Sugunan A, et al. The effectiveness of a simplified core stabilization program (TRICCS-trivandrum community-based core stabilisation) for community-based intervention in chronic non-specific low back pain. *Journal of Orthopaedic Surgery and Research* 2019; 14:86.

47. Huppe A, Zeuner C, Karstens S, et al. Feasibility and long-term efficacy of a proactive health program in the treatment of chronic back pain: A randomized controlled trial. *BMC Health Services Research* 2019; 19:714.

48. Foster NE, Mullis R, Hill JC, et al. Effect of stratified care for low back pain in family practice (IMPaCT Back): A prospective population-based sequential comparison. *Annals of Family Medicine* 2014; 12:102-111.

49. Hill JC, Whitehurst DGT, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): A randomised controlled trial. *Lancet (London, England)* 2011; 378:1560-1571.

50. Whitehurst DGT, Bryan S, Lewis M, et al. Exploring the cost-utility of stratified primary care management for low back pain compared with current best practice within risk-defined subgroups. *Annals of the Rheumatic Diseases* 2012; 71:1796-1802.

51. Hall JA, Jowett S, Lewis M, et al. The STarT Back stratified care model for nonspecific low back pain: A model-based evaluation of long-term cost-effectiveness. *Pain* 2021; 162:702-710.

52. Cherkin D, Balderson B, Wellman R, et al. Effect of low back pain risk-stratification strategy on patient outcomes and care processes: The match randomized trial in primary care. *Journal of General Internal Medicine* 2018; 33:1324-1336.

53. Turk DC, Rudy TE. Toward an empirically derived taxonomy of chronic pain patients: Integration of psychological assessment data. *Journal of Consulting and Clinical Psychology* 1988; 56:233-238.

54. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain inventory (WHYMPI). *Pain* 1985; 23:345-356.

55. Rudy TE (ed). *Multiaxial Assessment of Pain: Multidimensional Pain Inventory Computer Program User's Manual,* Version 2.0, University of Pittsburgh, Pittsburgh, PA 1989.

56. Verra ML, Angst F, Staal JB, et al. Reliability of the multidimensional pain Inventory and stability of the MPI classification system in chronic back pain. *BMC Musculoskeletal Disorders* 2012; 13:155.

57. Bergström G, Bodin L, Jensen IB, et al. Long-Term, non-specific spinal pain: Reliable and valid subgroups of patients. *Behaviour Research and Therapy* 2001; 39:75-87.

58. Verra ML, Angst F, Staal JB, et al. Differences in pain, function and coping in Multidimensional Pain Inventory subgroups of chronic back pain: A one-group pretest-posttest study. *BMC Musculoskeletal Disorders* 2011; 12:145.

59. Turk DC, Rudy TE. The robustness of an empirically derived taxonomy of chronic pain patients. *Pain* 1990; 43:27-35.

60. Junghaenel DU, Broderick JE. Validation of the MPI patient profiles by partners and healthcare providers. *Pain* 2009; 144:130-138.

61. Nelli JM, Nicholson K, Fatima Lakha S, et al. Use of a modified Comprehensive Pain Evaluation Questionnaire (CPEQ): Characteristics and functional status of patients on entry to a tertiary care pain clinic. *Pain Research & Management* 2012; 17:75-82.

62. Burns JW, Kubilus A, Bruehl S, et al. A fourth empirically derived cluster of chronic pain patients based on the Multidimensional Pain Inventory: Evidence for repression within the dysfunctional group. *Journal of Consulting and Clinical Psychology* 2001; 69:663-673.

63. Hopwood CJ, Creech SK, Clark TS, et al. Optimal scoring of the Multidimensional Pain Inventory in a chronic pain sample. *Journal of Clinical Psychology in Medical Settings* 2008; 15:301-307.

64. Choi Y, Mayer TG, Williams M, et al. The clinical utility of the Multidimensional Pain Inventory (MPI) in characterizing chronic disabling occupational musculoskeletal disorders. *Journal of Occupational Rehabilitation* 2013; 23:239-247.

65. Geisser ME, Robinson ME, Henson CD. The Coping Strategies Questionnaire and chronic pain adjustment: A conceptual and empirical reanalysis. *The Clinical Journal of Pain* 1994; 10:98-106.

66. McMorrow E, Walsh R, O'Keeffe D, et al. Differential outcomes to a pain management programme based on coping style. *The Irish Journal of Psychology* 2010; 31:11-26.

67. Asmundson GJG, Norton RG, Allerdings MD. Fear and avoidance in dysfunctional chronic back pain patients. *Pain* 1997; 69:231-236.

68. Johansson E, Lindberg P. Low back pain patients in primary care: Subgroups based on the Multidimensional Pain Inventory. *International Journal of Behavioral Medicine* 2000; 7:340-352.

69. Rusu AC, Hasenbring M. Multidimensional Pain Inventory derived classifications of chronic pain: Evidence for maladaptive pain-related coping within the dysfunctional group. *Pain* 2008; 134:80-90.

70. Etscheidt MA, Steger HG, Braverman B.

Multidimensional Pain Inventory profile classifications and psychopathology. *Journal of Clinical Psychology* 1995; 51:29-36.

71. Nitch SR, Boone KB. Normal personality correlates of chronic pain subgroups. *Journal of Clinical Psychology in Medical Settings* 2004; 11:203-209.

72. Lousberg R, Greonman N, Schmidt A. Profile characteristics of the MPI-DLV clusters of pain patients. *Journal of Clinical Psychology* 1996; 52:161-167.

73. Carmody TP. Psychosocial subgroups, coping, and chronic low-back pain. *Journal of Clinical Psychology in Medical Settings* 2001; 8:137-148.

74. Bergstrom G, Bergstrom C, Hagberg J, et al. A 7-year follow-up of multidisciplinary rehabilitation among chronic neck and back pain patients. Is sick leave outcome dependent on psychologically derived patient groups? *European Journal of Pain (London, England)* 2010; 14:426-433.

75. Bergstrom C, Hagberg J, Bodin L, et al. Using a psychosocial subgroup assignment to predict sickness absence in a working population with neck and back pain. *BMC Musculoskeletal Disorders* 2011; 12:81.

76. Bergstrom G, Jensen IB, Bodin L, et al. The impact of psychologically different patient groups on outcome after a vocational rehabilitation program for long-term spinal pain patients. *Pain* 2001; 93:229-237.

77. Gerdle B, Cervin M, Rivano Fischer M, et al. Outcomes of interdisciplinary pain rehabilitation across subgroups of the Multidimensional Pain Inventory – a study from the Swedish Quality Registry for Pain Rehabilitation. *Pain Practice* 2021; 21:662-679.

78. Verra ML, Angst F, Brioschi R, et al. Effectiveness of subgroup-specific pain rehabilitation: A randomized controlled trial in patients with chronic back pain. *European Journal of Physical and Rehabilitation Medicine* 2018; 54:358-370.

79. Buchner M, Neubauer E, Zahlten-Hinguranage A, et al. The influence of the grade of chronicity on the outcome of multidisciplinary therapy for chronic low back pain. *Spine* 2007; 32:3060-3066.

80. Ware J, Kosinski M, Dewey J, et al. A manual for users of the SF-8 health survey. Lincoln, RI: *Quality Metric Incorporated* 2001: 4-19.

81. Arnow BA, Hunkeler EM, Blasey CM, et al. Comorbid depression, chronic pain, and disability in primary care. *Psychosomatic Medicine* 2006; 68:262-268.

82. Walker BF, Muller R, Grant WD. Low back pain in Australian adults: Prevalence and associated disability. *Journal of Manipulative and Physiological Therapeutics* 2004; 27: 238-244.

83. Goubert L, Crombez G, De Bourdeaudhuij I. Low back pain, disability and back pain myths in a community sample: Prevalence and interrelationships. *European Journal of Pain (London, England)* 2004; 8:385-394.

84. Hauser W, Wolfe F, Henningsen P, et al. Untying chronic pain: Prevalence and societal burden of chronic pain stages in the general population - a cross-sectional survey. *BMC Public Health* 2014; 14:352.

85. Muller G, Pfinder M, Clement M, et al. Therapeutic and economic effects of multimodal back exercise: A controlled multicentre study. *Journal of Rehabilitation Medicine* 2019; 51:61-70.

86. Von Korff M, DeBar LL, Krebs EE, et al. Graded chronic pain scale revised: Mild, bothersome, and high-impact chronic pain. *Pain* 2020; 161:651-661.

87. Von Korff M, Scher AI, Helmick C, et al. United States National Pain Strategy for Population Research: Concepts, definitions, and pilot data. *J Pain* 2016; 17:1068-1080.

88. Dahlhamer J, Lucas J, Zelaya C, et al. Prevalence of chronic pain and high-impact chronic pain among adults—United States, 2016. *Morb Mortal Weekly Rep* 2018; 67:1001-1006.

89. Krebs EE, Lorenz KA, Bair MJ, et al. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. *J Gen Intern Med* 2009; 24:733-738.

90. Boonstra AM, Preuper HRS, Balk GA, et al. Cut-Off points for mild, moderate, and severe pain on the visual analogue scale for pain in patients with chronic musculoskeletal pain. *Pain®* 2014; 155:2545-2550.

91. Gerbershagen HJ, Rothaug J, Kalkman C, et al. 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.

92. Hirschfeld G, Zernikow B. Variability of "optimal" cut points for mild, moderate, and severe pain: Neglected problems when comparing groups. *PAIN®* 2013; 154:154-159.

93. Leavitt F, Garron DC. The detection of psychological disturbance in patients with low back pain. *Journal of Psychosomatic Research* 1979; 23:149-154.

94. Leavitt F, Garron DC, Whisler WW, et al. Affective and sensory dimensions of back pain. *Pain* 1977; 4:273-281.

95. Sanders SH. Cross-Validation of the Back Pain Classification Scale with chronic, intractable pain patients. *Pain* 1985; 22:271-277.

96. Leavitt F, Garron DC. Validity of a Back Pain Classification Scale for detecting psychological disturbance as measured by the MMPI. *Journal of Clinical Psychology* 1980; 36: 186-189.

97. McNeill TW, Sinkora G, Leavitt F. Psychologic classification of low-back pain patients: A prognostic tool. *Spine* 1986; 11:955-959.

98. Von Korff M, Miglioretti DL. A prognostic approach to defining chronic pain. *Pain* 2005; 117:304-313.

99. Miglioretti DL. Latent transition regression for mixed outcomes. *Biometrics* 2003; 59:710-720.

100. Von Korff M, Dunn KM. Chronic pain reconsidered. *Pain* 2008; 138:267-276.

101. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica* 1983; 67:361-370.

102. Dunn KM, Croft PR, Main CJ, et al. A prognostic approach to defining chronic pain: Replication in a UK primary care low back pain population. *Pain* 2008; 135:48-54.

103. Turner JA, Shortreed SM, Saunders KW, et al. Optimizing prediction of back pain outcomes. *PAIN®* 2013; 154:1391-1401.

104. Lheureux A, Berquin A. Comparison between the STarT Back Screening Tool and the Örebro Musculoskeletal Pain Screening Questionnaire: Which tool for what purpose? A semi-systematic review. *Annals of Physical and Rehabilitation Medicine* 2019; 62:178-188.

105. Muller S, Thomas E, Dunn KM, et al. A prognostic approach to defining chronic pain across a range of musculoskeletal pain sites. *The Clinical Journal of Pain* 2013; 29:411.

106. Thomas E, Dunn KM, Mallen C, et al. A prognostic approach to defining chronic pain: Application to knee pain in older adults. *Pain* 2008; 139:389-397.

Appendix A. *Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews Checklist*

Section	Item	PRISMA-ScR Checklist Item	Reported on Page #
TITLE			
Title	1	Identify the report as a scoping review.	1
ABSTRACT			
Structured Summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	4-6
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or patients, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	6
METHODS			
Protocol and Registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	N/A
Eligibility Criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	6-7
Information Sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	7
Search	8	Present the full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Appendix A
Selection of Sources of Evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	7
Data Charting Process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data Items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	7
Critical Appraisal of Individual Sources of Evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	23
Synthesis of Results	13	Describe the methods of handling and summarizing the data that were charted.	7
RESULTS			
Selection of Sources of Evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	8 and Fig. 1
Characteristics of Sources of Evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	8-19 and Table 1
Critical Appraisal Within Sources of Evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	N/A
Results of Individual Sources of Evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	8-19 and Table 1
Synthesis of Results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	19-22
DISCUSSION			
Summary of Evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	19-22

Appendix A (cont.). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews Checklist*

Section	Item	PRISMA-ScR Checklist Item	Reported on Page #
Limitations	20	Discuss the limitations of the scoping review process.	22-23
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	23-24
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	1

Abbreviations: JBI, Joanna Briggs Institute; PRISMA-ScR, Preferred Reporting Items for Systematic Reviews and Meta-analyses Extension for Scoping Reviews.
* Where sources of evidence (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.
† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with information sources (see first footnote).
‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4,5) refer to the process of data extraction in a scoping review as data charting.
§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).
From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and explanation. Ann Intern Med 2018; 169:467-473.

Appendix B. *Medline Search*

OVID Medline

September 7, 2021
English, human, abstract
1. (clusters or clustering).ab.
2. (categoris* or categoriz*).ab.
3. (classif* not international classification).ab.
4. (stratif* or grading or taxonomy or graded or "chronic pain grade*").ab.
5. subgrouping.ab.
6. 1 or 2 or 3 or 4 or 5
7. exp Back Pain/ or exp Chronic Pain/ or (exp *Pain Measurement/ and "back pain".ab.)
8. 6 and 7
9. limit 8 to (abstracts and english language and humans)