

Report of the Task Force on Research Standards for Chronic Low-Back Pain

Submitted to the
NIH Pain Consortium Executive Committee
November 18, 2013

Authors: Richard A. Deyo, M.D., M.P.H.; Samuel F. Dworkin, D.D.S., Ph.D.; Dagmar Amtmann, Ph.D.; Gunnar Andersson, M.D., Ph.D.; David Borenstein, M.D.; Eugene Carragee, M.D.; John Carrino, M.D., M.P.H.; Roger Chou, M.D.; Karon Cook, Ph.D.; Anthony DeLitto, P.T., Ph.D.; Christine Goertz, D.C., Ph.D.; Partap Khalsa, D.C., Ph.D.; John Loeser, M.D.; Sean Mackey, M.D., Ph.D.; James Panagis, M.D.; James Rainville, M.D.; Tor Tosteson, Sc.D.; Dennis Turk, Ph.D.; Michael Von Korff, Sc.D.; Debra Weiner, M.D.

Affiliations: Oregon Health and Sciences University; the University of Washington; Rush University Medical Center; George Washington University; Stanford University; Johns Hopkins University; Northwestern University; University of Pittsburgh; Palmer College of Chiropractic; National Center for Complementary and Alternative Medicine; National Institute for Arthritis, Musculoskeletal and Skin Diseases; New England Baptist Hospital; Geisel School of Medicine at Dartmouth; and Group Health Research Institute.

The statements, conclusions, and recommendations contained in this document reflect opinions of the meeting participants and are not necessarily intended to represent the official position of any Federal agency, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Table of Contents

Executive Summary	2
Introduction	4
Methods	5
Results	10
Discussion	16
References	17
Appendices	24
Appendix 1. Research Task Force (RTF) Meeting Agendas	25
1.1 Agenda for March 5–6, 2012	26
1.2 Agenda for October 1–2, 2012	29
1.3 Agenda for March 11–12, 2013	32
Appendix 2. Pre-RTF Meeting Background and Information Distribution	35
2.1 NIH Pain Consortium Charge to the RTF	36
2.2 Advancing Research on cLBP: Developing Standards for Research	38
2.3 Some Questions to Consider for Meeting 1	40
2.4 Rationale for Developing a Biopsychosocial, Dual-Axis RDC/cLPBS Diagnostic and Classification System at the Present Time	41
2.5 First Meeting Overview and Charge to Workgroups	43
2.6 Chronic Low-Back Pain: Classification; Research Needs and Recommendations	47
2.7 Overview of Minimal Dataset for Research on cLBP	56
2.8 Preamble and Recommendations of the NIH Task Force on Research Standards for Chronic Low-Back Pain	62
Appendix 3. Surveying the RTF: Survey Data on Preferences and Actions	64
3.1 Prioritizing Minimal Dataset Items: History, Demographics, Physical Findings	65
3.2 Prioritizing Minimal Dataset Items: Psychological, Behavioral, and Psychosocial	70
3.3 Feasibility of RTF Developing an RDC/cLBP Diagnostic and Classification System	72
Appendix 4. Recommended Multidimensional Minimal Dataset for Research on cLBP	73

EXECUTIVE SUMMARY

Background: Chronic low-back pain (cLBP) is common and has a major societal impact. Despite rapidly increasing use of medications, injections, and surgery, functional disability has increased in recent decades. Many patients who have procedures to correct putative causes continue to have pain. Further, we often cannot identify mechanisms to explain the major negative impact cLBP has on the lives of many patients. Such cLBP is often termed nonspecific, idiopathic, or mechanical, and may in fact be due to varied and multiple biologic and behavioral etiologies.

In 2009 and 2010, the National Institutes of Health (NIH) Pain Consortium convened two workshops on low-back pain research, noting that researchers use varied inclusion criteria, definitions, baseline assessments, and outcome measures. This impedes comparing studies, replicating findings, pooling data, resolving conflicts, and achieving consensus. It was recommended that NIH establish research standards on cLBP. The NIH Pain Consortium subsequently charged a Research Task Force (RTF) to:

- Consider the state of research relevant to standards for clinical research on cLBP
- Review definitions, diagnostic criteria, and outcome measures for clinical research
- Develop a draft set of standards for research on cLBP
- Engage the research community and government agencies in developing research standards
- Chart a plan for incorporating standards into research studies and making future revisions.

Approach: Co-chairs with complementary expertise were selected, along with 14 additional members who had varied scientific and clinical expertise. The RTF evolved a three-stage work plan, each with a 2-day meeting and intervening literature review. Between meetings, the co-chairs surveyed members by e-mail regarding key elements. These principles emerged:

- The process should be evidence-based and use a biopsychosocial model of chronic pain.
- Data should be useful for patients with degenerative disorders (e.g., herniated disc, lumbar stenosis) as well as those without clear pathoanatomy.
- Patients with underlying systemic or specific diseases were not the target of the Task Force.
- Patients with no clear pathoanatomy should not be assumed to have “psychogenic” pain.
- Classifying cLBP by impact is more feasible and potentially useful than classifying solely by pathophysiology. “Impact” includes pain intensity, interference, and physical function.
- A brief minimal uniform dataset should be reported in all studies of chronic back pain.
- The dataset should be relevant for population, observational, and interventional research.
- An investigator could substitute more detailed and precise measures for a particular domain but should report data for each domain of the minimal dataset.
- Research standards should evolve; we propose a potential research agenda for refinement.

Results: The RTF made six recommendations regarding standards for research on cLBP:

1. Definition of cLBP: The RTF recommended two questions to define chronic pain: (1) How long has back pain has been a problem? (2) What fraction of days in the past 6 months involved back pain? A patient with pain on at least half the days in the past 6 months would have accumulated at least 3 month’s worth of pain days, and this was the recommended definition.

2. Classification of cLBP by Impact: “Impact” was defined by pain intensity, pain interference with normal activities, and functional status. These items have major prognostic and discriminatory importance. Impact is calculated from 9 items of the 29-item Patient Reported Outcomes Measurement Information System (PROMIS) short form.

Using PROMIS data from patients with cLBP, the RTF Impact Classification showed strong correlations with legacy functional measures and was associated with patient satisfaction. Impact scores improved over time, as expected. Effect sizes and standardized response means suggested the Impact Classification was more responsive than the Roland Disability Index.

3. Minimal Dataset: Medical history and examination included demographics, involvement in workers’ compensation, work status, education, comorbidity, and previous treatment. For some of these, we adopted the Common Data Elements implemented by the National Institute of Neurological Disorders and Stroke. Physical examination items were reserved for studies of invasive interventions or of older adults. No laboratory or imaging tests were highly ranked by the RTF because of their weak associations with patient symptoms or function. However, magnetic resonance imaging was recommended for studies of surgical interventions. Key self-report domains (in addition to pain and pain-related interference) were physical function, depression, sleep disturbance, and catastrophizing. The short form PROMIS measures were thought to offer the best tradeoff of length with psychometric validity.

4. Outcome Measures: Many parts of the minimal dataset, such as PROMIS measures, are also appropriate as outcome measures. However, primary outcomes of clinical studies will vary, depending on study aims. Thus, the RTF did not recommend a minimal outcome dataset. However, the RTF recommended reporting a “responder” analysis in addition to reporting mean scores of outcome measures. This amounts to determining the “cumulative distribution function” of responders, reported as the percentage of responders at each cutoff value of the outcome score for treatment and control groups.

5. Recommendations for Research on the Proposed Standards: The RTF recommended new research to improve prognostic stratification of patients with cLBP; refine and test composite outcome measures for increasing the clinical importance of study results; undertake patient stakeholder assessment of relevant outcomes; and further evaluate psychometric properties of the minimal dataset.

6. Dissemination: Upon adoption of recommendations by the NIH Pain Consortium, the RTF recommends dissemination to the broad research community. This would include publication of a report in multiple professional journals and presentations at professional meetings.

Conclusion: The RTF believes these recommendations will advance the field, help to resolve controversies, and facilitate future research addressing the genomic, neurologic, and other mechanistic substrates of cLBP. We expect the RTF recommendations will become a dynamic document and undergo continual improvement.

INTRODUCTION

The Institute of Medicine (2011) has identified chronic pain as a U.S. societal problem of enormous impact. It affects about 100 million adults and has an estimated annual cost of \$635 billion, including for direct medical expenditures and loss of work productivity.¹ Low-back pain (LBP) that limits daily activity has a worldwide lifetime prevalence of about 39 percent and a similar annual prevalence of 38 percent.² It occurs from adolescence through the elderly. Most people having LBP experience recurrent episodes.³ The use of all interventions—including surgery, pharmacologic, and nonpharmacologic approaches—for treatment of chronic LBP (cLBP), sometimes referred to as cLBP syndrome (cLBPS), increased from 1995 to 2010; despite this, the prevalence of symptoms and expenditures also continued to increase.⁴⁻⁶

LBP is a symptom. There is now growing evidence, however, that in its chronic form (cLBP) it can progress, like other chronic pain conditions,⁷ beyond a symptomatic state to a complex condition unto itself. This can include persistent anatomical and functional changes in the central nervous system,⁸⁻¹⁰ in addition to changes in the back (e.g., degenerative spinal changes and atrophy and/or asymmetry of paraspinal muscles).¹¹⁻¹³ Although some patients with cLBP have clear pathoanatomic etiologies, for most there is no clear association between their pain and an identifiable pathology of the spine and its associated soft tissues (i.e., intervertebral discs, ligaments, joint capsules, and muscles).¹⁴ Furthermore, it is often not possible to identify mechanisms to account for the appreciable negative impact cLBP has on the lives of many sufferers.¹⁵ Such pain is often termed nonspecific, idiopathic, mechanical, or due to instability, and it may in fact be due to different and multiple biologic and behavioral etiologies in different individuals.¹⁶

A range of different classes of interventions have been developed and tested in adults with cLBP. These include spine surgery, injections into all the structures of the back, pharmacologic regimens approved by the U.S. Food and Drug Administration (FDA), psychological interventions (e.g., cognitive-behavioral treatment), manual therapies (e.g., spinal manipulation/mobilization, massage), exercise, nutritional supplements (e.g., glucosamine, herbs) and lifestyle-adjustment and self-management approaches.¹⁷⁻²⁰ Many of these interventions have shown some clinical benefit, but few appear to consistently provide substantial, long-term reductions in pain with increased function.¹⁷⁻²⁰

A critical issue for advancing research on cLBP is the challenge of comparing results from the many classes of interventions. In 2009 and 2010, the National Institutes of Health (NIH) Pain Consortium convened two workshops on LBP research that invited experts from the relevant scientific and clinical fields to provide NIH with research recommendations. These experts noted that, often, prior clinical studies have used variable criteria for determining whom to include and exclude, varying case definitions for LBP and its chronicity or recurrence, and inconsistent baseline assessments, stratification criteria, and outcome measures. As a result, it has been difficult to compare studies of similar or competing interventions, replicate findings, pool data from multiple studies, resolve conflicting conclusions, develop multidisciplinary consensus, or even achieve consensus within a single discipline regarding interpretation of findings. Key recommendations from these workshops on how to advance cLBP research were to establish research standards on cLBP and to have NIH facilitate and enable this process.

In response, the NIH Pain Consortium established a subcommittee, the Steering Committee for a Research Task Force on Standards of Research for cLBP. The Committee was comprised of representatives from the following NIH institutes/centers: NCCAM, NIA, NIAMS, NICHD, NIDA, NIDCR, NINDS, and NINR. The Steering Committee developed the goals for the Research Task Force (RTF), identified what scientific and clinical expertise would be needed, selected two co-chairs, and added 14 invited experts from outside NIH. The Steering Committee provided two representatives (James Panagis, M.D., M.P.H., and Partap Khalsa, D.C., Ph.D.) in *ex-officio* (i.e., non-voting) capacity to the RTF. Finally, NIH organized and funded (as costs for meeting space, hotel rooms, travel, and per diems) three face-to-face meetings of the RTF, held in March 2012, October 2012, and March 2013, in the greater Washington, DC, area.

The charge by the NIH Pain Consortium to the RTF was to develop a set of standards for clinical research on cLBP that would address the following:

- Consider the state of existing research relevant to the development of standards for clinical research on cLBP.
- Conduct a comprehensive review of existing case definitions, diagnostic criteria, and outcome measures that are relevant for clinical research on cLBP.
- Develop a draft set of standards for research on cLBP.
- Engage the broader research community and representatives from relevant government agencies in developing these standards for research on cLBP.
- Chart a general plan for their incorporation into research studies and their future modification.

This charge focuses completely on developing standards to be used in research—not in coding, billing, or other purposes in clinical settings. Appendix 2.1 provides more details on the charge, and Appendix 2.2 an overview of the goals, scope, and methods of the project as envisioned at its inception.

METHODS

Creating the RTF. After identifying the scientific/clinical expertise that would be needed on the RTF, the Steering Committee selected two co-chairs who would have complementary leadership expertise. Richard Deyo, M.D., M.P.H., was chosen because of his expertise in LBP research, and Samuel Dworkin, D.D.S., Ph.D., because of his prior leadership in developing Research Diagnostic Criteria (RDC) for another chronic pain condition, temporomandibular disorders. The co-chairs, in consultation with the Steering Committee, reached consensus on selecting the RTF members (listed in Table 1 below) to achieve the needed scientific/clinical expertise.

Table 1: Task Force Members, Affiliations, and Expertise		
Members	Affiliations	Expertise
Co-chairs		
Richard A. Deyo, M.D., M.P.H.	Departments of Family Medicine, Internal Medicine, and Public Health, Oregon Health and Science University and Clinical Investigator, Kaiser Center for Health Research	Primary care, health services research
Samuel F. Dworkin, D.D.S., Ph.D.	Departments of Oral Medicine and Psychiatry and Behavioral Sciences, University of Washington	Temporomandibular joint disorders, chronic pain, clinical psychology
Task Force Members		
Gunnar Andersson, M.D., Ph.D.	Department of Orthopaedic Surgery, Rush University Medical Center	Orthopaedic spine surgery
David Borenstein, M.D.	Department of Rheumatology, George Washington University	Rheumatology
Eugene Carragee, M.D.	Department of Orthopaedic Surgery, Stanford University School of Medicine	Orthopaedic spine surgery
John Carrino, M.D., M.P.H.	Department of Radiology, Johns Hopkins University School of Medicine	Musculoskeletal radiology
Roger Chou, M.D.	Departments of Medicine, and of Medical Informatics and Clinical Epidemiology, Oregon Health and Science University	General internal medicine, systematic review
Anthony DeLitto, P.T., Ph.D.	Department of Physical Therapy, University of Pittsburgh School of Health Rehabilitation	Physical therapy
Christine Goertz, D.C., Ph.D.	Palmer College of Chiropractic	Chiropractic care, epidemiology
John Loeser, M.D.	Department of Neurological Surgery, University of Washington	Neurosurgery, pain management
Sean Mackey, M.D., Ph.D.	Department of Anesthesia, Stanford University School of Medicine	Pain management, functional brain imaging
James Rainville, M.D.	Physical Medicine and Rehabilitation, New England Baptist Hospital and Tufts University	Spine rehabilitation
Tor Tosteson, Sc.D.	Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth	Biostatistics
Dennis Turk, Ph.D.	Department of Anesthesiology and Pain Medicine, University of Washington	Pain medicine, psychology
Michael Von Korff, Sc.D.	Group Health Research Institute, Seattle	Epidemiology
Debra Weiner, M.D.	Departments of Medicine, Psychiatry, and Anesthesiology, and the Clinical and Translational Science Institute, University of Pittsburgh, and the Geriatric Research, Education and Clinical Center, VA Pittsburgh Healthcare System	Rheumatology, geriatrics

Work Plan: The RTF evolved a three-stage work plan, with each stage including a 2-day meeting that would yield a summary of recommendations. Appendix 1 provides the agendas for all meetings, and Appendix 2 the background information supplied to members prior to meetings.

Stage 1. The first meeting took place on March 5–6, 2012. Pre-meeting documents were provided (see Appendices 2.1–2.5). Meeting 1 opened with remarks by the NIAMS and NCCAM Directors, Stephen Katz, M.D., Ph.D., and Josephine Briggs, M.D., respectively. Among their comments, they emphasized the nature of chronic back pain as a highly prevalent and costly public health challenge. They also noted the existence of many stakeholders in this problem, including individuals with back pain; health care systems; clinicians; drug and device makers; Federal, state, and regulatory agencies; and third-party payers. Issues such as lost productivity, litigation, and compensation for back pain were also key. Important research needs described included standardized methods to facilitate reporting chronic back pain research, comparing reports, and replicating results; case definitions; and standardized baseline and outcome measures. Existing diagnostic and classification systems were noted as potential models for cLBP research, and the research focus of the Task Force was emphasized.

This first meeting used small group discussion and plenary sessions to clarify and identify where the RTF could be most effective in its recommendations. The members sought consensus on the following issues and strategies:

- Potential benefit of a biopsychosocial model of chronic pain to drive research recommendations
- Emphasis on interdisciplinary research designs as a high priority
- Requirement for measurable variables and criteria, with operational definitions
- Need for reliable, valid, and clinically-useful research measures and criteria
- Suitability of recommended approaches for population, observational, and clinical research
- Potential for developing a multiaxial set of research diagnostic criteria (RDC) for cLBP, based on existing models for other chronic pain-related conditions (see Appendix 2.4 for a document on the rationale).

Three workgroups focused respectively on physical findings and objective signs of pathology; behavioral, psychological, and psychosocial function; and measures of prognosis. All contributed to the following topics:

- Developing a research definition for cLBP
- Identifying predictors of chronicity and poor outcomes
- Identifying minimal datasets to characterize study subjects at baseline, using a dual axis model (physical and psychosocial). This would be analogous to the model utilized in the RDC for temporomandibular disease.²¹

The RTF noted that intended users of the proposed research standards would be investigators submitting grant applications to NIH, but that the standards would be available and encouraged for all researchers. The research standards could potentially allow cLBP phenotypes to be uncovered, based on physical and psychosocial findings.

Stage 1 Summary. The RTF decided that it could not respond in detail to every component of the NIH Pain Consortium's charge. For example, producing explicit evidence-based diagnostic criteria for

conditions such as spinal stenosis, sciatica, or spine “instability” would be impossible, given the current available resources and lack of professional consensus.

Thus, members recommended that the co-chairs draft documents for the second meeting with these objectives:

1. Achieve consensus on a research-oriented definition of cLBP.
2. Pursue evidence-based measures and criteria to classify subjects by the personal impact of cLBP rather than by pathophysiology or pathoanatomy.
3. Pursue the feasibility of developing reliable and valid predictors of chronicity.
4. Reach consensus on a minimal dataset that could be recommended for NIH-supported cLBP research and that could be easily tabulated in published articles.
5. Reevaluate the feasibility of a pathology-based diagnostic system for subsets of cLBP.
6. Develop a strategy for NIH-funded cLBP research (see Appendix 2).
7. Plan a meeting for dissemination of recommendations and feedback from stakeholders in industry and in the governmental, professional, research, and advocacy communities.

Stage 2. The agenda for Meeting 2 is attached as Appendix 1.2. In preparation, the co-chairs prepared and distributed a series of surveys to RTF members on key issues from Meeting 1 and then provided the results (see Appendix 3). Those surveys and a review of existing literature included the following:

- *Axis I Survey: Candidate Measures for a Minimal Axis I Dataset (Objective and Medical History).* Members were asked to rank the importance of potential baseline descriptors for patients with cLBP. These included items of medical history, comorbidity, physical examination, and laboratory and imaging tests. Patients with clear-cut herniated disc and sciatica, lumbar stenosis, ankylosing spondylitis, and underlying systemic disease (such as cancer or infection) were presumed to be excluded.
- *Axis II Survey: Candidate Measures for a Minimal Dataset for Rating of Axis II (Self-Report of Pain, Interference, Function, Psychological, and Psychosocial Measures).* Task Force members were asked to rank the importance of measures of pain-related behavioral, emotional, and psychosocial domains influencing the expression of cLBP. This would accompany the Axis I minimal dataset.
- *Survey on Feasibility of Developing Research Diagnostic Criteria for Subsets of Nonspecific Chronic Low-Back Pain.* Findings from several past NIH Pain Consortium workshops included the observation that inability to categorize subtypes of back pain patients hinders back pain research and care, and, with respect to studies, interpretation, comparison, and replication. Part of the charge from the Pain Consortium was to consider developing a Research Diagnostic Classification system (i.e., criteria for subsets of nonspecific cLBP). This survey asked Task Force members to assess the feasibility of such an effort (see Appendix 3).

- *Review of Existing Literature on Back Pain Classification and Prognosis.*
The Task Force did not undertake a systematic literature review. However, it considered previous work on back pain taxonomy,²²⁻³⁷ prognostic classification,³⁸⁻⁷² pain and psychosocial measures,⁷³⁻⁹⁶ and outcome assessment.⁹⁷⁻¹¹¹ This literature informed its deliberations and recommendations.

Meeting 2 aimed at reaching consensus regarding the seven objectives listed in the *Stage 1 Summary* above. The RTF also heard presentations of two related NIH efforts. The first was NINDS' effort to create Common Data Elements to be used by all researchers that it supports. The second was the Patient Reported Outcomes Measurement Information System (PROMIS), which includes several measures directly relevant to the RTF.

Stage 2 Summary. Major accomplishments in Stage 2 were as follows:

1. **Defined cLBP, its impact, and the predictors of chronicity.** Finalizing these issues would require more data, but there was good agreement among members that it is important to assess pain impact as a means of categorizing cLBP.
2. **Defined elements for the minimal dataset.** Judging by the survey responses, candidate items for the Axis I and Axis II components of the minimal dataset were well received. Special attention was directed to the possible use of PROMIS measures.
3. **Refined the scope for research recommendations/research strategies for NIH-supported cLBP research.** The RTF decided to narrow its research recommendations to those needed to further evaluate and refine the evidence basis for the proposed definitions and minimal dataset.
4. **Reached consensus on developing RDC/cLBP.** There was consensus that developing pathophysiologic diagnostic criteria for subsets of nonspecific LBP was unfeasible at present.

The Task Force had discussed holding a future meeting for stakeholders on the recommendations to be developed. It decided to postpone this meeting until the actual recommendations are available.

Stage 3. An important issue that had emerged from Meeting 2 was the need for more data measuring the impact and prognosis of cLBP. To this end, five RTF members worked together to refine the working definitions of cLBP and its impact, and evaluate candidate items for predictors of LBP chronicity. The resulting document, "Classifying Chronic Low-Back Pain; Definition, Assessment of Impact, Rationale, and Measures," is attached as Appendix 2.6.

The co-chairs revised the minimal dataset to address concerns about overall length, measures for special research purposes, and outcome measures. This produced a refined and more practical version of the minimal dataset's pain descriptions, and emotional and psychosocial components. The co-chairs then compiled this material into a comprehensive document, which was distributed to RTF members as part of a package including the agenda (see Appendices 1.3, 2.6, and 2.7) prior to Meeting 3.

The objective of Meeting 3 was to finalize the RTF's recommendations. The meeting reached consensus on a definition of cLBP; key principles (see Table 2 on page 8); the final content of the minimal dataset (see Appendix 4, "Recommended Multidimensional Minimal Dataset for Research on cLBP"); and strategies for obtaining feedback and support for its recommendations through

presentations. These presentations would take place at meetings of research and professional organizations and consultation with:

- NIH Pain Consortium and relevant NIH institutes and centers
- Governmental agencies, including the FDA, CMS, CDC, etc.
- Editors of journals important to NIH pain researchers.

Stage 3 Summary. The final recommendations in “Results” below are a first step towards creating Standards for Research in cLBP. We anticipate that these Standards, if adopted, will receive ongoing evaluation and review, which will yield successively improved iterations.

RESULTS

The RTF developed and utilized a set of key principles (see Table 2) to guide formulation of its major recommendations.

Table 2: Key Principles Utilized by the Chronic Low-Back Pain Research Task Force

1. Use an evidence-based approach that incorporates a biopsychosocial model of chronic pain.
2. Data should be useful for a wide range of conditions, including patients thought to have degenerative spinal disorders (e.g., herniated disc or lumbar stenosis) as well as those without identified pathoanatomy.
3. Patients with underlying systemic or specific diseases are not the target of the Task Force, including cancer, spinal infections, fractures, and inflammatory spondylopathies such as ankylosing spondylitis.
4. Patients with no identified pathoanatomy should not be assumed to have “psychological,” “psychogenic,” or “somatoform” pain.
5. Given the current state of knowledge, classifying cLBP by its impact is more feasible and potentially useful than attempting classification solely by pathoanatomy or physiology. Impact of cLBP will tentatively be defined in terms of pain intensity, interference with activities, and physical function.
6. A minimal uniform dataset should be reported in all studies of chronic back pain. It should be brief, so that investigators can supplement it with key measures for specific research questions.
7. The dataset should include both biomedical and psychosocial variables.
8. An investigator could substitute more detailed, precise, and well-validated measures for a particular domain, but should report data for each domain of the minimal dataset.
9. Additional “core” items would be recommended for specific study aims or populations, such as surgical trials or elderly populations.
10. A prognostic dimension for the classification of chronic low back pain would be desirable, but more evidence is needed before an explicit recommendation will be made.
11. Research standards should evolve, and the RTF’s recommendations will suggest a potential research agenda for refining the research standards.

Definition of cLBP. The RTF considered various potential elements for this definition, including time with pain, days with pain, severity of pain, varying durations of pain (e.g., 3 months or 6 months), and the problem of intermittent, recurrent symptoms.

Elements of Final Definition. The RTF concluded that two questions should be used to define cLBP:

1. How long has your back pain been a problem?
2. What fraction of days in the past 6 months involved back pain?

Pain severity was not made part of the definition. A patient with pain on at least half the days in the past 6 months would have accumulated at least 3 months' worth of pain days, and the RTF concluded that this would be the recommended definition. A drawing to illustrate the low back would be included in a patient questionnaire, indicating the space between the lower posterior margin of the rib cage and the horizontal gluteal fold. Appendix 4, "Recommended Multidimensional Minimal Dataset for Research on cLBP," incorporates the precise wording recommended for this definition of cLBP.

Classification of cLBP by Impact. The RTF recommended classifying cLBP by its personal impact, rather than by anatomic or physiologic factors. It approved a proposal that this classification of impact should consist of a combination of three domains from PROMIS:

1. Pain intensity
2. Pain interference with normal activities
3. Functional status.

These items have substantial research support to readily validate their discriminatory and prognostic importance.

After considerable discussion about formal prognostic scales for subclassification, such as the Keele STarT Back Screening Tool, the RTF decided there remained substantial uncertainty about generalizability. Thus, the RTF recommends further research in this area, and included several items of the STarT Back instrument in the minimal dataset, but chose not to require it for classification purposes.

This classification of cLBP by impact would be appropriate whether or not there appears to be contributory degenerative pathoanatomy. Even when pathoanatomic conditions are thought to contribute to symptoms and dysfunction, they often coexist and overlap, and sometimes fail to respond to specific interventions. Thus, the impact classification seems to be a useful addition to pathoanatomic or physiologic classification.

The impact can be calculated from 9 items of the 29-item PROMIS short form, and this is referred to as the RTF Impact Classification. Using the raw PROMIS scores, the usual scoring of the Physical Function items is reversed. Therefore, for each item in the Impact Classification, a score of 1 is least severe and 5 most severe, except for the single item on pain intensity, which ranges from 0–10. Thus, scores on the Impact Classification range from 8 (least impact) to 50 (greatest impact). Appendix 4, "Recommended Multidimensional Minimal Dataset for Research on cLBP," incorporates the precise wording recommended for the definition of cLBP and includes the specific items used to assess pain impact.

Testing of the Proposed Impact Classification Score. Because the proposed impact score is a novel combination of three constructs (pain intensity, interference, and function), the RTF undertook a preliminary assessment of its validity and performance. This project used existing PROMIS data from 218 patients with low-back pain, with or without leg pain, who underwent epidural steroid injections. Given this intervention, an improvement in average functional scores was expected. That dataset included legacy measures of back pain-related physical function, including the Oswestry Disability Index and the Roland-Morris Disability Questionnaire (RMDQ).

Table 3 demonstrates the association of the RTF Impact Classification with the legacy measures, the distribution of RTF Impact Classification scores, and measures of responsiveness over time. Among the major findings:

- The RTF Impact Classification showed strong correlations with legacy measures.
- Score changes on the RTF Impact Classification correlated more strongly with patient satisfaction at follow-up than did changes on the RMDQ.
- In this rather severely affected sample, baseline RTF Impact scores were almost equally distributed among mild, moderate, and severe impacts (Table 3).
- Scores on the RTF measure improved over time, as expected.
- The RTF recommends that investigators simply report actual scores, along with any categorization(s) of their choice. In this test, although the cutoffs used for mild, moderate, and severe scores are intuitive and potentially useful, they are also relatively arbitrary.
- Measures of effect size and standardized response mean suggested that the RTF Impact Classification was more responsive than the RMDQ (Table 3).

Construct Validation: Correlation of RTF Impact Classification with Legacy Measures of Physical Function, Baseline (Spearman R)	Oswestry Disability Index	Roland-Morris Disability Index
RTF Impact Classification score	.806	.661
Construct Validation: Correlation of Score Changes with Patient Satisfaction with Treatment at Followup (Spearman R, absolute value)	Change, Roland-Morris Disability Index	Change, RTF Impact Classification score
Patient satisfaction index, scored 1-4	.148	.251
Distribution of RTF Impact Classification Scores	Baseline (N=218), % of subjects	Followup (N=170), % of subjects
RTF Impact Classification score 8-27 (mild)	30%	63%
RTF Impact Classification score 38-34 (moderate)	34%	18%
RTF Impact Classification score 35 or greater (severe)	36%	19%
Mean RTF Impact Classification score (SD)	32 (8.3)	25 (9.7)
Responsiveness	Effect Size (Change/Baseline SD)	Standardized Response Mean (Change/SD of change)
RTF Impact Classification	0.69	0.75
Roland-Morris Disability scale	0.39	0.41

Proposed Minimal Dataset—Medical History, Physical Examination, Diagnostic Testing. In the survey of RTF members regarding items for a minimal dataset (see Appendix 3), the most highly ranked items of medical history and examination included demographics, involvement in workers' compensation or legal claims, work status, education, various measures of comorbidity, and previous treatment history. For many of these measures, the RTF adopted the format of the NINDS Common Data Elements system.

The key comorbid conditions were judged to be smoking status, obesity, widespread pain symptoms, and substance abuse. The Two-item Conjoint Screen (TICS) was judged to be an adequate and suitably brief screen for substance abuse. The key items of treatment history were thought to be history of surgical interventions and use of opioid analgesics.

Measures from the physical examination generally ranked lower than the items of medical history. However, the most highly ranked of the physical examination measures were straight leg raising for patients with leg pain; hip internal rotation as a measure of hip arthritis (a potential cause of LBP); and lower extremity strength. There was general agreement that such physical examination items could be reserved for studies of invasive interventions (straight leg raising and lower extremity strength) or of older adults (hip examination). Thus, physical examination measures would not be required of all epidemiological studies, for example.

No laboratory or imaging tests were highly ranked, because of the widely recognized weak association between degenerative spine changes on imaging and patient symptoms or function.¹⁴ However, magnetic resonance imaging (MRI) was considered the most valuable of potential tests, and there was agreement that it should be required in studies of surgical interventions.

Proposed Minimal Dataset—Self-report of Functional Status, Psychosocial Factors, and Mood Disturbance. With regard to other self-report measures, there was discussion first of the domains to be included, potential sources of items, and desirable number of items. The key domains that the RTF selected for the proposed minimal dataset were physical function, depression, sleep disturbance, and catastrophizing. This was based on survey of and discussions by members (see Appendix 3 for completed survey data). The RTF found these constructs the most important for a wide range of patients with chronic back pain, with or without specific pathoanatomic diagnoses. Other constructs—such as anxiety, fatigue, and satisfaction with social role—were also considered important, but for reasons of parsimony, were not included in the dataset.

After considering a wide range of potential instruments for assessing these domains, the RTF concluded that the 29-item, short-form PROMIS measures¹¹³ offered the best tradeoff between length and psychometric validity. Therefore, it recommended use of that form's relevant scales, which include four items for each domain. Computer adaptive testing (CAT), using the entire PROMIS item bank, would offer an acceptable or even preferable alternative for investigators who have access to those tools.

There was agreement that if investigators preferred well-validated, lengthier legacy measures of these domains—such as the Oswestry or RMDQ for physical function, or the Patient Health Questionnaire (PHQ-9) or Beck Depression Inventory for depression—then their use would be acceptable. But if such substitutions are made, all other recommended domains should still be assessed.

The RTF suggested that investigators may find it useful to consult the PROsetta Stone Web site, a project supported by NCI-funded investigators at Northwestern University.¹¹⁴ This Web site provides a

“crosswalk” between scores on the PROMIS measures and scores on several legacy measures, such as the Brief Pain Inventory, the Center for Epidemiologic Studies Depression Scale (CES-D Scale), the PHQ-9, and the Short Form Health Survey (SF-36).

The resulting proposed minimal dataset is presented in Appendix 4, “Recommended Multidimensional Minimal Dataset for Research on cLBP,” in a format suitable for clinical research use by chronic low-back pain and non-pain control subjects.

Proposed Supplemental Data for Specific Situations. For studies of invasive therapies such as spine surgery, the RTF recommended that physical examination and imaging data be added to the minimal dataset. Straight leg raising, lower extremity reflexes, and lower extremity strength as indicators of radiculopathy were recommended as a minimum physical examination. Lumbar MRI was recommended in such studies as the minimal imaging evaluation.

In older adults, there is increased likelihood of hip osteoarthritis contributing to LBP. Thus, for studies of adults exclusively over age 65, the RTF recommended tests of hip rotation, to help identify potential osteoarthritis.

In studies focused on behavioral or mood correlates of chronic back pain, the RTF recommended that investigators be free to incorporate additional measures. These could include, and are not limited to, assessments relevant to the specific study interests—e.g., measures of emotional status, physical function, pain behaviors, substance abuse, personal abuse, or quality of life. It is important that such measures have published reliability, validity, sensitivity, and specificity data at least equal to or greater than those of the recommended PROMIS items. In addition, they should have population-based normative data available to be included when relevant to the reported design and data-analytic methods. The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) statement is one starting point for selecting desired supplemental measures.⁹⁷

Outcome Recommendations. The RTF recognized that much of the baseline minimal dataset, such as the PROMIS measures, were also perfectly appropriate as outcome measures. It was also recognized that the primary outcome of a clinical study might vary, depending on study aims. For example, some might focus on pain relief, but others might focus on return to work, physical function, mood, or need for subsequent therapy. Thus, the RTF did not make a recommendation regarding a minimal outcome dataset.

Reporting of outcomes. An important discussion centered on this topic. There was a general agreement that for (at least theoretically) continuous measures such as pain or function, not only should mean scores and score changes be reported, but also the proportion of subjects achieving certain thresholds (e.g., the proportion achieving a pre-specified minimal clinically important change). Investigators have proposed minimally important differences in PROMIS short forms, at least in the context of cancer therapy.¹¹⁵ The FDA refers to calculating the percentage of study participants who achieve such landmarks as a “responder analysis.”¹¹⁶

Other expert panels have suggested, for example, that a 30-percent improvement in pain or function might be a clinically important difference, and recommended reporting the proportion of subjects with this degree of improvement.¹¹⁷ Statistical analysts have suggested potential problems with the use of percentage changes,¹¹⁸ but one might alternatively specify a certain number of points as the relevant change, or the percentage of subjects reaching some threshold pain level (e.g., a pain score less than 3 out of 10).

An attractive option to the RTF was reporting the “cumulative distribution function” of responses for the treatment and control group. This is a continuous plot of the proportion of patients at each scale score who experience change at that level or better. This amounts to calculating the percentage of responders at each value of the outcome score. This approach acknowledges the lack of consensus on the approach for establishing a responder threshold, and provides information for any given threshold.¹¹⁶

Composite outcome measures. The RTF also discussed the potential for use of composite outcome measures. One member noted that it is common in studies of osteoarthritis to require improvement in pain score, functional status, and global self-assessment before judging treatment to be successful. Such composites are often required in FDA trials for drug or device approval. For example, “success” in trials of artificial disc replacement required functional improvement of 15 points on the Oswestry scale, improvement of quality of life on the SF-36, proper radiographic placement, and absence of new neurologic deficits or revision surgery. Such composites offer the potential advantage of defining success in terms that are more clearly clinically important, and not merely statistically significant.

However, the RTF concluded that in the absence of data on performance of such composite measures for LBP, it could not make a recommendation about composite outcome measures. Instead, this was recommended as an important topic for future research.

Timeframes for outcome measures. The RTF chose not to make specific recommendations for the timing of outcome assessments, because appropriate timing would vary depending on intervention. For some treatments (e.g., analgesics or spinal manipulation), the goal may be short-term relief, and for others (e.g., surgery), more often long-term relief. For studying patients with chronic pain, longer-term followup is generally preferred.

Adverse events. Reporting of adverse events was also recognized as an important outcome measure. Because likely adverse events vary enormously with the nature of an intervention, the RTF did not make recommendations for reporting specific adverse events. There was general agreement that for most intervention studies, it would be desirable to specify certain adverse events in advance and measure them prospectively, along with open-ended reporting of unanticipated events.

Recommendations for Research on the Proposed Standards. The RTF identified several important gaps in knowledge relevant to its work on classification and the minimal dataset, and encouraged further research on them:

Prognosis. Improving prognostic stratification (i.e., the ability to classify patients by subgroups based on their conditions) of patients with cLBP is important clinically to help guide the nature and intensity of therapy, and for researchers so that they can adjust for confounding and improve study comparability. Recent projects such as STarT Back have made important advances in this regard.⁵⁵⁻⁵⁸ Their generalizability to interventions and populations outside primary care remains uncertain, however, and additional work is needed.

Composite outcome measures. An ongoing frustration in the field has been the seeming lack of progress in reducing back-related disability at a population level. In part, this may be a result of claiming treatment efficacy based on statistically significant but clinically trivial results. More work is needed to understand how certain outcome scores are associated with major events such as return to work.

Composite outcome measures, such as those requiring simultaneous improvement in pain, function, and global self-assessment, may move the field closer to important outcomes. However, more data are needed to determine the performance of such measures in terms of validity, reliability, responsiveness, and prognostic value.

Patient stakeholder assessment. Little work has addressed the outcomes judged most important by patients with cLBP. Such outcomes may vary with demographic features and diagnosis.

Psychometric properties of the proposed minimal dataset. Extensive effort has been made to validate the PROMIS measures,^{104-109, 120-122} but there is modest information on their performance specifically in the context of cLBP. Further data on the precision of the domains is important (e.g., the optimal number of items), as well as data on responsiveness to change and sensitivity to small differences. Creating a “crosswalk” of scores with legacy measures such as the Oswestry and Roland-Morris disability questionnaires is also important.

Dissemination of the Report of the NIH Research Task Force on Recommendations for Standards of Research into Chronic Low-Back Pain. The NIH Pain Consortium has accepted this report and is recommending that all NIH institutes and centers require that all grant applications proposing clinical studies of cLBP utilize the research standards set forth in this report. Further, NIH encourages all other agencies that sponsor research on this condition to similarly incorporate these research standards for their respective studies.

DISCUSSION

Consistent with its charge from NIH, the RTF strove to have its recommendations serve as a set of standards for conducting research into the complex, intertwined factors that influence the onset, natural history, and clinical course of cLBP. This condition remains one of the most important and costly affecting the U.S. population. As adopted by NIH, the proposed research standards have the potential to result in standardized methods to identify cLBP research cases, provide clinical researchers with a core set of evidence-based measures deemed critical for incorporation into any cLBP scientific report, and allow cLBP published reports to be compared.

The new research standards should also improve the comparability of research studies on cLBP, facilitate pooling data from multiple studies (e.g., for meta-analyses), and improve the ability to define phenotypes among patients with LBP. They represent standards that will be useful not only for the conduct of cLBP research, but also for the reporting of scientific findings. They do so by allowing comparable core summary statistics to be included in all published reports, yet not interfering with collection of specific measures needed to address specific research questions.

After extended review and discussion, the RTF concluded that, given the current state of scientific evidence on cLBP, it was not realistic to create operationally defined RDC for subsets of cLBP. While creation of such criteria has proven beneficial to research for some other conditions (e.g., temporomandibular joint disorders,²¹ Alzheimer’s Disease,¹²³ and others), the multifactorial nature of most cases of cLBP decreased enthusiasm for attempting to do so in this condition. However, creation of a minimal dataset that would be utilized in all future studies of cLBP would achieve many of the same goals.

In summary, the RTF has recommended a definition of cLBP and classifying it in terms of its impact in addition to any presumed pathoanatomic diagnosis. Impact is conceived as a combination of pain intensity, interference with activities, and physical function. The RTF has also recommended a uniform minimal dataset, with recommendations for medical history, physical examination, diagnostic tests, and self-report measures of physical function, depression, sleep disturbance, pain intensity, and interference. Finally, recommendations have been made for reporting patient outcomes.

The RTF believes these recommendations can realistically help to advance the field, help to resolve controversies, and facilitate future research addressing the genomic, neurologic, and other mechanistic substrates of cLBP. Furthermore, it can help to reveal the biologic-behavioral interfaces that confound our present day understanding of cLBP and its evidence-based management.

It is anticipated that the RTF recommendations will become a dynamic document, and that the proposals are likely to undergo continual improvement. The research agenda that is proposed should facilitate this evolution.

REFERENCES

1. Pizzo PA, Clark NM, Carter-Pokras O, et al. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*. Committee on Advancing Pain Research, Care, and Education, Institute of Medicine. Washington, DC: National Academies Press; 2011.
2. Hoy D, Bain C, Williams G, et al. A systematic review of the global prevalence of low back pain. *Arthritis and Rheumatism*. 2012;64(6):2028-2037.
3. Hoy D, Brooks P, Blyth F, et al. The epidemiology of low back pain. *Best Practice & Research Clinical Rheumatology*. 2010;24(6):769-781.
4. Deyo RA, Mirza SK, Turner JA, et al. Overtreating chronic back pain: time to back off? *Journal of the American Board of Family Medicine*. 2009;22:62-68.
5. Rajaei SS, Bae HW, Kanim LE, et al. Spinal fusion in the United States: analysis of trends from 1998 to 2008; *Spine*. 2012;37:67-76.
6. Kenan K, Mack K, Paulozzi L. Trends in prescriptions for oxycodone and other commonly used opioids in the United States, 2000-2010. *Open Medicine*. 2012;6(2):e41-e47.
7. Tracey I, Bushnell MC. How neuroimaging studies have challenged us to rethink: is chronic pain a disease? *Journal of Pain*. 2009;10(11):1113-1120.
8. Seminowicz DA, Wideman TH, Naso L, et al. Effective treatment of chronic low back pain in humans reverses abnormal brain anatomy and function. *Journal of Neuroscience*. 2011;31(20):7540-7550.
9. Baliki MN, Petre B, Torbey S, et al. Corticostriatal functional connectivity predicts transition to chronic back pain. *Nature Neuroscience*. 2012;15(8):1117-1119.
10. Rodriguez-Raecke R, Niemeier A, Ihle K, et al. Structural brain changes in chronic pain reflect probably neither damage nor atrophy. *PLoS One*. 2013;8(2):e54475.
11. Beneck GJ, Kulig K. Multifidus atrophy is localized and bilateral in active persons with chronic unilateral low back pain. *Archives of Physical Medicine and Rehabilitation*. 2012;93(2):300-306.
12. Battié MC, Niemelainen R, Gibbons LE, et al. Is level- and side-specific multifidus asymmetry a marker for lumbar disc pathology? *Spine Journal*. 2012;12(10):932-939.
13. Barker KL, Shamley DR, Jackson D. Changes in the cross-sectional area of multifidus and psoas in patients with unilateral back pain: the relationship to pain and disability. *Spine*. 2004;29(22):e515-e519.

14. Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. *Radiologic Clinics of North America*. 2012;50(4):569-585.
15. Bunzlj S, Watkins R, Smith A, et al. Lives on hold: a qualitative synthesis exploring the experience of chronic low-back pain. *Clinical Journal of Pain*. 2013;Jan 30 [Epub ahead of print].
16. Negrini S, Zaina F. The chimera of low back pain etiology. A clinical rehabilitation perspective. *American Journal of Physical Medicine & Rehabilitation*. 2013; 92: 93-97
17. Chou R, Huffman LH. Medications for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Annals of Internal Medicine*. 2007;147(7):492-504.
18. Chou R, Huffman LH. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Annals of Internal Medicine*. 2007;147(7):492-504.
19. Chou R, Atlas SJ, Stanos SP, et al. Nonsurgical interventional therapies for low back pain: a review of the evidence for an American Pain Society clinical practice guideline. *Spine*. 2009;34(10):1078-1093.
20. Chou R, Baisden J, Carragee EJ, et al. Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. *Spine*. 2009;34(10):1094-1109.
21. Dworkin SF. Research diagnostic criteria for temporomandibular disorders: current status & future relevance. *Journal of Oral Rehabilitation*. 2010;37(10):734-743.
22. Bogduk N. On the definitions and physiology of back pain, referred pain, and radicular pain. *Pain*. 2009; 147(1-3):17-19.
23. Genevay S, Atlas SJ, Katz JN. Variation in eligibility criteria from studies of radiculopathy due to a herniated disc and of neurogenic claudication due to lumbar spinal stenosis: a structured literature review. *Spine*. 2010;35(7):803-811.
24. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA*. 2008;299(6):656-664.
25. Waddell G. 1987 Volvo Award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine*. 1987;12(7):632-644.
26. Shah RV. Spine pain classification: the problem. *Spine*. 2012;37(22):1853-1855.
27. Stanton TR, Latimer J, Maher CG, et al. A modified Delphi approach to standardize low back pain recurrence terminology. *European Spine Journal*. 2011;20(5):744-752.
28. Konstantinou K, Hider SL, Jordan JL, et al. The impact of low back-related leg pain on outcomes as compared with low back pain alone: a systematic review of the literature. *Clinical Journal of Pain*. 2013;29(7):644-654.
29. Deyo RA. Diagnostic evaluation of LBP: reaching a specific diagnosis is often impossible. *Archives of Internal Medicine*. 2002;162(13):1444-1448.
30. Abraham I, Killackey-Jones B. Lack of evidence-based research for idiopathic low back pain: the importance of a specific diagnosis. *Archives of Internal Medicine*. 2002;162(13):1442-1444.
31. Fairbank J, Gwilym SE, FranceJC, et al. The role of classification of chronic low back pain. *Spine*. 2011;36(21 Suppl):S19-S42.
32. Apeldoorn AT, Bosmans JE, Ostelo RW, et al. Cost-effectiveness of a classification-based system for sub-acute and chronic low back pain. *European Spine Journal*. 2012;21(7):1290-1300.
33. Croft P, Dunn KM, Von Korf M. Chronic pain syndromes: you can't have one without another. *Pain*. 2007;131(3):237-238.
34. Loisel P, Vachon B, Lemaire J, et al. Discriminative and predictive validity assessment of the Quebec Task Force classification. *Spine*. 2002 27(8):851-857.

35. Quebec Task Force on Spinal Disorders. Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. Chapter 3: Diagnosis of the problem (the problem of diagnosis). *Spine*. 1987;12(7 Suppl):S16-S21.
36. Hall H, McIntosh G, Boyle C. Effectiveness of a low back pain classification system. *Spine Journal*. 2009;9(8):648-657.
37. Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Annals of Internal Medicine*. 2004;141(12):920-928.
38. Von Korff M. Assessment of chronic pain in epidemiological and health services research: empirical bases and new directions. In: Turk DC, Melzack R, eds. *Handbook of Pain Assessment*. 3rd ed. New York, NY: Guilford Press; 2011:455-473.
39. Salovey P, Seiber WJ, Smith AF, et al. *Reporting Chronic Pain Episodes on Health Surveys*. Washington, DC: Centers for Disease Control, National Center for Health Statistics; 1992. HHS publication no. (PHS) 92-1081.
40. Stewart WF, Lipton RB, Simon D, et al. Validity of an illness severity measure for headache in a population sample of migraine sufferers. *Pain*. 1999;79(2-3):291-301.
41. Von Korff M, Dunn KM. Chronic pain reconsidered. *Pain*. 2008;138(2):267-276.
42. Von Korff M, Miglioretti DL. A prognostic approach to defining chronic pain. *Pain*. 2005;117(3):304-313.
43. 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(1-2):48-54.
44. 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(2):389-397.
45. Mallen CD, Peat G, Thomas E, et al. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *British Journal of General Practice*. 2007;57(541):655-661.
46. Friedman BW, Mulvey L, Davitt M, et al. Predicting 7-day and 3-month functional outcomes after an ED visit for acute nontraumatic low back pain. *American Journal of Emergency Medicine*. 2012;30(9):1852-1859.
47. Von Korff M, Ormel J, Keefe FJ, et al. Grading the severity of chronic pain. *Pain*. 1992;50(2):133-149.
48. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Annals of the Academy of Medicine*. 1994;23(2):129-138.
49. Krebs EE, Lorenz KA, Bair MJ, et al. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. *Journal of General Internal Medicine*. 2009;24(6):733-738.
50. Edeleen MO, Saliba D. Correspondence of verbal descriptor and numeric rating scales for pain intensity: an item response theory calibration. *Journals of Gerontology A: Biological Sciences and Medical Sciences*. 2010;65(7):778-785.
51. Jensen MP, Smith DG, Ehde DM, et al. Pain site and the effects of amputation pain: further clarification of the meaning of mild, moderate, and severe pain. *Pain*. 2001;91(3):317-322.
52. Kapstad H, Hanestad BR, Langeland N, et al. Cutpoints for mild, moderate and severe pain in patients with osteoarthritis of the hip or knee ready for joint replacement surgery. *BMC Musculoskeletal Disorders*. 2008;9:55.
53. Gerbershagen HJ, Rothaug J, Kalkman CJ, et al. Determination of moderate-to-severe postoperative pain on the numeric rating scale: a cut-off point analysis applying four different methods. *British Journal of Anaesthesia*. 2011;107(4):619-626.

54. Jones KR, Vojir CP, Hutt E, et al. Determining mild, moderate, and severe pain equivalency across pain-intensity tools in nursing home residents. *Journal of Rehabilitation Research and Development*. 2007;44(2):305-314.
55. 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(5):632-641.
56. Hill JC, Dunn KM, Main CJ, et al. Subgrouping low back pain: a comparison of the STarT Back Tool with the Orebro Musculoskeletal Pain Screening Questionnaire. *European Journal of Pain*. 2010;14(1):83-89.
57. Hill JC, Vohora K, Dunn KM, et al. Comparing the STarT back screening tool's subgroup allocation of individual patients with that of independent clinical experts. *Clinical Journal of Pain*. 2010;26(9):783-787.
58. Hill JC, Whitehurst DG, 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*. 2011;378(9802):1560-1571.
59. Beneciuk JM, Bishop MD, Fritz JM, et al. The STarT Back Screening Tool and individual psychological measures: evaluation of prognostic capabilities for low back pain clinical outcomes in outpatient physical therapy settings. *Physical Therapy*. 2013;93(3):321-333.
60. Bruyere O, Demoulin M, Brereton C, et al. Translation validation of a new back pain screening questionnaire (the STarT Back Screening Tool) in French. *Archives of Public Health*. 2012;70(1):12.
61. Field J, Newell D. Relationship between STarT Back Screening Tool and prognosis for low back pain patients receiving spinal manipulative therapy. *Chiropractic & Manual Therapies*. 2012;20(1):17.
62. 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(5):722-732.
63. Gusi N, del Pozo-Cruz B, Olivares PR, et al. The Spanish version of the "STarT Back Screening Tool" (SBST) in different subgroups. *Atencion Primaria*. 2011;43(7):356-361.
64. Kongsted A, Johannesen E, Leboeuf-Yde C. Feasibility of the STarT back screening tool in chiropractic clinics: a cross-sectional study of patients with low back pain. *Chiropractic & Manual Therapies*. 2011;19:10.
65. Morso L, Albert H, Kent P, et al. Translation and discriminative validation of the STarT Back Screening Tool into Danish. *European Spine Journal*. 2011;20(12):2166-2173.
66. Wideman TH, Hill JC, Main CJ, et al. Comparing the responsiveness of a brief, multidimensional risk screening tool for back pain to its unidimensional reference standards: the whole is greater than the sum of its parts. *Pain*. 2012;153(11):2182-2191.
67. Dunn KM, Von Korff M, Croft P. Defining chronic pain by prognosis. In: Hasenbring MI, Rusu AC, Turk DC, eds. *From Acute to Chronic Back Pain: Risk Factors, Mechanisms, and Clinical Implications*. 1st ed. New York, NY: Oxford University Press; 2012:21-40.
68. Martell BA, O'Connor PG, Kerns RD, et al. Systematic review: opioid treatment for chronic back pain: prevalence, efficacy, and association with addiction. *Annals of Internal Medicine*. 2007;146(2):116-127.
69. Muller S, Thomas E, Dunn KM, et al. A prognostic approach to defining chronic pain across a range of musculoskeletal conditions. *Clinical Journal of Pain*. (In press).
70. Von Korff M, Shortreed S, Saunders KW, et al. Comparison of prognostic screeners for back pain risk stratification. (Under review).

71. Turner JA, Shortreed SM, Saunders KW, et al. Optimizing prediction of back pain outcomes. *Pain*. 2013;154(8):1391-1401.
72. Wideman TH, Hill JC, Main CJ, Lewis M, Sullivan MJ, Hay EM. Comparing the responsiveness of a brief, multidimensional risk screening tool for back pain to its unidimensional reference standards: the whole is greater than the sum of its parts. *Pain*. 2012;153(11):2182-2191.
73. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R. eds. *Handbook of Pain Assessment*. 2nd ed. New York, NY: Guilford Press; 2001:15-34.
74. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Annals of the Academy of Medicine*. 1994;23(2):129-138
75. Dworkin RH, Turk DC, Revicki DA, et al. Development and initial validation of an expanded and revised version of the Short-form McGill Pain Questionnaire (SF-MPQ-2). *Pain*. 2009;144(1-2):35-42.
76. Margolis RB, Tait RC, Krause SJ. A rating system for use with patient pain drawings. *Pain*. 1986;24(1):57-65.
77. 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(2):141-144.
78. Fairbank J, Couper J, Davies JB, et al. The Oswestry Low Back Pain Disability Questionnaire. *Physiotherapy*. 1980;66(8):271-273.
79. Fordyce WE. *Behavioral Methods for Chronic Pain and Illness*. St. Louis, MO: Mosby; 1976.
80. Keefe FJ, Williams DA, Smith SJ. Assessment of pain behaviors. In: Turk DC, Melzack RJ, eds. *Handbook of Pain Assessment*. 2nd ed. New York, NY: Guilford Press; 2001:170-190.
81. Kerns RD, Haythornthwaite J, Rosenberg R, et al. The Pain Behavior Checklist (PBCL) – function, structure and psychometric properties. *Journal of Behavioral Medicine*. 1991;14(2):155-169.
82. Kori S, Miller R, Todd D. Kinisophobia: a new view of chronic pain behavior. *Pain Management*. 1990;Jan/Feb:35-43.
83. Waddell G, Newton M, Henderson I, et al. A Fear Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*. 1993;52(2):157-168.
84. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine*. 2002;64(2):258-266.
85. Kroenke K, Spitzer R, Williams JW. The PHQ-9: validity of a brief depression measure. *Journal of General Internal Medicine*. 2001;16(9):606-616.
86. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*. 1983;67(6):361-370.
87. Radloff LS. The CES-D scale: a self-report depression scale for research in general populations. *Applied Psychological Measurement*. 1977;1(3):385-401.
88. Beck AT, Steer RA, Ball Ret al. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*. 1996;67(3):588-597.
89. Spitzer RL, Kroenke K, Williams JB, et al. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*. 2006;166(10):1092-1097.
90. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. *Psychological Assessment*. 1995;7(4):524-532.
91. Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and adjustment. *Pain*. 1983;17(1):33-44.
92. Weathers FW, Ford J. Psychometric properties of the PTSD Checklist (PCL–C, PCL–S, PCL–M, PCL–PR). In: Stamm BH, ed. *Measurement of Stress, Trauma, and Adaptation*. Lutherville, MD: Sidran Foundation & Press: 1996:250-252.
93. Von Korff M, Ormel J, Keefe FJ, et al. Grading the severity of chronic pain, clinical section. *Pain*. 1992;50(2):133-149.

94. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). *Medical Care*. 1992;30(6):473-483.
95. Ware JE, Kosinski M, Keller SD. A 12-item Short-Form Health Survey. Construction of scales and preliminary tests of reliability and validity. *Medical Care*. 1996;34(3):220-233.
96. Hurst NP, Kind P, Ruta D, et al. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *British Journal of Rheumatology*. 1997;36(5):551-559.
97. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113(1-2):9-19.
98. Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2009;146(3):238-244.
99. Atlas S, Deyo RA, van den Ancker M, et al. The Maine-Seattle back questionnaire: a 12-item disability questionnaire for evaluating patients with lumbar sciatica or stenosis: derivation and validation. *Spine*. 2003;28(16):1869-1876.
100. Patrick DL, Deyo RA, Atlas SJ, et al. Assessing health-related quality of life in patients with sciatica. *Spine*. 1995;20(17):1899-1909.
101. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *Journal of Clinical Epidemiology*. 1992;45(6):613-619.
102. Cook KF, Choi SW, Crane PK, et al. Letting the CAT out of the bag: comparing computer adaptive tests and an 11-item short form of the Roland-Morris disability questionnaire. *Spine*. 2008;33(12):1378-1383.
103. Cherkin DC, Deyo RA, Street J, et al. Predicting poor outcomes for back pain seen in primary care using patients' own criteria. *Spine*. 1996;21(24):2900-2907.
104. Rose M, Bjorner JB, Becker J, et al. Evaluation of a preliminary physical function item bank supports the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *Journal of Clinical Epidemiology*. 2008;61(1):17-33.
105. Amtmann D, Cook KF, Jensen MP, et al. Development of a PROMIS item bank to measure pain interference. *Pain*. 2010;150(1):173-182.
106. Garcia SF, Cella D, Clauser SB, et al. Standardizing patient-reported outcomes assessment in cancer clinical trials: a PROMIS initiative. *Journal of Clinical Oncology*. 2007;25(32):5106-5112.
107. Buysse DJ, Yu L, Moul DE, et al. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. *Sleep*. 2010;33(6):781-792.
108. Pilkonis PA, Choi SW, Reise SP, et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS): depression, anxiety, and anger. *Assessment* 2011;18(3):263-283.
109. Hahn EA, DeVillis RF, Bode RK, et al. Measuring social health in the Patient-Reported Outcomes Measurement Information System (PROMIS): item bank development and testing. *Quality of Life Research*. 2010;19(7):1035-1044.
110. Froud R, Eldridge S, Kovacs F, et al. Reporting outcomes of back pain trials: a modified Delphi study. *European Journal of Pain*. 2011;15(10):1068-1074.
111. Tubach F, Ravaud P, Martin-Mola E, et al. Minimal clinically important improvement and patient acceptable symptomatic state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic back pain, hand osteoarthritis, and hip and knee osteoarthritis: the REFLECT multinational study. *Arthritis Care & Research*. 2012;64(11):1699-1707.
112. Brown RL, Leonard T, Saunders LA, Papasouliotis O. A two-item conjoint screen for alcohol and other drug problems. *Journal of the American Board of Family Practice*. 2001;14(2):95-106.

113. Accessing PROMIS short form instruments. Assessment Center Web site. Accessed at <http://www.assessmentcenter.net/PromisForms.aspx> on October 1, 2013.
114. Linking Tables. PROsetta Stone Web site. Accessed at www.prosettastone.org/LinkingTables/Pages/default.aspx on October 1, 2013.
115. Yost KJ, Eton DT, Garcia SF, et al. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *Journal of Clinical Epidemiology*. 2011;64(5):507-516.
116. McLeod LD, Coon CD, Martin SA, et al. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. *Expert Review of Pharmacoeconomics and Outcomes Research*. 2011;11(2):163-169.
117. Farrar JT, Young JP, LaMoreaux L, et al. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001;94(2):149-158.
118. Vickers AJ. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. *BMC Medical Research Methodology*. 2001; 1:6.
119. Zigler J, Delamarter R, Spivak JM, et al. Results of the prospective, randomized, multicenter Food and Drug Administration Investigational Device Exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. *Spine*. 2007;32(11):1155-1162.
120. Revicki DA, Cook KF, Amtmann D, et al. Exploratory and confirmatory factor analysis of the PROMIS pain quality item bank. *Quality of Life Research*. 2013;Jul 9. [Epub ahead of print].
121. Askew RL, Kim J, Chung H, et al. Development of a crosswalk for pain interference measured by the BPI and PROMIS pain interference short form. *Quality of Life Research*. 2013;Mar 29 [Epub ahead of print].
122. Kim J, Chung H, Amtmann D, et al. Measurement invariance of the PROMIS pain interference item bank across community and clinical samples. *Quality of Life Research*. 2013;22(3):501-507.
123. Sarazin M, de Souza LC, Lehericy S, et al. Clinical and research diagnostic criteria for Alzheimer's disease. *Neuroimaging Clinics of North America*. 2012;22(1):23-32.

APPENDICES

Appendix 1. Research Task Force (RTF) Meeting Agendas

- 1.1 Agenda for March 5–6, 2012
- 1.2 Agenda for October 1–2, 2012
- 1.3 Agenda for March 11–12, 2013

Appendix 2. Pre-RTF Meeting Background and Information Distribution

- 2.1 NIH Pain Consortium Charge to the RTF
- 2.2 Advancing Research on cLBP: Developing Standards for Research
- 2.3 Some Questions to Consider for Meeting 1
- 2.4 Rationale for Developing a Biopsychosocial, Dual-Axis RDC/cLPBS Diagnostic and Classification System at the Present Time
- 2.5 First Meeting Overview and Charge to Workgroups
- 2.6 Chronic Low-Back Pain: Classification; Research Needs and Recommendations
- 2.7 Overview of Minimal Dataset for Research on cLBP
- 2.8 Preamble and Recommendations of the NIH Task Force on Research Standards for Chronic Low-Back Pain

Appendix 3. Surveying the RTF: Survey Data on Preferences and Actions

- 3.1 Prioritizing Minimal Dataset Items: History, Demographics, Physical Findings
- 3.2 Prioritizing Minimal Dataset Items: Psychological, Behavioral and Psychosocial
- 3.3 Feasibility of RTF Developing an RDC/cLBP Diagnostic and Classification System

Appendix 4. Recommended Multidimensional Minimal Dataset for Research on cLBP

APPENDIX 1

Research Task Force (RTF) Meeting Agendas

APPENDIX 1.1

Agenda for March 5–6, 2012

NIH Task Force Meeting on Research Standards for Chronic Low-Back Pain

DoubleTree by Hilton
 8120 Wisconsin Avenue, Bethesda, MD 20814
 March 5–6, 2012

Time	Topic
Mon., 3/5 8:30 a.m.	Opening: NIH welcome , introduction of other NIH people present; meeting logistics <ul style="list-style-type: none"> • Review of NIH-supported events that led to this meeting • Nature and size of the problem • Rationale for the Task Force from the NIH perspective • Introductions
9:00	Overview: The state-of-the-art of standards for research: classifying cLBP and its outcomes <ul style="list-style-type: none"> • Emphasis on problems with available research and with doing research into cLBP • Areas of concern and the diverse stakeholders: People in pain: health care delivery systems; varied clinicians; drugs and devices; regulatory agencies; compensation for back pain • Why emphasis on case definitions, diagnostic criteria, and outcome measures • Review of some available diagnostic/classification systems and outcome measurement strategies • Why a research focus for the Task Force
9:30	Group discussion: Most impactful aspects of research to address; scope of Task Force work; reaching consensus on approach
10:30	Break
10:50	Critical factors and perspectives in developing research-oriented diagnostic criteria for cLBP: Based on items listed in “Advancing Research on cLBP: Developing Standards for Research”
11:15	Group discussion Arrive at consensus on developing an evidence-based and research-focused system, the RDC/cLBP, using perspective just presented: <ol style="list-style-type: none"> 1. Based on a biopsychosocial model 2. Interdisciplinary research designs a high priority 3. Need for measurable variables and criteria-operational definitions 4. Emphasis on reliability of criteria 5. Suitability for population, observational, and clinical research 6. Dual axis (or multiple axes)
12:15 p.m.	Lunch

Time	Topic
1:30	<p>Review the available scientific literature for data applicable to an RDC/cLBP</p> <ul style="list-style-type: none"> • Provide illustrative examples of putative, suitably-defined criteria and variables; 5-minute presentations by Task Force members from orthopedics, physical therapy, manipulative therapy • Discussion of current state-of-the art: areas of consensus or disagreement; availability of scientific data on reliability and validity of classification schemes • Logistics, room designations; relevant literature, designating a recorder to report to larger group
2:15	<p>Create Work Groups of about five members each</p> <ol style="list-style-type: none"> 1. Axis I variables—the physical variables and diagnostic criteria measured objectively 2. Axis II variables—the psychological and psychosocial variables measured by self-report 3. Other axes and/or critical variables to consider for inclusion in RDC/cLBP <ul style="list-style-type: none"> • Prognosis; predictors of chronicity • Genetics and epigenetics • Central neuroprocessing <p>Charge to Work Groups:</p> <ul style="list-style-type: none"> • Consider available scientific evidence to identify delivery systems with electronic records; NHANES; published databases/data sets • Identify available databases and research studies suitable for mining, e.g.: <ul style="list-style-type: none"> ○ Useful case definitions of cLBP ○ Most common subtypes of cLBP and assessment of scientific data for their classification ○ Available diagnostic systems—whole or in part ○ Available operationally defined physical, psychological, and psychosocial variables • Identify significant gaps in available scientific evidence relevant to each of the above
3:15	Break
3:30	Work Group presentations to assembled Task Force.
5:00	Summary and group discussion: Major conclusions from day's activities, highlighting where there is agreement over what's known, suitable for inclusion; critical gaps.
5:30	Conclude session

Time	Topic
<p>Tue. 3/6 8:30 a.m.</p>	<p>Attempt to develop an RDC/cLBP system</p> <p>The objective of this exercise is to shed light on where difficulties are encountered when attempting such a scientifically based diagnostic and classification system for cLBP. The results will hopefully sharpen the focus on next steps to be taken to accomplish the principal goal of the Task Force.</p> <p>Use data made available through this meeting, the scientific literature, and/or clinical judgment (if the group’s consensus is that no scientific evidence is available but the variable or criterion being considered is important).</p> <p>It is suggested that each Work Group follow a similar work plan, using the major topics identified below to arrive at their respective sets of diagnostic or classificatory criteria:</p> <ul style="list-style-type: none"> • Select most commonly occurring sub-types for which RDC/cLBP should be created. • Categorize the most common subtypes according to a descriptive or etiologic system (muscle, disk, disability, etc.). • Specify the necessary variables that have to be assessed to arrive at diagnostic/classification criteria for each subtype. • Specify the measurement instruments—including physical examination procedures, imaging types, lab tests, self-report, and observational measures—to be used to quantify each variable being assessed. • Specify how variables will be grouped to generate the diagnostic/classification criteria for each subtype. • Specify at each step what is known about the availability, reliability, and the validity of each of the diagnostic measurement instruments and where unknown, specify the type of scientific evidence needed to justify using the measuring instrument and to allow incorporating the variables and criteria into the RDC/CLBP.
10:30	Work Group reports to the assembled Task Force with Q & A
12:00 p.m.	Lunch
Time	Topic
1:00	<p>Integrate Axis I and Axis II</p> <p>Provide a conceptual model of cLBP, which meaningfully integrates Axis I physical findings data with Axis II self-report data to allow a meaningful, comprehensive understanding of the cLBP patient, and promote evidence-based decisions regarding tailored Axis I and Axis II treatments. Create an integrated definition of treatment “success.”</p>
2:30	<p>Summarize progress and develop a strategy for conducting research into:</p> <ul style="list-style-type: none"> • Reliability, validity, and clinical utility of putative RDC/cLBP system and outcome measures • Clinical effectiveness and efficacy of the RDC/cLBP in predicting its natural history and controlling its clinical course to relieve pain and disability associated with cLBP as one of the most important, expensive, and high-impact public health problems worldwide.
3:30	Wrap up and plan next step(s)
4:00	Adjourn

APPENDIX 1.2

Agenda for October 1–2, 2012

NIH Pain Consortium: Research Task Force on Research Standards for Chronic Low-Back Pain

NIH Conference Center, Terrace Level – Rooms 508 and 509 5635 Fishers Lane, Rockville, MD 20852
October 1–2, 2012

Meeting Agenda	
Objectives: Monday, 10/1/2012	
I. Final consensus on defining cLBPS and predicting chronicity	
II. . A/B Final consensus on Axis I/II minimal datasets	
Objectives: Tuesday 10/2/2012	
III. Develop a Research Agenda for cLBPS:	
A. Reaching consensus on an RDC/cLBPS	
B. Developing a biopsychosocial program of research into the diagnosis, prevention, and management of cLBPS for NIH-funded researchers	
C. Planning a stakeholders’ meeting	
Day 1	I. Finalize definitions of cLBPS; confirm, refine, finalize minimal datasets
8:30-8:45 a.m.	Opening: NIH welcome by Josephine Briggs, M.D. (NCCAM) and James Panagis, M.D. (NIAMS) <ul style="list-style-type: none"> • Rationale for the Research Task Force (RTF) from the NIH perspective
8:45-9:00	Overview of agenda (Deyo) <ol style="list-style-type: none"> 1. Final consensus on definition of Chronic Low Back Pain Syndrome (cLBPS) 2. Finalize minimal databases as prerequisite to any next steps for the RTF 3. Initiate strategies for developing a comprehensive research agenda for cLBPS, including review of RDC/cLBPS status
9:00-10:30	Consensus on final definition of cLBPS and issues in predicting chronicity (Von Korff)
10:30-10:50	Break
10:50-11:20	IIA. Axis I minimal dataset: Summarize activities and results; overall review of Axis I dataset (Deyo).
11:20 a.m.-1:00 p.m.	RTF group: Confirms, refines, and finalizes Axis I minimal dataset. Includes potential of Axis I for predictors of chronicity. Moderator-led (Deyo)
1:00-2:30	Lunch
2:30-3:00	IIB. Axis II minimal dataset: Summarize activities and results; define terms; identify major issues confronting Axis II dataset development and use. Includes predicting chronicity measure. (Turk)
3:00-4:45	RTF group: Confirms, refines, and finalizes Axis II minimal dataset. Includes predicting chronicity measure. Moderator-led (Turk and Dworkin).
4:45-5:15	Summary of the day: Status of definitions of cLBPS, chronicity, and Axis I/II datasets. (RD/SFD)

NIH Pain Consortium: Research Task Force on Research Standards for cLBPS Meeting Agenda: Tuesday, October 2, 2012	
III. Develop an NIH Research Agenda for cLBPS: A. Reaching consensus on an RDC/cLBPS B. Developing a biopsychosocial program of research into the diagnosis, prevention, and management of cLBPS for NIH-funded researchers C. Planning a stakeholders' meeting	
8:30-9:00 a.m.	IIIA. RDC/cLBPS: Arriving at an RFT consensus (Dworkin): Can and should the RTF undertake development of a biopsychosocial RDC/CLPBS system 1. Findings from the RTF Survey Data on developing an RDC/cLBPS 2. Group discussion: implications for research 3. Final consensus determination: to RDC/cLBPS or not?
9:00-10:30	IIIB. Developing a biopsychosocial program of research into the diagnosis, prevention, and management of cLBPS for NIH-funded researchers. Moderator-led (Deyo, Dworkin) RTF participation: Using Table below as a guide, each member will: 1. Identify 1 to 3 critical gaps in current knowledge in their field with regard to cLBPS: <ul style="list-style-type: none"> • Epidemiology, orthopedics, neurology, rehab medicine, behavioral medicine, health services, imaging, genetics, central brain processes
10:30-10:50	Break
10:50 a.m.-12:00 p.m.	IIIB (continued). Developing a biopsychosocial program of research. Moderator-led (Deyo, Dworkin) 2. Identify most-suitable research methods and key assessment/diagnostic and treatment outcome measures to close those gaps : <ul style="list-style-type: none"> • Epidemiologic, observational, experimental, and RCT studies to identify phenotypes, risk and predictive factors for Axis I and Axis II disease/illness progression; longitudinal, multi-institutional, cross-discipline methods • Outline potential research programs as indicators of RTF priorities for NIH
12:00-1:15	Lunch
1:15-2:30	III B. Stakeholders' Meeting: Initial Planning Need, Rationale, Timing (Killen and Khalsa) 1. Summary of RTF Activities: Research Standards and Research Agenda for cLBPS as basis for the stakeholders' meeting <ol style="list-style-type: none"> a. Biopsychosocial research perspective b. cLBPS defined c. Axis I and Axis II minimal datasets d. Consensus on RDC/CLBP 2. Are we ready for a meeting? 3. Logistics: potential invitees; roles for selected invitees; dates and venues 4. TBD
2:30-3:00	Summarize the day: Priorities for research questions and methods; plan for stakeholders' meeting; next steps—RTF Meeting #3 and/or stakeholders' meeting (Deyo, Khalsa)
3:00	Meeting adjourned

Table: Summarizing RTF recommendations: Moderator-led solicitation from RTF specialists of critical gaps in research in their respective areas: What new assessment, classification and outcome data is needed?

Research content domain	Critical gaps or research questions to be addressed	Suggested type of research study	Key assessment/ diagnostic and/or treatment outcome variables
Epidemiology			
Behavioral medicine			
Brain neuroscience			
CAM			
Chiropractic			
Genetics and epigenetics			
Health services			
Neurology			
Neurologic surgery			
Orthopedics			
Physical therapy			
Rehabilitation medicine			
Statistics and methodology			
Other			

APPENDIX 1.3

Agenda for March 11–12, 2013

NIH Pain Consortium, Research Task Force on Research Standards for Chronic Low Back Pain

Neuroscience Building, Conference Room D
 6001 Executive Boulevard, Rockville, MD
 20852
 March 11–12, 2013

Meeting Agenda	
Objectives: Monday, March 11, 2013	
I. Final consensus/approval: cLBP research definition; measuring cLBP impact; predictors of chronicity II. Final consensus/approval: Minimal multidimensional cLBP baseline research databases	
Objectives: Tuesday, March 12, 2013	
III. Recommendations for research on Task Force measures IV. Plan for a stakeholders' meeting (tentatively planned for period of November 2013 to February 2014)	
Monday, March 11, 2013	
8:30-8:50 a.m.	Opening: NIH welcome: Introduction of NIH attendees and guests NIH perspective on Task Force Progress and Objectives (what needs to get done this meeting)
8:50-9:10	Overview of agenda and progress to date (Deyo) 1. Defining cLBP, measuring its impact and risk levels for an unfavorable outcome 2. A multi-dimensional minimal database for baseline assessment in cLBP research 3. A comprehensive NIH research agenda for cLBPs—filling the gaps, setting priorities 4. Plan for today and tomorrow: the need for final resolutions
9:10-12:00	I. Final consensus/approval: cLBP research definition; measuring cLBP impact; predicting cLBP chronicity 9:10-9:45. 1. Consensus definition of cLBPS 9:45-10:30. 2. Measuring impact levels (Moderator: Deyo) 10:30-10:50. Break 10:50-12:15. 3. Risk of unfavorable outcomes—Predictors of chronicity (Moderator: Von Korff)
12:00-1:30 p.m.	Lunch
1:30-2:30	II. Final consensus/approval: minimal multidimensional cLBP baseline research databases: Part 1. Demographic and social status Part 2. Physical findings and medical history Group discussion and final recommendations (Moderator: Deyo)

2:30-3:40	II. Final consensus/approval: minimal multidimensional cLBP baseline research databases: Part 3. Pain report; behavioral, emotional, and psychosocial self-report measures Group discussion and final recommendations (Moderator: Dworkin) Part 4. Additional measures Group discussion and final recommendations (Moderators: Deyo and Dworkin)
3:40-4:00	Break
4:00-5:15	II. Final consensus/approval: minimal database-Part 5. Outcome Measures 4:00-4:20. Patient-Centered Outcomes Research Institute (PCORI). Overview of “Working Group for the Treatment Options for Back Pain-Targeted PCORI Funding Announcements” (Christine Goertz) 4:20-4:30. European COMET Initiative on Outcomes (Raymond Ostelo) 4:30-5:15. Part 5. Minimal Database: Outcome Measures and IMMPACT Introduction and Overview. Group discussion and final recommendations (Moderator: Turk)
5:15-5:45	Summary of the day: Have we met required needs to finalize recommendations: 1. Define cLBP; assess impact; determine risk for poor outcomes 2. Finalize minimal database: Part 1-Part 5. (Moderators: Deyo, Dworkin, Turk, Von Korff)

NIH Pain Consortium, Research Task Force on Research Standards for Chronic Low Back Pain

Neuroscience Building, Conference
Room D 6001 Executive Boulevard,
Rockville, MD March 11–12, 2013

Meeting Agenda	
II. (continued) Outcomes recommendation	
III. Recommendations for research on Task Force measures	
IV. Planning a stakeholders’ meeting	
8:00-8:15 a.m.	Overview of the day: What we need to accomplish (Deyo, Dworkin)
8:15-9:30	II. Final consensus/approval: minimal multidimensional cLBP baseline research databases: Part 5. Outcomes Recommendation Group discussion and final recommendations (Moderator: Turk)
9:30-11:30	III. Research recommendations for Task Force measures: (Moderators: Deyo, Dworkin) <ul style="list-style-type: none"> • Studies emerging from RTF recommendations: prevalence, distribution, reliability, validity of suggested measures included in minimal multidimensional dataset • General recommendations for addressing gaps and emerging areas in cLBP research

	<ul style="list-style-type: none"> Coordinating efforts with related groups regarding comprehensive standards for cLBP research
11:30 a.m.- 12:30 p.m.	Lunch
12:30-2:30	<p>IV. Plan for a stakeholders' meeting—based on recommendations of the RTF (Moderators: Killen, Panagis, and Khalsa)</p> <p>Need, Rationale, Timing</p> <ol style="list-style-type: none"> Summary of RTF activities: research standards and research agenda for cLBPs as basis for the stakeholders meeting <ol style="list-style-type: none"> Biopsychosocial research perspective cLBPs defined, measuring impact of cLBP, and defining risk levels for poor prognosis of chronicity Axis I and Axis II minimal datasets Recommendations for a broad-based NIH agenda for research into the prevention, diagnosis and classification, and treatment of cLBP Logistics: potential invitees; roles for selected invitees; dates and venues TBD
2:30-2:50	Summarize the day: Priorities for research questions and methods; plan for Stakeholders' meeting; next steps for the RTF (Deyo, Khalsa)
3:00	Meeting adjourned

APPENDIX 2

Pre-RTF Meeting Background and Information Distribution

APPENDIX 2.1

NIH Pain Consortium Charge to the RTF

March 5, 2012

Background

By any measure, the symptom of chronic low-back pain (cLBP) is a problem of enormous public health significance. Unfortunately, current best practices for its diagnosis and treatment are only partially successful. Furthermore, a clear cause can be identified and specifically treated in only a small minority of cases, and, given present knowledge, most cases are termed “idiopathic.”

Two NIH workshops convened in 2009 and 2010 focused on research needs and challenges with respect to back pain. Both identified the inability to compare, contrast, or pool data across different studies as a fundamental obstacle to progress in understanding the clinical problem of cLBP and identifying better approaches to its management. This assessment has been confirmed by many discussions with other scientists and clinicians as well who are engaged in back pain research. Specific issues include the following:

1. Inconsistent case definitions, study-eligibility criteria, and stratification criteria (e.g., for duration of the problem, severity of pain, degree of functional limitation, etc.)
2. Inconsistent diagnostic algorithms and procedures
3. Inconsistent outcome measures (e.g., for symptoms, function, and other objective measures).

The NIH Pain Consortium has endorsed a proposal that NIH convene a Chronic Low-Back Pain Research Task Force. The Task Force would oversee a longitudinal process that seeks to address these issues, initially through developing a set of standards for clinical research on cLBP.

The charge to the Task Force includes the following provisions:

- Consider the state of existing research relevant to the development of standards for clinical research on cLBP
- Conduct a comprehensive review of existing case definitions, diagnostic criteria, and outcome measures that are relevant for clinical research on cLBP
- Develop a draft set of standards for research on cLBP
- Engage the broader research community and representatives from relevant government agencies in developing these draft standards
- Chart a general plan for incorporation of the final standards into research studies and for their future modification.

Other Considerations

Although the primary focus of this effort is aimed at better understanding of, and management strategies for, the problem of idiopathic cLBP, the Task Force may need to consider research and approaches to other cLBP categories.

The process followed by the Task Force should be driven by evidence and focused primarily on the needs of research needed to move the field forward. One model is the Research Diagnostic Criteria (RDC) process, which has been used to advance research in a number of other complex clinical conditions (e.g., temporomandibular joint disorders, complex regional pain syndrome, premenstrual tension, interstitial cystitis syndrome, chronic fatigue syndrome, insomnia, Parkinson's disease, Alzheimer's disease, tardive dyskinesia, and vascular dementia).

It is expected that the Task Force will need to call on additional scientific or clinical expertise as its work unfolds.

Implementation

An initial face-to-face meeting of the Task Force will be held in Bethesda, Maryland, on March 5–6, 2012.

APPENDIX 2.2

Advancing Research on cLBP: Developing Standards for Research

March 5, 2012

The Chronic Low-Back Pain Research Task Force will address the broad issue of developing standards for research into chronic low-back pain (cLBP) by examining the research literature on this condition. The goal is to facilitate standardizing patient classification, and defining and measuring risk factors, prognostic factors, treatment outcomes, disease mechanisms, and other aspects of research design. Within this broad groundwork, an important focus will be to generate an initial set of Research Diagnostic Criteria (RDC) for the most common forms of chronic low-back pain (cLBP; this project will be referred to as RDC/cLBP).

The **research** nature of these criteria will be emphasized. In any form, they must currently be based as much on description of observable findings that cluster together as on underlying etiologic mechanisms. Their creation is viewed as a useful step toward placing the diagnosis of cLBP on a more scientific, and hence research-compatible, basis. The RDC are intended neither to replace the ICD-9 (or 10) coding system nor to be used for billing purposes, but rather to foster more consistent and comparable research studies and populations. The Task Force will provide guidance and develop materials to support this broad charge and disseminate its work in due process to the broadest possible array of stakeholders.

Goal and Scope of the RDC/cLBP Project

The long-term goal in creating these RDC/cLBP is to make available a scientifically-acceptable diagnostic and classification system that is reliable, valid, and useful in research. The scope of the project will encompass those cLBP conditions in adults for which information of sufficient reliability and validity exists to develop working case definitions, using physical examination, clinical laboratory tests, imaging, and self-report data-gathering procedures.

This project is envisioned as a start of a much longer process to include validation, evolution, and revalidation over time and solicitation of input from diverse stakeholders (e.g., patient advocacy groups, the Food and Drug Administration, pharmaceutical companies, medical device manufacturers, etc.). Its results will initially be published in an appropriate journal, and additional dissemination efforts implemented as well. NIH plans to sponsor future workshops/symposia, to include participation from diverse stakeholders, that will highlight studies validating or invalidating the initial RDC and propose next-generation iterations.

Overview of Methods

The methods used to derive the RDC/cLBP represent an advance over those currently available, as follows:

1. *An interdisciplinary effort:* The RDC/cLBP will represent the consensus of a team of recognized researchers whose areas of expertise range from basic biology to clinical and biobehavioral sciences.

2. *Broad-spectrum research*: The RDC/cLBP format should serve the needs of a wide spectrum of BP research, from epidemiologic studies gathering data by interview or medical records, to research in clinical settings (including treatment trials but also studies examining etiology, prevention, diagnosis and management of cLBP), to biomechanical and other basic-science studies.
3. *Use of epidemiologic data*: Existing epidemiologic and evidence-based clinical data will be used to guide the initial selection and operationalization of these RDC/cLBP. In addition, original data collection may be called for with the greatest reliability and validity.
4. *Operational definition of terms*: The RDC/cLBP will be stated in operational or measurable terms to maximize reproducibility among investigators, which will facilitate comparison of results.
5. *Specification of examination methods*: Detailed examination specifications will be provided, allowing clinical data associated with RDC/cLBP criteria to be gathered using standardized examination and interview methods.
6. *Reliability of measurement*: The reliability of clinical methods and measures will be established and will serve as the basis for selecting specific clinical measurement methods with established sensitivity and specificity.
7. *Multiaxis system*: Conceptually, an approach using at least two axes will be taken, to allow for an integrated biologic and psychosocial characterization of the back pain patient. For example, to assess the extent of pain, disability and dysfunction, one axis could consist of physical diagnoses based on objective clinical findings, and another axis could include one or more of the following:
 - Cognitive, emotional, and behavioral status, as evaluated through self-report and observation
 - Potential predictors of chronicity
 - Genetic and epigenetic factors
 - Central neuroprocessing.

APPENDIX 2.3

March 5, 2012

Some Questions to Consider for Meeting1

Case Definitions

When a clinician tells you a patient has chronic low-back pain (cLBP), what do you understand that to mean—that is, how many different categories or types would you need in order to include the most common presentations resulting in the label of cLBP?

1. What characterizes the presentation of the most common forms of back pain—what are the key clinical variables—the signs and symptoms?
 - Physical examination: limitations in movement, neurological deficits, evidence of radiculopathy?
 - Clinical tests: laboratory, imaging?
 - Observational and self-reported findings: psychological disability and/or psychosocial dysfunction, work disability, previous interventions?
 - Pain: intensity, temporal characteristics (chronic, fluctuating, intermittent), localization, response to movement, duration?
2. What recognizable patterns emerge by which clinicians cluster the signs and symptoms—that is, how are signs and symptoms currently organized into clinical constellations or syndromes that constitute a clinical case?
 - Can we define muscle and/or disk and/or joint/bony structure (e.g., stenosis) disorders? Discogenic pain? Radiculopathy? Instability?
3. Are the various clinical case types reliably assessed? Will different examiners come to the same conclusion?
4. Is there independent scientific evidence for the validity of case definitions? Do the case definitions or diagnostic criteria distinguish cases from non-cases and differentially distinguish one case type of cLBP from another? Do they predict treatment response or prognosis?
 - Does clinical examination and/or diagnostic testing distinguish cases of sprain, strain, spasm, facet joint pain, discogenic pain, or other conditions with evidence-based reliability and validity?
5. What are the most commonly used/highly regarded diagnostic systems clinicians rely on to arrive at a case definition or diagnosis? Examples might include sciatica or not, centralization of pain or not, discogenic pain or not, instability or not, spondylolisthesis or not, etc.
 - To what extent are these evidence-based or “scientific”?
 - What are the major shortcomings of the available diagnostic schemes that explain your interest in the Task Force?
6. Is chronicity, the “sine qua non” of cLBP, well enough understood?
 - What biologically plausible mechanisms could explain the emergence and maintenance of chronicity?

APPENDIX 2.4

Rationale for Developing a Biopsychosocial, Dual-Axis RDC/cLBPS Diagnostic and Classification System at the Present Time

1. The impetus for a biopsychosocial RDC/cLBPS derives from NIH Pain Consortium workshop conclusions that the present inability to categorize types of subjects included in published reports results in critical limitations in interpretation of findings, comparisons among studies, or replication of results.
2. Development of an RDC/cLBPS, if it could be accomplished for perhaps the two or three most common subtypes of cLBPS, would allow refinement of study design, reduce ambiguity in interpretation of published reports, and facilitate more reliable and valid independent replication of published findings.
3. The ability to develop a reliable and valid RDC/cLBPS requires the ability to derive operationally-defined cLBPS subtypes (evidence-based wherever possible) separate from those chronic low-back pain conditions excluded from cLBPS, such as disc herniation, spinal stenosis, spine instability, etc.,
4. If different treatments are thought more appropriate for some cLBPS patients than others, a reasonable inference is that clinicians have in mind differential diagnostic criteria for those subtypes that justify differentiation of treatment decisions.
5. Does the RTF consider that the most common cLBPS Axis I and/or Axis II subtypes can reliably be differentiated, at least descriptively, across the following domains (i.e., is there a commonly occurring musculoskeletal form of cLBPS that can be differentiated from a discogenic or neuropathic subtype; are there subtypes that incorporate both Axis I and Axis II domains)?

Axis I: Disease and disease progression	Axis II: Illness and illness progression
Musculoskeletal	Pain (e.g., intensity, location, duration)
Discogenic	Pain behaviors (e.g., splinting, bracing, limping, ROM)
Articular (bony joint components)	Level of psychosocial dysfunction (e.g., AOL, drug impairment)
Neuropathic	Psychological status (e.g., depression, anxiety, co-morbid pains, MUPS)
Rheumatoid (inflammatory disease)	Treatments (e.g., number and outcomes)
Central (brain) processing (e.g., central sensitization effects on peripheral and higher centers processes)	Compensation and litigation
Neurochemical (e.g., changes in prevalent	Other
Epigenetics (e.g., emergent gene variants)	
Other	

1. Alternatively, there may be some other schema known to the RTF that can reliably identify cLBPS subtypes, at least as an initial effort.
2. For the present, it is recognized that any attempt to develop RDC/cLBPS will be based on a descriptive system of signs and symptoms rather than on etiologic mechanisms of action.

3. Accordingly, emphasis is placed on *reliability* of putative diagnostic or classificatory criteria, based on physical examination findings and psychosocial measures of pain and pain-related disability that have known reliability and, to the extent possible, known validity.
4. It seems reasonable to expect that the first round of putative diagnostic criteria would be based on findings derived from the Axis I and Axis II minimal databases arrived at by the RTF.
5. A high priority for any research agenda that includes use of such an RDC/cLBPS must include scientific investigations into the reliability and validity of the RDC/cLBPS system itself, on an iterative basis so subsequent revisions are always evidence-based.

Rationale for *not* considering development of a biopsychosocial, dual-axis RDC/cLPBS diagnostic and classification system at the present time:

1. The scientific need and utility of an RDC/cLBPS is acknowledged, but the RTF consensus is that there is not sufficient evidence on which to base a diagnostic or classification of cLBPS; that is, subtypes of cLBPS cannot be reliably defined at present.
2. The RTF considers that the scientific need and utility of an RDC/cLBPS has not been adequately established.
3. The scope of work required is beyond the funding and personnel resources of the RTF at present.
4. A higher priority would be to specify a comprehensive NIH-funded research agenda that would encompass the research necessary to allow future development of an evidence-based RDC/cLBPS.

APPENDIX 2.5

First Meeting Overview and Charge to Workgroups

March 5, 2012

The purpose of the Task Force is to address the broad issue of developing standards for research into cLBP. Towards this end, we hope to improve the research literature on chronic back pain by helping to standardize patient classification, outcome measurement, and perhaps other aspects of research design. We will not finish the process by the end of this meeting and that should not be our expectation. Rather, it would be extremely rewarding if we were to make at least some progress on identifying the areas of greatest need (and greatest impact) and advance some major fronts. We imagine that these may include case definitions, identification of critical variables, agreeing on a common set of principles and perspectives to guide the work, considering multiple axes of classification, gathering reliable data, creating a diagnostic and classification scheme, integrating objective findings with self-report subjective data, and creating an operational definition of treatment success.

Progress in this regard will be uneven along those fronts but we will be better able to understand where we are and what has to be done to make further advances in the lagging areas. This meeting is the start of that process and how far we get along each front will be revealed at the end of the meeting. Much of the focus for this initial meeting will be on creating “research diagnostic criteria” for chronic back pain, but we also hope to flesh out a larger agenda.

Customary welcome:	Introductions, logistics of the day and business-end of the meeting
NIH review:	Events and factors leading to creation of the Task Force
Introduction of Drs. Deyo and Dworkin; group self-introductions	
Dr. Deyo:	Why are we here, and what are the problems we need to address? Emphasis on the need for case definitions of cLBP subtypes; issues in outcome assessment
Group discussion:	Preferred scope of Task Force activities; possible approaches
Dr. Dworkin:	Some critical factors and perspectives for developing Research Diagnostic Criteria for Chronic Low Back Pain (RDC/cLBP)
Group discussion:	Can consensus agreement be reached on the 6 factors presented (biopsychosocial model; interdisciplinary research; measurable operational definitions; reliability; suitable for population and clinical research; dual or multiple axes)?

Reviewing what we know:	Informing Work Groups about existing classification schemes; considering how to obtain needed data
Work groups:	Tell us what we know and don't know in order to develop RDC/cLBP criteria for Axis I (physical), Axis II (psychosocial); possible other critical factors not yet adequately integrated; development of outcome measures and definition of success
Group discussion:	Reports from the Work Groups on what we know, what we need to know, and how to get the needed data where possible
Summary	
An attempt at developing draft criteria:	We used small groups to develop critical operational criteria for Axis I diagnoses of major subtype and Axis II (psychological, behavioral, and psychosocial components of cLBP and treatment response). The subset of "special topics" may be incorporated into the Axis I and Axis II groups for creating draft RDC/cLBP diagnostic and classification criteria. We made similar efforts for outcome measures.
Integrating multiple axes:	Develop a model or approach that allows Axis I diagnoses to be interpreted or evaluated in the context of Axis II findings, with consideration of other factors of potential classification relevance, such as, epigenetics and central neuroprocessing data.
Summarizing progress and next steps:	To date, the Task Force has summarized and identified where important gaps remain. We are developing specific questions for literature synthesis that may be needed (perhaps as contract work). We are enrolling relevant critical stakeholders in the next steps.

Charge to Work Groups

Work Group:

Conducting Research on Assessment and Utility of Psychosocial Factors in Chronic Low-Back Pain

Dennis Turk, Ph.D., Group Leader

Monday, March 5, 2012

2:15-3:15 p.m.

Charge to the Axis II (Psychological/Psychosocial) Work Group

For each of the items listed below, please identify the most critical issues involved, the state-of-the-art of current knowledge and useful measures, and types of future research needed:

- Evaluate the evidence-based justification for including Axis II constructs and measures in chronic pain diagnostic systems. Consider also if case definitions of cLBP should include Axis components?
- Identify cLBP diagnostic systems that include Axis II variables or criteria and evaluate their scientific quality and comprehensiveness.
- Identify available databases and research studies suitable for mining of Axis II constructs and measures, especially identifying available electronic records and published databases/datasets (e.g., NHANES, etc.).
- How are Axis II assessment/classification measures related to Axis II outcome measures?
- Evaluate or suggest developing alternative testable approaches for integrating Axis I and Axis II finding with a view to identifying comprehensive and research-useful phenotype(s) of the cLBP patient.
- Identify the most relevant types of information and research designs to aid in further validating Axis II construct and measures.

FYI: The second Axis II Work Group session is on Tuesday, March 6 from 8:30-10:30 a.m., led by Deb Weiner. The second session will emphasize identifying which constructs, variables, and measures relevant to the domains of Axis II should/can be incorporated into an evidence-based system of research diagnostic criteria for cLBP, which the Task Force is evolving.

It's a tall order for both Work Groups, we know. We have no expectation that each of the tasks above will be carried to conclusion—that prospect remains the goal of the Task Force's continuing efforts. We do hope, however, the Work Group will be able to crystallize what is known and what needs to be known about how psychological and psychosocial factors influence the presentation, diagnosis, and management of cLBP.

**Work Group:
Conducting Research on Factors Related to Chronicity and Functional Impairment
in Chronic Low-Back Pain**

Michael Von Korff, Sc.D., Group Leader Monday, March 5, 2012
2:15-3:15 p.m.

- **Biologic, psychologic, social, and temporal components of chronicity**
- **Predictors of chronicity**

For each of the items listed below, please identify the most critical issues involved, the state-of-the-art of current knowledge and useful measures, and types of future research needed:

- What is the level of current scientific evidence to support the respective biologic, psychologic, and psychosocial contributions to initiating and maintaining chronicity of LBP?
- What is the level of current scientific evidence in support of identifying reliable and valid predictors of chronicity?
- Should chronicity be incorporated into the research definition of the subtypes of cLBP?
- Identify available databases and research studies suitable for mining the issue of chronicity.
- Identify the most relevant types of information and research designs to aid in further defining parameters of chronicity, especially predictors of chronicity.

FYI, the second Chronicity Work Group session is on Tuesday, March 6 from 8:30-10:30 a.m., led by Gunnar Andersson. The second session will emphasize identifying which constructs and measures of chronicity and its prediction should be incorporated into the system of research diagnostic criteria for cLBP, which the Task Force is evolving. The future work of the Task Force will be based at least in part on findings from all the Work Groups taking place at this initial RTF meeting.

It's a tall order, we know. We have no expectation that each of the tasks above will be carried to a satisfying conclusion—that prospect remains the goal of the Task Force's continuing efforts. We do hope, however, the Work Group will be able to crystallize what is known and what needs to be known regarding how factors influencing chronicity contribute so uniquely and pervasively to the diagnosis and management of LBP.

APPENDIX 2.6

Chronic Low-Back Pain: Classification; Research Needs and Recommendations

March 5, 2012

Classifying Chronic Low-Back Pain

Richard Deyo, M.D., M.P.H., and Michael Von Korff, Sc.D.

Prepared for the NIH Research Task Force on Standards of Research for Chronic Low-Back Pain

Chronicity—Low-back pain will be classified as chronic or nonchronic based on whether low-back pain has been present on at least half the days in the prior 6 months.¹ Prior research has found correlations of $r \geq 0.67$ between recall of pain days measured by daily diary over time periods of 2 weeks² and 3 months.³ Pain days has consistently been found to predict long-term pain outcomes for diverse pain conditions, including back pain.⁴⁻⁹ Requiring at least half the days in 6 months assures that all persons defined as having chronic pain have had at least 90 days of back pain. Persons with chronic (and nonchronic) back pain will be further differentiated by back pain impact (severity) and prognosis.

Impact—Low-back pain impact (severity) will be assessed by NIH PROMIS pain intensity and pain interference ratings. A PROMIS pain interference item will be rated using widely-used 0-10 interference ratings¹⁰⁻¹⁴ in place of PROMIS verbal descriptors. We propose new social role disability items to identify persons with major social role disability (e.g., work disability).

The proposed cut-points for differentiating mild, moderate, and severe low-back pain are based on research indicating that pain ratings ≤ 3 reflect mild pain/interference generally acceptable to patients; pain ratings of 4 to 6 reflect moderate pain/interference; and pain ratings of ≥ 7 reflect severe pain/interference.¹⁵⁻¹⁹ Persons endorsing one or more items indicating major social role disability will be classified as disabled.

Prognosis—Back pain prognosis will be assessed using a brief, validated set of prognostic screening items, drawing on items adapted from the STarTBack screening tool²⁰⁻³¹ and the Chronic Pain Risk Score.^{1,4-7,32}

Generalizability to other musculoskeletal pain conditions—Research suggests that the proposed classification of chronic low-back pain could be generalized to other musculoskeletal pain conditions.^{4,6,8,33}

Classification of Chronic Low-Back Pain

Chronicity	
a. Chronic	Back pain present at least half the days in the prior 6 months
b. Not chronic	Back pain present fewer than half the days in the prior 6 months
Impact	
a. Low impact	Sum of average pain intensity and pain-related interference <8
b. Moderate impact	Sum of average pain intensity and pain-related interference ≥ 8 and ≤ 12
c. Severe impact, not disabled	Sum of usual pain intensity and pain-related interference ≥ 13 , not disabled
d. Disabled	One or more indicators of social role disability: [Off work \geq one month, Receiving work disability benefits, Usually inactive, Usually Housebound, Unable to care for self (cooking, dressing, bathing) due to low back pain]

Prognosis¹

1. Low risk² of unfavorable outcome in 6+ months
2. Medium risk² of unfavorable outcome in 6+ months
3. High risk² of unfavorable outcome in 6+ months

1. *Unfavorable outcome = Impact levels 3 and 4 -OR- Unfavorable outcome = Impact levels 2-4. A narrow definition of an unfavorable outcome would be appropriate for more restrictive targeting of an intervention based on risk stratification.*
2. *Risk can be determined by a screening scale score, a predicted probability of an unfavorable outcome, or by an algorithm that classifies risk levels. The specific methods and criteria for classification risk level are to be determined.*

Minimum Item Set To Assess Back Pain Chronicity, Impact, and Prognosis

Chronicity assessment

1. Have you had low-back pain on at least half the days in the past 6 months?

- Yes
- No

Impact assessment

2. In the past 7 days, how would you rate your low-back pain on average? [NIH PROMIS item]

- | | | | | | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|---|
| <input type="checkbox"/> |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 100 |
| No pain | | | | | | | | | Worst Imaginable Pain
(Alternative: Pain as bad as could be) |

3. In the past 7 days, how much did low-back pain interfere with your day-to-day activities?
[NIH PROMIS item with Brief Pain Inventory 0-10 rating. PROMIS rating is: not at all, a little bit, somewhat, quite a bit, very much]

- | | | | | | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Does not interfere | | | | | | | | | Completely interferes |

[New items to identify major social role disability. Agree/disagree responses are used to make responding to these items as simple as possible; also matches the STarT Back format]

4. Over the past month, **I have been resting or lying down most of the time** due to low-back pain.

- Agree
- Disagree

5. Over the past month, **I have rarely gotten out of the house to shop or do things with family or friends** due to low-back pain.

- Agree
- Disagree

6. Over the past month, **I have been unable to take care of myself**, including cooking, dressing, bathing, and other self-care activities, due to low-back pain.

- Agree
- Disagree

7. **I have been off work or unemployed for 1 month or more** due to low-back pain
- Agree
 - Disagree
 - Does not apply
8. **I receive or have applied for disability or workers' compensation benefits because I am unable to work** due to low-back pain.
- Agree
 - Disagree
 - Does not apply
9. Please check each location where you have had pain in the past week. Check each pain that applies to you:
- Back pain
 - Neck pain
 - Headache
 - Facial or jaw pain
 - Stomach pain
 - Chest pain
 - Genital/urinary pain
 - Pelvis or groin pain
 - Hip or buttocks pain
 - Arm pain
 - Hand pain
 - Shoulder pain
 - Leg or knee pain
 - Foot pain
 - Pain all over your body
 - Other pain site

Additional items to assess prognosis:

[Q10-16: Modified STarT Back items]

Thinking about the past 2 weeks, check whether you agree or disagree that the following statements apply to you.

- *10. My back pain has **spread down my leg(s)** during the past 2 weeks.
- Agree
 - Disagree

11. I have only **walked short distances** because of my back problem.
- Agree
- Disagree
12. In the past 2 weeks, I have **dressed more slowly** than usual because of my back pain.
- Agree
- Disagree
- *13. It's **not really safe** for a person with my back problem to be physically active.
- Agree
- Disagree
- *14. **Worrying thoughts** about my back condition have been going through my mind a lot of the time.
- Agree
- Disagree
15. I feel that **my back pain is terrible** and **it's never going to get any better**.
- Agree
- Disagree
16. In general, I have **not enjoyed** all the things I used to enjoy.
- Agree
- Disagree

**Question wording slightly modified from the original STarT Back items*

Discussion Points: Brief Assessment

1. The NIH PROMIS pain intensity and interference ratings are proposed to measure impact, substituting a widely-used 0-10 rating for the pain interference item in place of the NIH PROMIS verbal descriptor scale. The 0-10 rating is more widely used, and facilitates summary scoring with the NIH PROMIS 0-10 pain intensity rating.
2. Items 10-16 are from the STarT Back screening tool with minor modifications. STarT Back items assessing back pain bothersomeness and neck/shoulder pain were omitted from this brief assessment. Back pain bothersomeness is adequately assessed by items 2 and 3, and neck/shoulder pain is adequately assessed by item 9.
3. Items 11 and 12 from the STarT Back assess interference with activities that may be adequately assessed by the 0-10 interference rating (Item 3). These items could potentially be omitted.
4. Items 1-3 and 9 are similar to items found to predict long-term pain outcomes in studies using the Chronic Pain Risk Score.
5. An alternative wording for Item 1 would be: *Have you had back pain more than half the time in the past 6 months?* However, we are unaware of experience with this question wording, whereas questions about back pain days have been widely used and validity assessed.
6. Question 9 is included to assess diffuse pain. Diffuse pain has prognostic value comparable to pain severity and pain persistence.^{4,8,36}
7. Questions 4-8 are included to identify persons with substantial social role disability due to low-back pain. Items 7-8 assess work role disability. Items 4-6 assess social role disability appropriate for persons who are not in the paid labor force, (e.g., retired persons and homemakers). Pain interference ratings alone do not adequately identify persons with substantial social role disability due to back pain.

Research Needs and Recommendations

Impact and Prognosis Measures:

Based on available prognostic studies, the proposed set of items should yield prediction of unfavorable low-back pain outcomes (yes or no) with an area under the Receiver Operating Characteristic Curve in the 0.72 to 0.82 range. Existing research suggests that the brief screening assessment can detect back pain patients who will have an unfavorable future outcome (e.g., Impact levels 2-4) with reasonable specificity (80 percent or greater) but with lower sensitivity (50 percent to 65 percent).^{34, 35}

1. Further research is needed to estimate prediction models for risk stratification from the brief assessment. Although there are existing risk stratification methods (e.g., from the STarT Back or the Chronic Pain Risk Score), developing new risk models will likely be necessary. In prior research, risk of an unfavorable outcome has been quantified by a scale score,¹ by probabilities predicted from a regression model,⁴ or by a risk stratification algorithm.²⁰
2. The Chronic Pain Risk Score has been evaluated in populations including older persons, but outcome prediction with the STarT Back has not been evaluated among older back pain patients. Research evaluating prediction of back pain outcomes and risk stratification among the elderly is needed.
3. An unfavorable outcome could be defined by Impact levels 2-4, including persons with moderate pain intensity and/or interference with activities, or by Impact levels 3-4, focusing on

highly unfavorable outcomes. Prediction of severe outcomes might be useful for targeting high cost interventions (e.g., multidisciplinary rehabilitation services).

4. Further research is needed to determine whether a single prediction model can be used for diverse patient populations (e.g., general population, primary care, surgical patients, rehabilitation patients), or whether prediction models need to be population specific.
5. The proposed approach to classification and assessment of chronic pain might be applied to chronic musculoskeletal pain in general, not only low back pain.^{4, 6, 8, 33} Further research would be necessary.

Minimal Database:

We anticipate that comparing chronic back pain patients from different settings (e.g., primary care, surgical patients, or patients seeking complementary health care) will help to determine whether the data elements included in the minimal dataset have useful discriminative properties. Similarly, it may be valuable to compare characteristics of patients with varied degenerative spinal diagnoses, e.g., herniated disc, spinal stenosis, spondylolisthesis, or degenerative disc disease.

It will be valuable to compare some components of the minimal database with legacy measures that are not incorporated. For example, how do measures from the minimal dataset compare with scores from the Roland-Morris Disability Questionnaire, the Oswestry Disability Questionnaire, or the SF-36?

We anticipate that as more genetic markers of predisposition to chronic pain are identified, it will be important to determine if the minimal dataset discriminates among patients with varied genetic risk factors.

Outcome Measures:

Although we recommend a definition of “successful outcome,” we recognize that this is a relatively new and untested concept. There is a need to test the implications of our proposed definition and alternative definitions in terms of construct validity (e.g., association with return to work or use of opioids) and responsiveness to change.

References

1. Von Korff M. Assessment of chronic pain in epidemiological and health services research: empirical bases and new directions. In: Turk DC, Melzack R, eds. *Handbook of Pain Assessment*. 3rd ed. New York, NY: Guilford Press; 2011: 455-473.
2. Salovey P, Seiber WJ, Smith AF, et al. *Reporting chronic pain episodes on health surveys*. Washington DC: Centers for Disease Control, National Center for Health Statistics; 1992. HHS publication no. (PHS) 92-1081.
3. Stewart WF, Lipton RB, Simon D, et al. Validity of an illness severity measure for headache in an population sample of migraine sufferers. *Pain*. 1999;79(2-3):291-301.
4. Von Korff M, Dunn KM. Chronic pain reconsidered. *Pain*. 2008;138(2):267-276.
5. Von Korff M, Miglioretti DL. A prognostic approach to defining chronic pain. *Pain*. 2005;117(3):304-313.
6. 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(1-2):48-54.

7. 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(2):389-397.
8. Mallen CD, Peat G, Thomas E, et al. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *British Journal of General Practice*. 2007;57(541):655-661.
9. Friedman BW, Mulvey L, Davitt M, et al. Predicting 7-day and 3-month functional outcomes after an ED visit for acute nontraumatic low back pain. *American Journal of Emergency Medicine*. 2012;30(9):1852-1859.
10. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113(1-2):9-19.
11. Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2009;146(3):238-244.
12. Von Korff M, Ormel J, Keefe FJ, et al. Grading the severity of chronic pain. *Pain*. 1992;50(2):133-149.
13. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Annals of Academic Medicine*. 1994;23(2):129-138.
14. Krebs EE, Lorenz KA, Bair MJ, et al. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. *Journal of General Internal Medicine*. 2009;24(6):733-738.
15. Edeleen MO, Saliba D. Correspondence of verbal descriptor and numeric rating scales for pain intensity: an item response theory calibration. *Journals of Gerontology A: Biological Sciences and Medical Sciences*. 2010;65(7):778-785.
16. Jensen MP, Smith DG, Ehde DM, et al. Pain site and the effects of amputation pain: further clarification of the meaning of mild, moderate, and severe pain. *Pain*. 2001;91(3):317-322.
17. Kapstad H, Hanestad BR, Langeland N, et al. Cutpoints for mild, moderate and severe pain in patients with osteoarthritis of the hip or knee ready for joint replacement surgery. *BMC Musculoskeletal Disorders*. 2008;9:55.
18. Gerbershagen HJ, Rothaug J, Kalkman CJ, et al. Determination of moderate-to-severe postoperative pain on the numeric rating scale: a cut-off point analysis applying four different methods. *British Journal of Anaesthesia*. 2011;107(4):619-626.
19. Jones KR, Vojir CP, Hutt E, et al. Determining mild, moderate, and severe pain equivalency across pain-intensity tools in nursing home residents. *Journal of Rehabilitation Research and Development*. 2007;44(2):305-314.
20. 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(5):632-641.
21. Hill JC, Dunn KM, Main CJ, et al. Subgrouping low back pain: a comparison of the STarT Back Tool with the Orebro Musculoskeletal Pain Screening Questionnaire. *European Journal of Pain*. 2010;14(1):83-89.
22. Hill JC, Vohora K, Dunn KM, et al. Comparing the STarT back screening tool's subgroup allocation of individual patients with that of independent clinical experts. *Clinical Journal of Pain*. 2010;26(9):783-787.
23. Hill JC, Whitehurst DG, 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*. 2011;378(9802):1560-1571.
24. Beneciuk JM, Bishop MD, Fritz JM, et al. The STarT Back Screening Tool and individual psychological measures: evaluation of prognostic capabilities for low back pain clinical outcomes in outpatient physical therapy Settings. *Physical Therapy*. 2013;93(3):321-333.

25. Bruyere O, Demoulin M, Brereton C, et al. Translation validation of a new back pain screening questionnaire (the STarT Back Screening Tool) in French. *Archives of Public Health*. 2012;70(1):12.
26. Field J, Newell D. Relationship between STarT Back Screening Tool and prognosis for low back pain patients receiving spinal manipulative therapy. *Chiropractic & Manual Therapies*. 2012;20(1):17.
27. 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(5):722-732.
28. Gusi N, del Pozo-Cruz B, Olivares PR, et al. The Spanish version of the "STarT Back Screening Tool" (SBST) in different subgroups. *Atencion Primaria*. 2011;43(7):356-361.
29. Kongsted A, Johannesen E, Leboeuf-Yde C. Feasibility of the STarT back screening tool in chiropractic clinics: a cross-sectional study of patients with low back pain. *Chiropractic & Manual Therapies*. 2011;19:10.
30. Morso L, Albert H, Kent P, et al. Translation and discriminative validation of the STarT Back Screening Tool into Danish. *European Spine Journal*. 2011;20(12):2166-2173.
31. Wideman TH, Hill JC, Main CJ, et al. Comparing the responsiveness of a brief, multidimensional risk screening tool for back pain to its unidimensional reference standards: the whole is greater than the sum of its parts. *Pain*. 2012;153(11):2182-2191.
32. Dunn KM, Von Korff M, Croft P. Defining chronic pain by prognosis. In: Hasenbring MI, Rusu AC, Turk DC, eds. *From Acute to Chronic Back Pain: Risk Factors, Mechanisms, and Clinical Implications*. 1st ed. New York, NY: Oxford University Press; 2012:21-40.
33. Martell BA, O'Connor PG, Kerns RD, et al. Systematic review: opioid treatment for chronic back pain: prevalence, efficacy, and association with addiction. *Annals of Internal Medicine*. 2007;146(2):116-127.
34. Von Korff M, Shortreed S, Saunders KW, et al. Comparison of prognostic screeners for back pain risk stratification. (Under review).
35. Turner JA, Shortreed SM, Saunders KW, et al. Optimizing prediction of back pain outcomes. *Pain*. 2013;154(8):1391-1401.
36. Muller S, Thomas E, Dunn KM, et al. A prognostic approach to defining chronic pain across a range of musculoskeletal conditions. *Clinical Journal of Pain*. (In press).

APPENDIX 2.7

Overview of Minimal Dataset for Research on cLBP

Part 1. Demographic and Social Status

Part 2. Medical History and Physical Examination

Part 3. Pain Report, Behavioral, Emotional, and Psychosocial Domains

Part 4. Additional Measures

Part 5. Outcomes Measures

Objective: To recommend to NIH a minimal multidimensional dataset of baseline descriptors and outcomes measures that assess the major biologic, psychologic, and psychosocial domains influencing the expression of cLBP for inclusion in NIH-funded cLBP research.

The salient aspects of the integrated baseline and outcomes minimal dataset are as follows:

- As described in the preamble to our proposed recommendations to NIH, the minimal dataset comprises an important subset of recommendations to NIH that have the common objective to recommend Standards for Research into cLBP.
- Intended users are NIH-funded researchers but it is available to all researchers.
- To the extent that current scientific knowledge allows, the multidimensional minimal dataset:
 - Encompasses demographic, patient history, biologic, psychologic, and psychosocial domains in order to elaborate cLBP phenotypes and critical cLBP research outcomes
 - Serves as a core set of basic baseline and outcomes measurements useful for all types of cLBP clinical research (i.e., epidemiology, observational, RCTs)
 - Encourages researchers to expand their cLBP inquiries beyond the recommended minimal dataset of baseline and outcomes measures
 - Comprises measures sensitive to changes over time and differences between study populations
 - Contains evidence-based measures and scales of established reliability, validity, and demonstrated clinical research utility.

Please carefully review the enclosed draft minimal dataset to insure it reflects the intentions of the RTF as gathered during our prior meetings. You are strongly encouraged to direct questions, comments, and critiques concerning any aspect of this component of the minimal data set to either or both of us. At the upcoming RTF meeting, March 11–12, 2013, the suggested minimal dataset will be presented for discussion and hopefully, final approval.

Once again, thank you for all your work!

Cordially,

Rick Deyo
Sam Dworkin

Creating The Multidimensional Minimal Dataset For cLBP Research

Objective: To create a single minimal dataset paper-pencil measure suitable for entry by research subjects

Part 1.

Please insert all six items on attached table (include notes) from Deyo labeled: “Demographics and Social Status (all but item 9 directly from NINDS Common Data Elements).”

Part 2.

Please insert all 10 items on attached table (include notes) from Deyo labeled “Medical History and Physical Examination.”

Part 3. Minimum Item Set To Assess Back Pain Chronicity, Impact, and Prognosis

Please insert items 1-16 (Von Korff) and items 17-37, which follow:

Chronicity assessment

1. Have you had low-back pain on at least half the days in the past 6 months?

- Yes
- No

Impact assessment

2. In the past 7 days, how would you rate your low-back pain on average? *[NIH PROMIS item]*

- | | | | | | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|---|
| <input type="checkbox"/> |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| No pain | | | | | | | | | Worst Imaginable Pain
(Alternative: Pain as bad as could be) |

3. In the past 7 days, how much did low-back pain interfere with your day-to-day activities?

[NIH PROMIS item with Brief Pain Inventory 0-10 rating. PROMIS rating is: not at all, a little bit, somewhat, quite a bit, very much]

- | | | | | | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Does not interfere | | | | | | | | | Completely interferes |

[New items to identify major social role disability. Agree/disagree responses are used to make responding to these items as simple as possible.]

4. Over the past month, **I have been resting or lying down most of the time** due to low-back pain.
- Agree
 Disagree
5. Over the past month, **I have rarely gotten out of the house to shop or do things with family or friends** due to low-back pain.
- Agree
 Disagree
6. Over the past month, **I have been unable to take care of myself**, including cooking, dressing, bathing, and other self-care activities, due to low-back pain.
- Agree
 Disagree
7. **I have been off work or unemployed for 1 month or more** due to low-back pain.
- Agree
 Disagree
 Does not apply
8. **I receive or have applied for disability or workers' compensation benefits because I am unable to work** due to low-back pain.
- Agree
 Disagree
 Does not apply

Additional items to assess prognosis

9. Please check each location where you have had pain in the past week. Check each pain that applies to you:
- Back pain
 Neck pain
 Headache
 Facial or jaw pain
 Stomach pain
 Chest pain
 Genital/urinary pain
 Pelvis or groin pain
 Hip or buttocks pain
 Arm pain
 Hand pain
 Shoulder pain
 Leg or knee pain
 Foot pain
 Pain all over your body
 Other pain site

[Q10-16: Modified STarT Back items]

Thinking about the past 2 weeks, check whether you agree or disagree that the following statements apply to you.

*10. My back pain has **spread down my leg(s)** during the past 2 weeks.

- Agree
- Disagree

11. I have only **walked short distances** because of my back problem.

- Agree
- Disagree

12. In the past 2 weeks, I have **dressed more slowly** than usual because of my back pain.

- Agree
- Disagree

*13. It's **not really safe** for a person with my back problem to be physically active.

- Agree
- Disagree

*14. **Worrying thoughts** about my back condition have been going through my mind a lot of the time.

- Agree
- Disagree

15. I feel that **my back pain is terrible** and **it's never going to get any better**.

- Agree
- Disagree

16. In general, I have **not enjoyed** all the things I used to enjoy.

- Agree
- Disagree

**Question wording slightly modified from the original STarT Back items*

Part 3 (continued). Minimum Item Set To Assess Physical Function (Sleep and Fatigue), Emotional Function (Depression, Anxiety, Substance Abuse), and Social Role (Satisfaction With Social Role)

Physical Function:

For sleep, please insert PROMIS 29 item 17.

For fatigue, please insert PROMIS 29 item 16.

Emotional Function:

19-22. For anxiety, please insert PROMIS 29 items 5-8.

23-26. For depression, please insert PROMIS 29 items 9-12.

27-30. For substance abuse, please insert: 4 items from C-Disk Screener or Opioid Assessment for patients with pain. TBD.”

Social Role:

31-34. For satisfaction with social role, please insert PROMIS 29 items 21-24.

Part 4. Outcomes Measures

35. For global impression of change, please insert the single item.

36. Roland-Morris or Oswestry (insert only the names of the two scales)

37. TBD

Part 5: Additional Items

Additional Items for Part 2:

Recommended Minimal Physical Examination and Imaging Measures for Specific Types of Research

- A. Recommended for studies of surgical intervention, injections, or ablative therapy:
 1. Straight leg raising if patient reports leg pain
 2. Lower extremity strength
 3. Lower extremity reflexes
 4. MRI results

- B. Recommended for studies of adults over age 65
 5. Tests of hip arthritis

Additional Items for Part 3:

Domain	Clinical Variable	Minimal Additional Dataset Measures (Note: enter only name of scale, not individual items), plus:
	Intensity	BPI
	Pattern	BPI
	Comorbidities	PHQ-15 non-pain 9 symptom item

Physical Function	Physical Disability	Roland or Oswestry (if not included as Outcome measure
		Roland or Oswestry if not included in Outcomes measures
		Ambul. Activity: Activity diary
	Interference	PROMIS 29
	Pain Behaviors	PROMIS 29
	Fear of Movement	Fear of Activity Beliefs or TSK Scales
Emotional Function	Depression	HADS
	Anxiety	HADS
	Catastrophize	Pain Catastrophizing Scale (TBD)
		Coping Strategies subscale (TBD)
	PTSD	PTSD checklist-(TBD)
Subst./Med. Abuse	SOAPP-R	
Quality of Life	HRQoL	EQ-5D-5L
	Satisf. Soc. Role	PROMIS 29
	Work/Injury Hx.	1. Disab. Comp:Patient Hx
		2. Litig: Patient Hx

Additional Items for Part 4: Outcomes Measures

TBD

APPENDIX 2.8

Preamble and Recommendations of the NIH Task Force on Research Standards for Chronic Low-Back Pain

March 5, 2012

Introduction

This Task Force was convened under the auspices of the NIH Pain Consortium, with the goal of creating standards for terminology, classification, data collection, and outcome assessment in research concerning patients with chronic low-back pain. The intent was emphatically not to create new standards for clinical care, coding, or billing purposes, but to bring greater consistency to clinical research on a challenging patient population.

The rationale for this focus was a concern widely expressed at NIH conferences, symposia, and workshops that clinical research on chronic low-back pain suffers from poorly defined patient populations, inconsistent terminology and definitions, and variable outcome measures. As a result, it is difficult to compare studies of similar or competing interventions, replicate findings, pool data from multiple studies, resolve conflicting conclusions, develop multidisciplinary consensus, or even achieve consensus within a single discipline regarding interpretation of findings.

Background

Examples of inconsistencies in terminology and measurement in back pain research abound. There is little consensus on operational definitions of terms such as spine instability, spondylolisthesis, sciatica, or radiculopathy.¹ Researchers define diseases like spinal stenosis differently in nearly every study.² Outcome measures range from radiographic findings to self-reported pain, and are often only loosely related to daily functioning or work disability. Studies report trivial improvements as “statistically significant” with little attention to recognizable impact on quality of life.

As a practical matter, population surveys indicate that, over the past decade, patients with back pain report increasing, rather than decreasing, levels of functional limitation and work disability.³ These findings suggest that the field of back pain research has not achieved the scientific progress necessary for consistent improvement in patient care. Researchers need strategies that are more consistent for defining patient samples, patient characteristics, and outcomes.

Task Force Approach

The Task Force felt the first need was to provide a practical definition of what some have labeled “non-specific” back pain: a large public health problem with a major impact. The intent was to exclude well-defined conditions such as spinal infections, cancer, and ankylosing spondylitis that can be associated with chronic low-back pain but whose treatment involves highly specialized interventions.

Some other conditions (largely degenerative changes) such as herniated disc with radiculopathy, spinal stenosis, and spondylolisthesis are associated with specific pathoanatomic changes, but the definitions of these conditions are variable and sometimes ambiguous; the conditions may coexist; patients may be labeled differently by different practitioners; and a given patient might move from a less specific to a more specific diagnosis over time (or vice versa). Furthermore, specific treatments may fail to relieve pain. Thus, many patients with these degenerative conditions may become part of the larger group with chronic low-back pain.

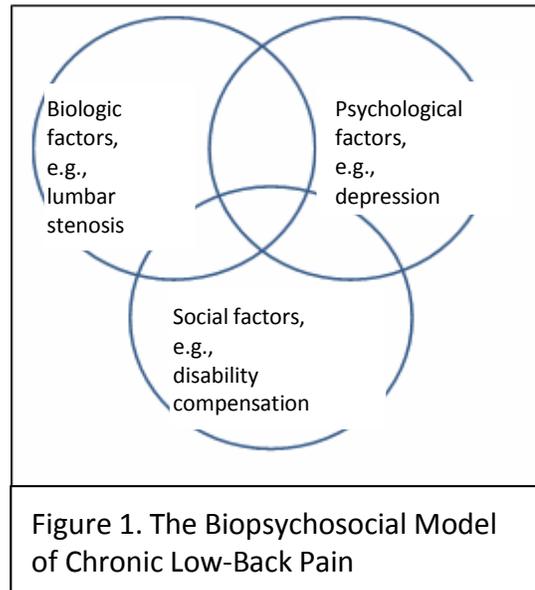


Figure 1. The Biopsychosocial Model of Chronic Low-Back Pain

Chronic pain is a biopsychosocial phenomenon and chronic low-back pain (cLBP) is no exception, as illustrated in Figure 1.⁴ We intend that the term *chronic low-back pain*, a symptom rather than a diagnosis, be used to reflect current uncertainty about etiology in an individual patient. We acknowledge the likelihood that some specific pathoanatomic or pathophysiologic conditions may eventually be differentiated within this group.

Although many researchers long for etiologic subclassification within this large group, Task Force members did not believe evidence-based etiologic subclassification was possible with the current state of knowledge. Instead, we proposed two strategies that—along with a basic definition of cLBP—would allow more consistent “phenotyping” of research subjects: (1) subclassifying those with cLBP according to functional impact and prognosis for functional recovery, and (2) requiring that research reports include a minimum standard dataset of baseline information. We also proposed a more consistent set of outcome measures.

Our hope is that the dataset and prognostic classification will be incorporated into intervention trials, cohort studies, epidemiologic surveys, diagnostic test evaluations, studies of etiologic risk factors, and a wide range of other studies. Furthermore, we hope the dataset and the prognostic classification will be used for studying cLBP of any etiology (e.g., stenosis, herniated disc), acknowledging that cLBP of any etiology is associated with biological, psychological, and social contributors.

Research Agenda: The Task Force proposes a research agenda to address gaps in our scientific knowledge about the impact, prognosis, phenotypic, and outcome measures proposed here.

References

1. Bogduk N. On the definitions and physiology of back pain, referred pain, and radicular pain. *Pain*. 2009;147(1-3):17-19.
2. Genevay S, Atlas SJ, Katz JN. Variation in eligibility criteria from studies of radiculopathy due to a herniated disc and of neurogenic claudication due to lumbar spinal stenosis: a structured literature review. *Spine*. 2010;35(7):803-811.
3. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA*. 2008;299(6):656-664.
4. Waddell G. 1987 Volvo Award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine*. 1987;12(7):632-644.

APPENDIX 3

Surveying the RTF: Survey Data on Preferences and Actions

APPENDIX 3.1

Prioritizing Minimal Dataset Items: History, Demographics, Physical Findings

Survey compiled and summarized by Dr. Richard Deyo

Axis I Survey Summary

Axis I Variables: The Top 20, by Category

(red = wide SD [>2.7], less consensus)

Demographics	Comorbidity	Physical Exam	Lab, Imaging	Current and Past Rx
Age, sex, race	Other pain sites	SLR if leg pain		Prev. fusion
Disability comp.	BMI	Hip rotation		Current opioid use
Legal claim	Comorbidity score	LE strength		Previous decompression
Employment status	Hx. Substance abuse			Exercise Rx
Education				Prev. injections
Smoking status				Use of PT
				Use of muscle relaxants

Axis I Variables: The Top 30 by Category

Demographics	Comorbidity	Physical Exam	Lab, Imaging	Current and Past Rx
Age, sex, race	Other pain sites	SLR if leg pain		Prev. fusion
Disability comp.	BMI	Hip rotation		Current opioid use
Legal claim	Comorbidity score	LE strength		Previous decompression
Employment status	Hx. Substance abuse	Other tests, hip arthritis		Exercise Rx
Education	Specific Comorbid dx's	LE reflexes		Prev. injections
Smoking status		Tender points		Use of muscle relaxants
				Past opioid use

				Use of antidepressants
				Use of NSAIDs
				Spinal manipulation
				Use of anticonvulsants
				Other complementary health approaches

	ALL ITEMS (n of responses)	Category	Mean Score	SD
1	Age, sex, race (14)	Demographics	1.86	1.23
2	Other pain sites (14)	Comorbidity	1.93	1.27
3	Previous fusion (14)	Past, current Rx	2.07	1.33
4	Current opioid use (14)	Past, current Rx	2.14	1.29
5	Previous decomp. (14)	Past, current Rx	2.21	1.72
6	Disability comp. (14)	Demographics	2.79	1.63
7	Legal claim (14)	Demographics	2.86	1.61
8	Employment status (14)	Demographics	3.00	1.62
9	BMI (14)	Comorbidity	3.14	1.51
10	Education (14)	Demographics	3.14	1.79
11	SLR if leg pain (13)	Physical exam	3.15	1.46
12	Comorbidity score (13)	Comorbidity	3.23	1.42
13	Exercise therapy (14)	Past, current Rx	3.36	2.06
14	Hx. substance abuse (14)	Comorbidity	3.5	1.99
15	Smoking status (14)	Demographics	3.64	1.34
16	Hip rotation (12)	Physical exam	3.92	2.35

17	Previous injections (14)	Current, past Rx	3.93	2.76
18	Use of PT (14)	Current, past Rx	3.93	2.76
19	LE strength (12)	Physical exam	4.00	2.37
20	Use of muscle relaxants (14)	Current, past Rx	4.07	2.97
21	Other tests, hip arthritis (12)	Physical exam	4.08	2.19
22	Past opioid use (14)	Current, past Rx	4.14	2.96
23	Use of antidepressants (14)	Current, past Rx	4.21	2.83
24	Use of NSAIDs (14)	Current, past Rx	4.21	2.94
25	Spinal manipulation (14)	Current, past Rx	4.29	2.49
26	Use of anticonvulsants (14)	Current, past Rx	4.36	2.90
27	Specific comorbid dx's (12)	Comorbidity	4.58	1.88
28	LE reflexes (12)	Physical exam	4.58	2.47
29	Other complementary health approaches (14)	Current, past Rx	4.64	2.68
30	Tender points (14)	Physical exam	4.93	2.46
31	Sensory exam (12)	Physical exam	5.17	2.76
32	Walking distance, durat. (13)	Physical exam	5.31	1.93
33	MRI ((13)	Lab	5.92	2.69
34	Walking speed (13)	Physical exam	5.92	2.84
35	Inflammatory markers (12)	Lab	6.08	2.57
36	Activity monitor (13)	Physical exam	6.08	2.36
37	Spine ROM (12)	Physical exam	6.25	2.45
38	Balance (13)	Physical exam	6.31	2.18
39	CT results (13)	Lab	6.54	2.11
40	Lifting, back strength (12)	Physical exam	6.58	2.02
41	Plain x-ray (13)	Lab	6.62	2.36
42	SLR if no leg pain (12)	Physical exam	6.75	2.90
43	Manual assessment (13)	Physical exam	6.77	2.45

44	Genomics (12)	Lab	6.83	2.59
45	HLA-B27 (12)	Lab	7.25	2.22
46	EMG/NCV (12)	Lab	7.58	1.78
47	Brain neuroimaging (13)	Lab	8.23	1.92
	“OTHER” NOMINATIONS: each with one nomination:			
	Heat/cold pain thresholds; quantitative sensory testing	Lab		
	Psychological therapies	Past, current Rx		
	Cognitive function, older adults	Physical exam		
	Walking program	Past, current Rx		
	TENS therapy	Past, Current Rx		

High Points of Axis I Survey Results
(See tables of top 20 items, top 30, all items)

- No lab or imaging findings were in the top 30 (out of 47) items.
- All of what I called “demographics” ranked high. Some overlap with Axis II items, and where we put them seems largely a matter of semantics (e.g., education, employment, disability compensation).
- Many items of current and past therapy rated high, but some were also the most variable in ratings. The ones in red are those with a Standard Deviation greater than an arbitrary 2.7, suggesting greater disagreements in rankings than other items.

Some reflections: I think the rankings are not too surprising, and generally make sense. It seems to be reasonable not to expect imaging results to be reported in every study of cLBP. Undoubtedly, surgical studies will be more likely to include imaging, and we could make that expectation explicit. A question in my mind is whether we expect every study (even, say, studies of acupuncture or massage) to have physical exam items like straight lower extremity strength. And how big is a reasonable “minimal dataset”? Whatever items we finally recommend will require careful specification of the exact questions, response options, or techniques. These are all issues for discussion.

I wonder if we should add an item on source of patients (e.g., primary care clinics, specialty clinic, general public, physical therapy clinic, chiropractic practice, etc.).

To get a crude sense of what using even the top 20 variables could mean for improving the literature, I searched four random issues of *Spine*, 2010, for clinical articles on adult patients (excluding animal studies, anatomic studies, radiographic studies, review articles, trauma, instrument development, or large database studies). I found eight eligible studies. Here's what they reported:

- Age and sex: all 8
- Race: 1
- Compensation status: 1
- Employment status: 1
- Smoking status: 2
- BMI: 7
- Comorbidity measure of some sort: 3 (some were just by exclusion)
- SLR: 1
- Previous decompression: 3
- Previous fusion surgery: 4 (all by exclusion)

Other pain sites; legal claims; history of substance abuse; current opioid use; current exercise therapy; previous injections; use of PT, muscle relaxants, antidepressants; past opioid use: NONE

(The single best study was the only one that was NIH-funded—a SPORT followup.)

APPENDIX 3.2

Prioritizing Minimal Dataset Items: Psychological, Behavioral, and Psychosocial

Prioritizing Minimal Dataset Items: Psychological, Behavioral, and Psychosocial

RTF Axis II Survey Response: N, Range, Mode, Mean (Ranked Within Groups), SD								
Domains	Measures	Scales	N	High	Low	Mode	Mean (Ranked)	SD
Pain	Intensity	Numerical 0-10 Rating	11	1	2	1	1.09	0.30
		PROMIS29	9	1	5	5	3.11	1.48
		BPI: Pain Scale	10	2	9	3	3.80	2.10
	Characteristics	Duration: Self-Report	9	1	10	4	5.00	2.83
		Location Pain Draw	11	1	5	1	1.91	1.21
		SF-MPQ-2-Constant, Intermittent	11	1	10	1	3.09	2.93
Chronicity	TBD	3	1	2	2	1.67	0.58	
Physical function includes sleep	Disability	Roland-Morris (RMDS)	10	1	6	1	2.20	1.52
		Oswestry (ODI)	11	1	6	1	2.45	1.79
		Phys. Func: PROMIS29	9	1	5	5	3.00	1.50
		Sleep: PROMIS29	9	1	7	5	3.67	2.18
		Fatigue: PROMIS29	9	1	7	5	3.44	2.09
	Ambulatory/ Activity	Active Diary	11	4	10	4	6.36	2.42
		Ambul. Monitor/ Actigraph	11	2	10	5	6.27	2.37
	Pain Behaviors	PROMIS Item Bbank	9	1	6	1	3.44	2.40
		Pain Behavior Checklist	11	2	10	4/10	6.09	2.37
		Observational Scoring	11	3	10	7	6.55	2.17
	Fear of Movement	Tampa Scale of Kines.	10	1	8	1/5	3.27	1.75
		Fear of Activity Beliefs	11	1	6	4	3.90	2.79
Somatic Symptoms	PHQ-15	10	1	8	5	3.60	2.53	
Psychol. status	Depression	PHQ-9	10	1	5	1	2.80	1.55
		PROMIS29	10	1	5	1/2	2.80	2.20
		HADS	10	1	5	4/5	3.80	1.23
		BDI	11	2	9	6	4.64	2.06
		CES-D	11	1	9	4	4.18	2.14
	Anxiety	PROMIS29 Anxiety	10	1	8	2	2.90	2.23
		GAD-7	10	1	8	4	3.80	2.52
		HADS	9	1	8	4	4.33	2.31
	Catastrophize	Pain Catastrophizing Scale	10	1	7	1	3.20	2.67
		Coping Strategies Subscale	11	1	6	1	3.27	2.01

Report of the Task Force on Research Standards for Chronic Low-Back Pain

	PTSD	PTSD Checklist	11	2	9	5	5.18	2.00
	Subst. Abuse	TBD	7	1	5	1	2.29	1.77
Psycho-social function	Pain	PROMIS29 Interfer	11	1	5	3	3.45	1.13
	Interference-AOL	BPI (Interference) Graded Chronic Pain Scale	10	1	8	1/3 4	3.50 3.70	2.30 1.64
	HRQoL	SF-36	11	2	6	4	3.91	1.33
		SF-12; 2 Items per SF-36 Scale	10	2	9	3	4.00	2.72
	Satisf. Soc. Role	PROMIS29	8	1	9	5	4.13	2.80
	EuroQOL	EQ5D	11	1	9	4	4.18	2.64
	Work/Injury Hx.	1. Disab. Comp; 2. Litig.	10	1	3	2	1.70	0.67
Outcomes	PGIC	Pt Global Impress Chng	9	1	3	2	2.00	0.71

APPENDIX 3.3

Feasibility of RTF Developing an RDC/cLBP Diagnostic and Classification System

Survey of the NIH Research Task Force on Standards for Chronic Low-Back Pain Research

Survey Objective: To determine if it is feasible/desirable for the RTF to initiate development of Research Diagnostic Criteria for subsets of chronic low-back pain syndrome (RDC/cLBPS)

Overview:

The impetus for a biopsychosocially-based RDC/cLBPS derives from conclusions at several NIH Pain Consortium workshops arguing that inability to categorize subtypes of back pain patients hinders progress in understanding and managing back pain and limits interpretation of published findings, comparisons among studies, and replication of results. As part of its charge from the NIH Pain Consortium, the RTF was directed to examine the possibility of developing an RDC/cLBPS system; i.e., criteria for subsets of the cLBP syndrome.

Survey Items: (Please use the scale from 1 to 5 below to indicate how much you agree or disagree with each of the items listed below):

1. Developing subsets of RDC/cLBPS would potentially allow refinement of research study designs, reduce ambiguity in interpretation of published reports, and facilitate more reliable and valid independent replication of published findings:

1. Strongly Agree 2. Agree 3. Undecided 4. Disagree 5. Strongly Disagree

2. For the present, any attempt to develop RDC/cLBPS will need to be based on a descriptive system of signs and symptoms rather than on etiologic mechanisms of action:

1. Strongly Agree 2. Agree 3. Undecided 4. Disagree 5. Strongly Disagree

3. Given current scientific knowledge, it is not possible to operationally define clinically relevant subsets of the cLBP syndrome (e.g., musculoligamentous vs. discogenic vs. articular vs. neuropathic):

1. Strongly Agree 2. Agree 3. Undecided 4. Disagree 5. Strongly Disagree

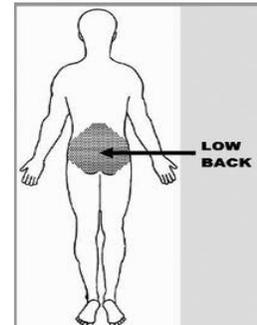
4. **Before attempting to operationally define subsets of cLBPS**, I believe the Task Force should develop standardized Axis I and Axis II datasets to encourage future research to define important subsets with adequate reliability and validity:

1. Strongly Agree 2. Agree 3. Undecided 4. Disagree 5. Strongly Disagree

APPENDIX 4

Recommended Multidimensional Minimal Dataset for Research on cLBP

Minimal Dataset
(PROMIS items marked with ¹; STarT Back or nearly identical items marked with ²; RTF Impact Classification items marked with *)



1. How long has low-back pain been an ongoing problem for you?

- Less than 1 month
- 1–3 months
- 3–6 months
- 6 months–1 year
- 1–5 years
- More than 5 years

2. How often has low-back pain been an ongoing problem for you over the past 6 months?

- Every day or nearly every day in the past 6 months
- At least half the days in the past 6 months
- Less than half the days in the past 6 months

3. In the past 7 days, how would you rate your low-back pain on average?*^{1,2}

- | | | | | | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| No pain | | | | | | | | | Worst Imaginable pain |

4. Has back pain spread down your leg(s) during the past 2 weeks?²

- Yes
- No
- Not sure

5. During the past 4 weeks, how much have you been bothered by ...

	Not bothered at all	Bothered a little	Bothered a lot
• Stomach pain	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Pain in your arms, legs, or joints other than your spine or back	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Headaches	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Widespread pain or pain in most of your body	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. Have you ever had a low-back operation?

- Yes, one operation
- Yes, more than one operation
- No

7. If yes, when was your last back operation?

- Less than 6 months ago
- More than 6 months but less than 1 year ago
- Between 1 and 2 years ago
- More than 2 years ago

8. Did any of your back operations involve a spinal fusion? (also called an arthrodesis)

- Yes
- No
- Not sure

In the past 7 days...	Not at all	A little bit	Somewhat	Quite a bit	Very much
9. How much did pain interfere with your day-to-day activities?*	<input type="checkbox"/>				
10. How much did pain interfere with work around the home?*	<input type="checkbox"/>				
11. How much did pain interfere with your ability to participate in social activities?*	<input type="checkbox"/>				
12. How much did pain interfere with your household chores?*	<input type="checkbox"/>				

13. Have you used any of the following treatments for your back pain? (Check all that apply)

	Yes	No	Not sure
• Opioid painkillers (prescription medications such as Vicodin, Lortab, Norco, hydrocodone, codeine, Tylenol #3 or #4, Fentanyl, Duragesic, MS Contin, Percocet, Tylox, OxyContin, oxycodone, methadone, tramadol, Ultram, Dilaudid)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If you checked yes, are you currently using this medication?.....	<input type="checkbox"/>	<input type="checkbox"/>	
• Injections (such as epidural steroid injections, facet injections)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Exercise therapy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Psychological counseling, such as cognitive-behavioral therapy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The next two questions are for people who normally work outside the home.

14. I have been off work or unemployed for 1 month or more due to low-back pain.

- Agree
- Disagree
- Does not apply

15. I receive or have applied for disability or workers' compensation benefits because I am unable to work due to low-back pain.

- Agree
- Disagree
- Does not apply

Physical Function	Without any difficulty	With a little difficulty	With some difficulty	With much difficulty	Unable to do
16. Are you able to do chores such as vacuuming or yard work? ^{*1}	<input type="checkbox"/>				
17. Are you able to go up and down stairs at a normal pace? ^{*1}	<input type="checkbox"/>				
18. Are you able to go for a walk of at least 15 minutes? ^{*1,2}	<input type="checkbox"/>				
19. Are you able to run errands and shop? ^{*1}	<input type="checkbox"/>				
In the past 7 days...	Never	Rarely	Sometimes	Often	Always
20. I felt worthless ¹	<input type="checkbox"/>				
21. I felt helpless ¹	<input type="checkbox"/>				
22. I felt depressed ¹	<input type="checkbox"/>				
23. I felt hopeless ¹	<input type="checkbox"/>				
In the past 7 days...	Very poor	Poor	Fair	Good	Very good
24. My sleep quality was ¹	<input type="checkbox"/>				
In the past 7 days...	Not at all	A little bit	Somewhat	Quite a bit	Very much
25. My sleep was refreshing ¹	<input type="checkbox"/>				
26. I had a problem with my sleep ¹	<input type="checkbox"/>				
27. I had difficulty falling asleep ¹	<input type="checkbox"/>				

28. It's not really safe for a person with my back problem to be physically active.²

- Agree
- Disagree

29. I feel that *my back pain is terrible and it's never going to get any better.*²

- Agree
- Disagree

30. Are you involved in a lawsuit or legal claim related to your back problem?

- Yes
- No
- Not sure

In the past year:

	Never	Rarely	Sometimes	Often
31. Have you drunk or used drugs more than you meant to?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. Have you felt you wanted or needed to cut down on your drinking or drug use?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

33. Age: _____ years (0–120)

34. Gender:

- Female
- Male
- Unknown
- Unspecified

35. Ethnicity: (*"X" ONLY one with which you MOST CLOSELY identify*)

- Hispanic or Latino
- Not Hispanic or Latino
- Unknown
- Not Reported

36. Race: (*"X" those with which you identify*)

- American Indian or Alaska Native
- Asian
- Black or African-American
- Native Hawaiian or Other Pacific Islander
- White
- Unknown
- Not Reported

37. Employment Status:

- Working now
- Looking for work, unemployed
- Sick leave or maternity leave
- Disabled due to back pain, permanently or temporarily
- Disabled for reasons other than back pain
- Student
- Temporarily laid off
- Retired
- Keeping house
- Other, Specify: _____
- Unknown

38. Education Level: *(select the highest level attained)*

- No high school diploma
- High school graduate or GED
- Some college, no degree
- Occupational/technical/vocational program
- Associate degree: academic program
- Bachelor's degree
- Master's degree (e.g., M.A., M.S., M.Eng., M.Ed., M.B.A.)
- Professional school degree (e.g., M.D., D.D.S., D.V.M., J.D.)
- Doctoral degree (e.g., Ph.D., Ed.D.)
- Unknown

39. How would you describe your cigarette smoking?

- Never smoked
- Current smoker
- Used to smoke, but have now quit

- 40. Height:** _____ inches centimeters measured self-reported
Weight: _____ pounds kilograms measured self-reported