

# Options for Documenting Functional Improvement in Conservative Care

## Purpose and Intended Use

### Table of Contents

#### [Summary Information](#)

#### [Clinical Resources](#)

- Progress Checklist
- Algorithm for Functional Measurement Options

#### [Scales Recommended by IICAC for Routine Use](#)

- Generic Musculoskeletal Scales
- Psychosocial Scales
- Regional Scales - Spine
- Regional Scales - Upper Extremity
- Regional Scales - Lower Extremity

#### [Evidence Summary for Functional Measurement](#)

#### [Issues plus Additional Scales](#)

- Functional Measurement Summary
- General Health Status and Quality of Life Measurement Instruments
- Generic Musculoskeletal Scales
- Psychosocial Scales
- Regional Scales
- Condition Specific Scales
- Routine Use Anchored and Numerical Scales
- Physical Capacity Measurement

#### [Additional Materials](#)

- Evidence and Methodology

#### [Citations](#)

This document updates a 2012 resource developed by the Industrial Insurance Chiropractic Advisory Committee (IICAC) of the Washington State Department of Labor and Industries. It provides concise summaries of published clinical and scientific literature regarding utility and effectiveness of commonly used conservative approaches for work-related musculoskeletal conditions; history, examination and special studies, recommendations for supportive, manual, and rehabilitative care including practical clinical resources (useable without licensing/charge in practice for non-commercial use). It is intended to inform care options and shared decision-making. It is not a standard of care, claim management standard, or a substitute for clinical judgment in an individual case. This practice resource does not change L&I coverage or payment.

Included in this summary are recommendations from IICAC regarding which functional and outcome tracking instruments and scales might be easily adopted into practice. In addition to validation for diagnostic, screening, or baseline assessments, IICAC factored in attributes including detection of meaningful clinical change, ease of administration and scoring, and ability to utilize them in clinical practice without cost or burdensome licensing requirements. A companion document that includes the recommended scales is available at the link below.

A comprehensive search of available scientific literature on instruments and strategies for documenting and tracking functional status was conducted by the Policy, Practice, and Quality (PPQ) Subcommittee of the IICAC and department staff during Spring 2012. Literature was reviewed, assessed for relevance and quality and summaries were drafted by consensus of the subcommittee with expert content input from consultants and reviewers, including the Industrial Insurance Medical Committee and selected relevant professional societies in March 2012. The updated resource was posted for public comment and revision, and approved for distribution by the IICAC in April 2012. A minor update to include updated sources was made and approved at its April 2014 meeting. This resource is expected to be updated periodically by the IICAC. Interested parties may submit new published scientific report for consideration for future revisions.

This and other practice resources are available for download at the State of Washington Department of Labor & Industries website. Contact information for public input and submission of studies for future revisions is available there.

<http://www.lni.wa.gov/ClaimsIns/Providers/ProjResearchComm/IICAC>

#### Subcommittee

J.F. Lawhead DC, Chair  
Robert Baker, DC  
Linda DeGroot, DC  
Michael Neely, DC

#### Consultants

Judy Turner, Ph.D.  
Steve Yeomans, DC

#### Department Staff

Reshma Kearney, MPH  
Bintu Marong, MS  
Robert D. Mootz, DC

## PRACTICAL APPLICATION POINTS

- Outcomes assessment scales provide a concise, valid way to track function and improvement in function. Meaningful change usually involves at least a 30% improvement in score. 50% improvement can typically be considered to be substantial.
- Anchored numerical scales are recommended for tracking routine progress, particularly pain interference with important activities.
- Regional or condition functional outcome scales should be routinely used at baseline and periodic follow-ups. More frequent follow-up is recommended with higher frequency care.
- Psychosocial scales help identify those at higher risk of chronicity and improvement in fear avoidance scales predicts later improvement.
- Several physical performance outcomes also have substantial reported reliability and clinical meaningfulness.

### Functional Improvement

- Ideally, care should contribute to better and faster improvement in function and return-to-work than natural progression. To determine degree of improvement, it is recommended that specific function and activity levels be documented before care begins and at periodic intervals as care is provided. Examples of valid and reliable patient self-report strategies and tools are included in this resource.

### Curative & Rehabilitative Care

- Washington State workers' compensation law mandates that the care workers receive is curative and/or rehabilitative (WAC 296-20-01002). In non-catastrophic cases, this has been operationalized by clinical documentation that demonstrates improved physical function (including return-to-work) is occurring.

### Maximal Medical Improvement (MMI)

- MMI occurs when no marked change in the workers' condition can be expected, with or without treatment. Fluctuations in pain and function may occur once MMI is reached. Over time, improvement or deterioration may occur once MMI is reached. Treatment that results only in temporary or transient changes is not considered proper and necessary. (WAC 296-20-01002)

### General Health/Biopsychosocial Status Measurement Summary

- Numerous instruments have been used to capture general health status. Instruments typically capture elements of physical and mental function attributable to the respondent's state of health. The most widely used validated examples include the SF-36, HSQ-36, SF-12, and HSQ-12.
- Increasing evidence has emerged that fear of activity and low recovery expectations are associated with poorer outcomes from common musculoskeletal conditions. Increasing attention to assessing and tracking certain mental health and psychosocial health status elements has resulted in using instruments (e.g., SBST-9, TSK-11, FABQ) to help determine which interventions should be considered and to assess improvement.

### Regional Functional Measurement Summary

- Many anatomic regional area instruments have been developed for the neck, back, and upper and lower limbs. These have the advantage of assessing impact of multiple affected sites with a single instrument. Examples include the QuickDASH, NDI, ODI, and LEFS.
- Instruments addressing a specific joint (e.g., SST for shoulder, FAAM for foot and ankle) have also been validated and sometimes offer more specificity and sensitivity to monitor response to interventions.

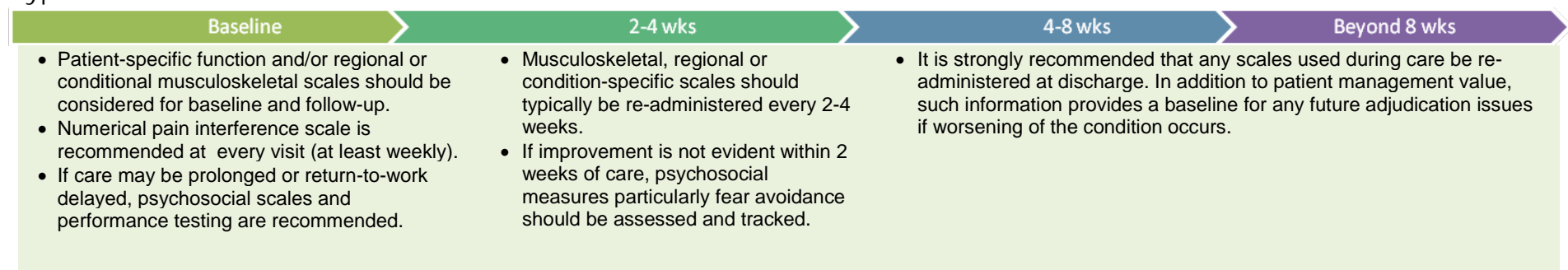
### Condition-specific Measurement Summary

- Instruments have also been developed and validated for a specific condition such as carpal tunnel syndrome, lateral epicondylitis, osteoarthritis, and many other conditions seen in occupational and primary care.

### Physical Performance Testing (PPT) Measurement Summary

- PPT may help assess/track conditioning particularly when recovery is not evident by 4-6 weeks.

## Typical Functional Measurement Thresholds



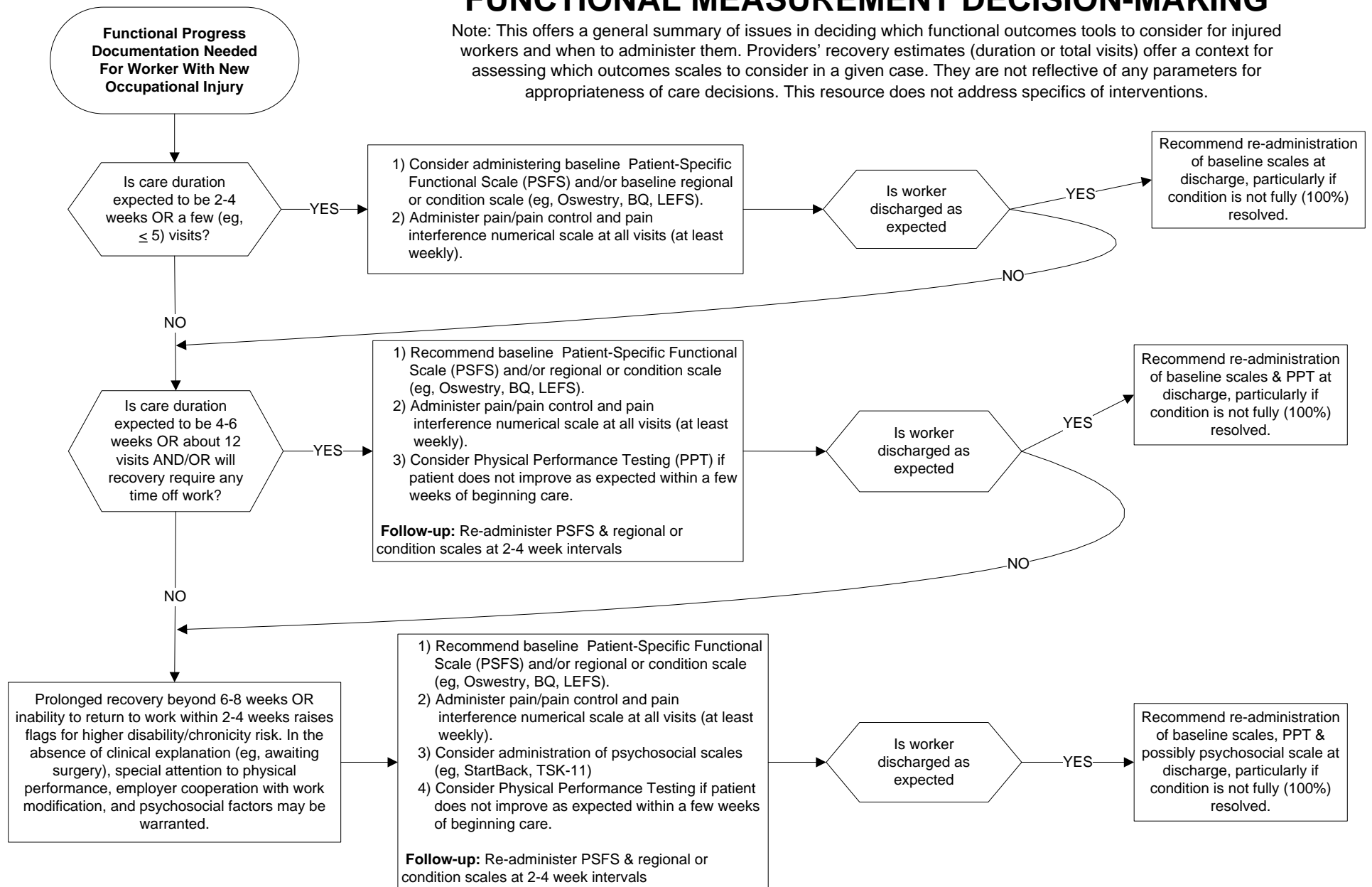
# FUNCTIONAL PROGRESS CHECKLIST

Voluntary educational / practice aid – Not an L&I documentation requirement

	Baseline	1-2 wks	3-6 wks	7-8 wks	Beyond 8 wks
Assessment / Progress	<b>Date:</b> _____  <b>Baseline Function Score:</b> _____  <b>Pain Interference*</b> 0 1 2 3 4 5 6 7 8 9 10 <i>None</i> <span style="float: right;"><i>Unable to do any activities</i></span>	<b>Date:</b> _____  <b>Function Score:</b> _____  <b>Pain Interference</b> 0 1 2 3 4 5 6 7 8 9 10 <i>None</i> <span style="float: right;"><i>Unable to do any activities</i></span>	<b>Date:</b> _____  <b>Function Score:</b> _____  <b>Pain Interference</b> 0 1 2 3 4 5 6 7 8 9 10 <i>None</i> <span style="float: right;"><i>Unable to do any activities</i></span>	<b>Date:</b> _____  <b>Function Score:</b> _____  <b>Pain Interference</b> 0 1 2 3 4 5 6 7 8 9 10 <i>None</i> <span style="float: right;"><i>Unable to do any activities</i></span>	
	<b>Self-control of pain**</b> 0 1 2 3 4 5 6 7 8 9 10 <i>Complete control of pain</i> <span style="float: right;"><i>No control of pain</i></span>	<b>Self-control of pain</b> 0 1 2 3 4 5 6 7 8 9 10 <i>Complete control of pain</i> <span style="float: right;"><i>No control of pain</i></span>	<b>Self-control of pain</b> 0 1 2 3 4 5 6 7 8 9 10 <i>Complete control of pain</i> <span style="float: right;"><i>No control of pain</i></span>	<b>Self-control of pain</b> 0 1 2 3 4 5 6 7 8 9 10 <i>Complete control of pain</i> <span style="float: right;"><i>No control of pain</i></span>	
	<b>Work Status</b> <input type="checkbox"/> Full Duty <input type="checkbox"/> Modified <input type="checkbox"/> None	<b>Work Status</b> <input type="checkbox"/> Full Duty <input type="checkbox"/> Modified <input type="checkbox"/> None	<b>Work Status</b> <input type="checkbox"/> Full Duty <input type="checkbox"/> Modified <input type="checkbox"/> None	<b>Work Status</b> <input type="checkbox"/> Full Duty <input type="checkbox"/> Modified <input type="checkbox"/> None	
	<b>Date</b> <b>Baseline Scale</b> <b>Score</b> _____ _____ _____	<b>Date</b> <b>Follow up Scale</b> <b>Score</b> _____ _____ _____	<b>Date</b> <b>Follow up Scale</b> <b>Score</b> _____ _____ _____	<div> <b>*Pain Interference:</b> "In the past week, how much did pain interfere with your daily activities?"   <b>**Self-control:</b> "In the past week, how much control were you able to have over your pain?"           </div>	
Outcome Scales / Tracking	<b>Musculoskeletal Questionnaires/Scales:</b> <input type="checkbox"/> Patient Specific Function (PSFS) <input type="checkbox"/> Bournemouth Back (BQ-Back) <input type="checkbox"/> Oswestry (ODI) <input type="checkbox"/> Roland (RMQ) <input type="checkbox"/> Bournemouth Neck (BQ-Neck) <input type="checkbox"/> Neck Disability (NDI) <input type="checkbox"/> Shoulder Pain and Disability (SPADI) <input type="checkbox"/> QuickDASH & work module (QDASH) <input type="checkbox"/> Upper Extremity Function (UEFI) <input type="checkbox"/> Upper Limb Function (ULFI) <input type="checkbox"/> Lower Extremity Function (LEFS)  <input type="checkbox"/> Other: _____	If return to work/good improvement is not seen by 2 weeks of care, consider a psychosocial scale to assess chronicity/disability risk, which may help inform care planning.  <b>Psychosocial Questionnaires/Scales:</b> <input type="checkbox"/> STarT Back-9 (SBST-9) <input type="checkbox"/> Patient Health (PHQ-9) <input type="checkbox"/> Fear-Avoidance Belief (FABQ) <input type="checkbox"/> Kinesiophobia (TSK-11) <input type="checkbox"/> Yellow Flags Disability (YFDQ)  <input type="checkbox"/> Other: _____	If return to work or adequate improvement is not seen by 4-6 weeks of care, physical performance tests can be used to assess a 'baseline level' of conditioning and to help target rehabilitation options. Repeat follow-up at 4-6 weeks may assess progress. If not yet included, a psychosocial scale may be considered.  <input type="checkbox"/> Short Physical Performance Battery (SPPB) for older patients <input type="checkbox"/> Back Physical Performance Battery (BPPB) <input type="checkbox"/> Static Neck Endurance  <input type="checkbox"/> Other: _____  <b>Patient's Name:</b> _____		

## FUNCTIONAL MEASUREMENT DECISION-MAKING

Note: This offers a general summary of issues in deciding which functional outcomes tools to consider for injured workers and when to administer them. Providers' recovery estimates (duration or total visits) offer a context for assessing which outcomes scales to consider in a given case. They are not reflective of any parameters for appropriateness of care decisions. This resource does not address specifics of interventions.



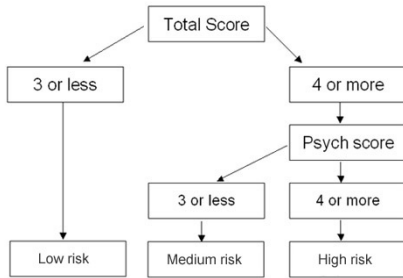
## OUTCOMES SCALES RECOMMENDED BY IICAC FOR ROUTINE USE

### Generic Musculoskeletal Scales

Outcome Scale	Description & Purpose	Administration	Scoring & Interpretation	Licensing
<b>Patient-Specific Functional Scale (PSFS)</b> <sup>1, 2</sup>	<p><b>For:</b> Back, neck, knee disability (validated); other musculoskeletal (not validated)</p> <p><b># Items:</b> 3-5 activities selected by patient</p> <p><b>Other:</b> Patient chooses ADL limitations most important to them.</p> <p><b>Link:</b>  <a href="http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf">http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf</a> </p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks.</p> <p><b>Completion Time:</b> Usually less than a minute (depends on number of ADLs chosen).</p>	<p>0-10 scale (Worst = 0 to best = 10). 3-5 ADLS used (Usually 3). Neck version adds 2 questions regarding pain (0-10 neck pain scale and 10-0 activity of daily living (ADL) pain-interference scale.</p> <p><b>Meaningful change:</b> Neck: 1 point for the average of 3 ADLs and for the pain limitation measure. For individual ADLS, 2 points.</p>	None

### Psychosocial Scales – Depression/Anxiety/Kinesiophobia

Outcome Scale	Description & Purpose	Administration	Scoring & Interpretation	Licensing
<b>Fear Avoidance Belief Questionnaire (FABQ)</b> <sup>3</sup>	<p><b>For:</b> Validated for chronic low back pain in an injured worker population<sup>3</sup> but may help identify acute back patients at risk of poor outcome.<sup>4</sup></p> <p><b># Items:</b> 16</p> <p><b>Other:</b> May be used for other conditions by modifying items 3 &amp; 11 to the condition the patient has. Includes two sections: Physical Activity (PA-5 questions) and Work Activity (WA-11 questions).</p> <p><b>Link:</b>  <a href="http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf">http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf</a> </p>	<p><b>Baseline:</b> Optional at intake in acute care, but recommended if suspicion of chronicity risk exists or meaningful improvement does not occur within 2-4 weeks.</p> <p><b>Follow-up:</b> After about 4-6 weeks following initiation of care.</p> <p><b>Completion Time:</b> Less than 10 minutes</p>	<p>Each Item has an agreement response scale (0 completely disagree- 3 unsure- 6 completely agree). The FABQ has a total score (sum all marked items -96 possible) and two subscales PA (items 2, 3, 4, 5; -24 possible) and WA (items 6, 7, 9, 10, 11, 12, 15 -42 possible) sections. Higher scores reflect higher fear avoidance beliefs.</p> <p><b>Meaningful change:</b> Not designed as a tracking instrument so meaningful change has not been determined for questionnaire as a whole. Has been shown to correlate with TSK-11 scores (Woby 2004). If used for tracking, 30-50% improvement is considered meaningful.</p>	None
<b>STarT Back Screening Tool-9 (SBST-9)</b> <sup>5-8</sup>	<p><b>For:</b> Non-specific back pain in primary care when chronicity is of concern. Items are drawn from several validated scales.</p> <p><b># Items:</b> 9</p> <p><b>Other:</b> Domains addressed include referred leg pain and co-morbid pain, disability, catastrophizing, fear avoidance, anxiety and depression. Validated in multiple settings, specialties and languages.</p> <p><b>Link:</b> <a href="http://www.keele.ac.uk/sbst/">http://www.keele.ac.uk/sbst/</a></p>	<p><b>Baseline:</b> Optional at intake in acute care, but recommended if suspicion of chronicity risk exists or meaningful improvement does not occur within 2 weeks.</p> <p><b>Follow-up:</b> After about 4 weeks following initiation of care.</p> <p><b>Completion Time:</b> Less than 5 minutes</p>	<p>Eight of the 9 items are agree/disagree with agree being a positive response. One question is a 5-point scale for which either of 2 responses are positive (very much and extremely). An overall score is made by summing all positive responses. Three or fewer positives represent a low chronicity risk. Four or more positives require looking at the distress subscale (last 5 items). Three or less positives in the last 5 represent medium chronicity risk while 4 or more reflect a high chronicity risk.</p>	None, provided the copyright and funding statement is used on the instrument

			 <p><b>Meaningful change:</b> This scale is designed for assessing chronicity risk and guiding intervention for non-responders. Given the nature of the questions and their origin in other instruments, re-administration of the tool would seem reasonable to assess if their risk of chronicity is diminishing.</p>	
<b>Tampa Scale for Kinesiophobia-11 (TSK-11)</b> <sup>9, 10</sup>	<p><b>For:</b> Assesses pain-related fear in back patients.</p> <p><b># Items:</b> 11</p> <p><b>Other:</b> A shortened version of the TSK-17 and 13, using only the questions with best psychometrics. Appears useful in spine care settings.</p> <p><b>Link:</b>  <a href="http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf">http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf</a> </p>	<p><b>Baseline:</b> Optional at intake in acute care, but recommended if suspicion of chronicity risk exists or meaningful improvement does not occur within 2 weeks.</p> <p><b>Follow-up:</b> After about 4 weeks following initiation of care.</p> <p><b>Completion Time:</b> Less than 5 minutes</p>	<p>11 statements are answered on a 4 point scale (1-4 disagree – agree) and the point value is summed. Score may be between 11 and 44. Higher scores reflect more anxiety and fear avoidance and correlate with greater likelihood of developing chronicity.</p> <p><b>Meaningful change:</b> A change of 4 points can be considered meaningful.</p>	None
<b>Yellow Flag Disability Questionnaire (YFDQ)</b> <sup>11</sup>	<p><b>For:</b> Any musculoskeletal condition</p> <p><b># Items:</b> 13</p> <p><b>Other:</b> Based on numerical scale questions (independently validated elsewhere) for domains of pain and function (work, sleep)</p> <p><b>Link:</b>  <a href="http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf">http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf</a> </p>	<p><b>Baseline:</b> Optional at intake in acute care, but recommended if suspicion of chronicity risk exists or meaningful improvement does not occur within 2 weeks.</p> <p><b>Follow-up:</b> After about 4 weeks following initiation of care.</p> <p><b>Completion Time:</b> Less than 5 minutes</p>	<p>Scored by adding the circled number on each item's scale (except item 3 which has anchors reversed and is scored as 10 minus the circled number). Includes score sheet with space for recoding all items' numeric responses at baseline and 5 follow-ups for easy reference for each item, domain, and entire score.</p> <p><b>Meaningful change:</b> Not determined specifically for this questionnaire. Generally, 30-50% improvement is considered meaningful. Greater than 65 point improvement is considered supportive of long term recovery.<sup>17</sup></p>	None

## Regional Scales – Spine

Outcome Scale	Description & Purpose	Administration	Scoring & Interpretation	Licensing
Bournemouth Questionnaire - Back	<b>For:</b> Back complaints	<b>Baseline:</b> At intake	Each of the seven functional items is scored on a 0-10 point numerical rating scale with a total of 70 points possible. A	None



<b>(BQ-Back)</b> <sup>12</sup>	<p><b># Items:</b> 7</p> <p><b>Other:</b> Based on the function questions of the Bournemouth Questionnaire (BQ). Addresses pain, pain interference on ADLs and psychosocial factors (anxiety, depression, locus of control). Initial and follow-up versions accommodate differences in context.</p> <p><b>Link:</b> <a href="http://www.aecc.ac.uk/research/bu-study.aspx">http://www.aecc.ac.uk/research/bu-study.aspx</a></p>	<p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks.</p> <p><b>Completion Time:</b> &lt; 5 minutes</p>	<p>lower score reflects less disability.</p> <p>In addition, the versions of the BQ available from the developing institution include several additional questions on change in medication use, bothersomeness of complaint in the past few days, and global assessment of improvement.</p> <p><b>Meaningful change:</b> A change of 17 points or 47% (follow-up score/baseline score x 100) on the BQ correlated significantly with the patient's sense of global improvement.<sup>13</sup></p>	
<b>Bournemouth Questionnaire - Neck (BQ-Neck)</b> <sup>14</sup>	<p><b>For:</b> Neck complaints</p> <p><b># Items:</b> 7</p> <p><b>Other:</b> Based on the function questions of the BQ. Addresses pain, pain-interference on ADLs, and psychosocial factors (anxiety, depression, locus of control).</p> <p><b>Link:</b> <a href="http://www.aecc.ac.uk/research/bu-study.aspx">http://www.aecc.ac.uk/research/bu-study.aspx</a></p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks</p> <p><b>Completion Time:</b> 1-2 minutes</p>	<p>Each of the seven functional items is scored on a 0-10 point numerical rating scale with a total of 70 points possible. A lower score reflects less disability.</p> <p>In addition, the versions of the BQ available from the developing institution include several additional questions on change in medication use, bothersomeness of the complaint in the past few days, and a global assessment of improvement.</p> <p><b>Meaningful change:</b> A change of 13 points or 34% (follow-up score/baseline score x 100) on the BQ correlated significantly with the patient's sense of global improvement.<sup>13</sup></p>	None
<b>Neck Disability Index (NDI)</b> <sup>15</sup>	<p><b>For:</b> Neck pain-related functional limitation</p> <p><b># Items:</b> 10</p> <p><b>Other:</b> Patterned after the ODI (below). Addresses pain level, pain interference with ADLs, sleep, etc.</p> <p><b>Link:</b> <a href="http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf">http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf</a></p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks.</p> <p><b>Completion Time:</b> 3-5 minutes</p>	<p>Each question has 6 responses scored on an ascending scale (0, 1, 2, 3, 4, 5). The 10 questions are totaled, and then divided by the number of points possible (50 if all questions are answered). This score is expressed as a percentage (by multiplying by 100)</p> <p>Scores range from 0-100% with higher being worse. Typical 'global' interpretation. A higher score means worse disability.</p> <ul style="list-style-type: none"> <li>0-20% minimal</li> <li>20-40% moderate</li> <li>40-60% severe</li> <li>60-80% housebound</li> <li>80-100% bedbound or exaggerating (indicates need for further assessment).</li> </ul> <p><b>Meaningful change:</b> Minimal detectable change (MDC) was reported to be 10% (approx. 5 points).<sup>16</sup> Expert consensus considers clinically meaningful change to be 30-50% (approx. 15 points).<sup>17</sup></p> <p>REF: <a href="http://www.tac.vic.gov.au/upload/OM-Ortho.pdf">http://www.tac.vic.gov.au/upload/OM-Ortho.pdf</a></p>	None

<b>Modified Oswestry Low Back Disability Index (ODI)</b> <sup>18, 19</sup>	<p><b>For:</b> Back pain-related disability and functional limitation.</p> <p><b># Items:</b> 10</p> <p><b>Other:</b> Addresses pain level, pain interference with ADL, sleep, etc. Original version includes a question on sex life which has been replaced by one on employment and homemaking on the modified version.</p> <p><b>Link:</b>  <a href="http://www.workcover.com/site/treat_home/outcome_measures_and_risk_screening_tools.aspx?str=functional%20patient%20scale%20specific">http://www.workcover.com/site/treat_home/outcome_measures_and_risk_screening_tools.aspx?str=functional%20patient%20scale%20specific</a> </p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks.</p> <p><b>Completion Time:</b> 3-5 minutes</p>	<p>Each question has 6 responses scored on an ascending scale (0, 1, 2, 3, 4, 5). The 10 questions are totaled, and then divided by the number of points possible (50 if all questions are answered). This score is expressed as a percentage (by multiplying by 100)</p> <p>Scores range from 0-100% with higher being worse. Typical 'global' interpretation. A higher score means worse disability.</p> <ul style="list-style-type: none"> <li>0-20% minimal</li> <li>20-40% moderate</li> <li>40-60% severe</li> <li>60-80% housebound</li> <li>80-100% bedbound or exaggerating (indicates need for further assessment).</li> </ul> <p><b>Meaningful change:</b> Minimal detectable change (MDC) has been reported to be a 10% change.<sup>18</sup> Meaningful change is typically considered to be 4-16 points or 30-50%.<sup>17, 20-22</sup></p>	None
<b>Roland-Morris Low Back Disability Questionnaire (RMQ)</b> <sup>23-26</sup>	<p><b>For:</b> Disability related to low back pain</p> <p><b># Items:</b> 24</p> <p><b>Other:</b> Lists ADLs for which the patient simply checks which ones are limited due to their back pain. Validity, utility, and comparability to other measures has been reported.</p> <p><b>Link:</b>  <a href="http://www.workcover.com/site/treat_home/outcome_measures_and_risk_screening_tools.aspx?str=functional%20patient%20scale%20specific">http://www.workcover.com/site/treat_home/outcome_measures_and_risk_screening_tools.aspx?str=functional%20patient%20scale%20specific</a> </p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks.</p> <p><b>Completion Time:</b> 1-2 minutes</p>	<p>All items indicated by the patient are summed for a low of 0 to a maximum score of 24. A lower score means less disability. A score &gt;13 points = significant disability (unfavorable outcome)</p> <p><b>Meaningful change:</b> In general, a 30% change in RMQ score can be considered meaningful with 50% considered substantial. Studies have reported that high initial RMQ scores require larger amounts of change to be considered meaningful, while smaller amounts of change may be meaningful in patients reporting lower initial scores.<sup>21, 22</sup></p>	None
<b>Regional Scales – Upper Extremity</b>				
<b>QuickDASH (Q-DASH)</b> <sup>27</sup>	<p><b>For:</b> Upper extremity musculoskeletal conditions that restrict function</p> <p><b># Items:</b> 11 + 4</p> <p><b>Other:</b> Addresses work- and sports-related activities</p> <p><b>Link:</b></p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks</p> <p><b>Completion Time:</b> 5-6 minutes</p>	<p>Two components are scored separately:</p> <p><b>Disability Section</b> (11 items scored 1-5). At least 10 of the items must be answered to score the test. Responses are summed and averaged to produce a score out of 5 possible. The value is transformed to a score out of 100 (to simplify comparisons) by subtracting 1 and multiplying by 25.</p> <p><b>Work Section</b> There are two optional versions of work and</p>	Registration (see link)



	<a href="http://www.dash.iwh.on.ca/conditions.htm">http://www.dash.iwh.on.ca/conditions.htm</a> .		high performance (4 items scored 1-5). All completed responses are summed and divided by 4, then multiplied by 25.  <b>Meaningful change:</b> Minimal clinically important difference has been reported to be 19 points with minimal detectable change being 11 points. <sup>28</sup>	
<b>Shoulder Pain &amp; Disability Index (SPADI)</b> <sup>29-32</sup>	<p><b>For:</b> General shoulder conditions that cause pain and disability (functional limitation) with various activities.</p> <p><b># Items:</b> 13</p> <p><b>Other:</b> Pain and difficulty is rated when it's at its worst and when engaging in various positions/activities when using the affected arm. Activities include: reaching, pushing, cleaning, placing objects in front or above, removing contents of back pocket, etc.</p> <p><b>Link:</b>  <a href="http://www.workcover.com/site/treat_home/outcome_measures_and_risk_screening_tool/s/links_to_outcome_measures_and_screening_tools.aspx?str=functional%20patient%20scale%20specific">http://www.workcover.com/site/treat_home/outcome_measures_and_risk_screening_tool/s/links_to_outcome_measures_and_screening_tools.aspx?str=functional%20patient%20scale%20specific</a> </p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks.</p> <p><b>Completion Time:</b> 2-3 minutes</p>	<p>Items are rated on a 10 point numerical scale (0 – 10)</p> <p><b>Pain Scale:</b> Add the scores of 5 pain items, divide by total possible score (50 if all items are answered), then multiply by 100 for a percentage score.</p> <p><b>Disability Scale:</b> Add the scores of 8 disability items, divide by total possible score (80 if all items answered) then multiply by 100 for a percentage score</p> <p><b>Total Score:</b> Add all item scores, divide by total possible score (130 if all items answered) then multiply by 100 for a percentage score.</p> <p>Possible % score is 0-100 for the pain scale, the disability scale and the total score. Higher scores mean worse pain and/or disability.</p> <p><b>Meaningful change:</b> Minimum detectable change (MDC) is 10% or 13 points.<sup>33</sup> Clinically meaningful change is recommended to be at least 30% (Pain scale: 18 points; Disability scale: 13 points).</p>	None
<b>Upper Extremity Functional Index (UEFI)</b> <sup>34</sup>	<p><b>For:</b> Upper limb orthopedic conditions that limit function</p> <p><b># Items:</b> 20</p> <p><b>Other:</b> Addresses work, housework, and recreational activities, grooming, dressing, pushing up, lifting, carrying, driving, sleeping, and throwing.</p> <p><b>Link:</b>  <a href="http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf">http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf</a> </p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 weeks or every 2-4 weeks.</p> <p><b>Completion Time:</b> 2-3 minutes</p>	<p>20 activities are each rated on a scale of 0 (extreme difficulty or unable to do the activity) to 5 (no difficulty) Possible score 0 – 80. A lower score means worse function.</p> <p><b>Meaningful change:</b> Minimal detectable change (MDC) is 10% (8 points). Meaningful change has been reported comparable to QuickDash at about 20% correlating with self-report of global improvement.<sup>35</sup></p>	None

### Regional Scales – Lower Extremity

<b>Outcome Scale</b>	<b>Description &amp; Purpose</b>	<b>Administration</b>	<b>Scoring &amp; Interpretation</b>	<b>Licensing</b>
<b>Foot and Ankle Ability</b>	<b>For:</b> Leg, ankle, and foot disorders	<b>Baseline:</b> At intake	Each item is scored on a five point scale with 4 being “No	None

<b>Measure (FAAM)</b> <sup>36-39</sup>	<p><b># Items:</b> 29</p> <p><b>Other:</b> Updated version of the FADI, minus 5 items (4 pain-related, 1 sleep-related). A self-reported region-specific instrument consisting of a 21-item ADL subscale and an 8-item Sports subscale.</p> <p><b>Link:</b>  <a href="http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf">http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf</a> </p>	<p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks</p> <p><b>Completion Time:</b> 5-6 minutes</p>	<p>Difficulty" and 0 being "Unable To Do." The lowest potential score of the Activities of Daily Living (ADL) subscale of the FAAM is 0 points, the highest 84 points. The lowest potential score of the Sports subscale of the FAAM is 0 points, the highest 32 points. Total score is converted into percentage. Higher percentage indicates higher level of physical function.</p> <p><b>Meaningful change:</b> Minimum clinically important difference (MCID) is reported as: ADL subscale: 8 points. Sports subscale: 9 points.<sup>37</sup> Minimal detectable change (MDC) for the ADL subscale: 5.7 points and the Sports subscale: 12.3 points.<sup>39</sup></p>	
<b>Lower Extremity Functional Scale (LEFS)</b> <sup>40-42</sup>	<p><b>For:</b> Lower extremity orthopedic conditions that limit function</p> <p><b># Items:</b> 20</p> <p><b>Other:</b> Especially useful for higher performance requirements. Focuses on activities and positions such as work/housework, ADLs (dressing, sitting, standing, squatting, walking, stair-climbing, running, hopping, lifting, moving in bed/bath.</p> <p><b>Link:</b>  <a href="http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf">http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FunctionalScales.pdf</a> </p>	<p><b>Baseline:</b> At intake</p> <p><b>Follow-up:</b> After 6-12 visits or every 2-4 weeks.</p> <p><b>Completion Time:</b> 2-3 minutes</p>	<p>Each of the 20 question's 5 possible responses are scored on an ascending scale (0 = Extreme Difficulty, 4= No Difficulty). Points are summed for a maximum possible score of 80. Lower score means worse function.</p> <p><b>Meaningful change:</b> Minimal detectable change (MDC) has been reported to be 9 points (10%).<sup>40</sup> Meaningful change can be assumed to be 30% or about 27 points.</p>	<p>None</p>

## EVIDENCE SUMMARY FOR FUNCTIONAL MEASUREMENT ISSUES & ADDITIONAL SCALES

### FUNCTIONAL MEASUREMENT SUMMARY

#### General Considerations

There are many things to consider when tracking and reporting functional improvement of injured workers. Validity and responsiveness of specific measures are most commonly addressed in the literature. If a questionnaire does not accurately measure what one thinks it does or if it is not responsive to change as the condition improves or worsens, it is not worth using in a practice setting. Tracking functional improvement using validated scales is becoming a best practice because it provides important information regarding the patient's progress over time. Other important factors include how meaningful changes in scale scores are to patients, if functional improvement scores reflect the patient's ability to return to normal activities including work, ease of use (administration, understandability for patients), and licensing issues.

#### Meaningful Clinical Change vs. Minimal

An instrument may display psychometric properties that are sensitive to change (minimal detectable change), even if the amount of change/improvement does not reflect any global improvement in the patient's ability to perform daily activities or the patient's perceived improvement. The concept of meaningful clinical change has become a focus of recent literature with comparisons of scale scores to

## Detectable Change

patient reports of global well-being, comparisons to activity capabilities, provider assessment of improvement, and international expert consensus regarding the magnitude of change.

- To date, no research was identified that correlates magnitude of scale score improvements with a patient's ability to get back to work following an occupational injury.
- Although detailed psychometrics are increasingly reported for published reports of various scales (as well as in this resource), as a general rule, based on substantial literature review and expert consensus, 30% change in most any scale can be considered to be "meaningful" and 50% change to be "substantial."<sup>17</sup>
- A standard for reporting psychometric properties of scales has emerged in the literature for measuring minimal detectable change at the 90% confidence level (MDC-90). This is typically reported as the number of points or percentage change required in a scale to be reliably detected. However, the MDC-90 statistic may not be a reflection of the amount of change (improvement) that either a patient or provider might think is important.
- Several studies have compared various outcomes scales to patient self-report of global ratings of change. A change of 50% on the Modified Oswestry Disability Index (ODI) correlated with global ratings of successful outcome by low back patients seeking PT care.<sup>20</sup> Changes of 47% and 37% in the Bournemouth Questionnaires (BQ) in back and neck pain patients respectively correlated with Patient's Global Impression of Change (PGIC) for improvement.<sup>13</sup>
- An international consensus panel that reviewed literature on visual analog scales, numerical rating scales, Roland Morris Disability Questionnaire (RMQ), the Oswestry Disability Index (ODI) and the Quebec Back Pain Disability Questionnaire (QBPD) concluded that generally a minimum of a 30% change in score is needed to conclude patient-reported improvement is clinically meaningful.<sup>17</sup>
- In low back pain patients, a change of 2 points (20%) on a numerical pain-rating scale correlated with perceived improvement of both patient and therapist using the Global Rating of Change scale.<sup>43</sup>
- Numerical pain-interference scales have been reported to detect minimal clinically important change in back pain patients in the 35-50% range for subacute patients and 25-45% for chronic patients.<sup>44</sup>
- Minimally important change (MIC) has frequently been determined by comparing a scale's reported change with either a patient's self-reported global assessment of improvement (e.g. much improved, not improved) and psychometric/statistical calculation of standard error of measurement (SEM). Several studies have demonstrated correlations between both approaches; however several factors appear to impact how MIC is properly interpreted. Baseline values, the nature of the condition, and direction of change influence how much change is important. For example, patients with a large amount of baseline functional disability on a Roland Morris Disability Questionnaire required a larger amount of change on the scale to consider it important than patients with lower baseline functional disability.<sup>44, 45</sup>
- There do not appear to be any baseline characteristics that predict if a patient will improve, however early improvement on self reported instruments, particularly a decline in fear avoidance scores does correlate with later improvement.<sup>46</sup>

## Ease of Use

Trade-offs often exist between how simple, short, and understandable a questionnaire is and how accurate and meaningful it is. While high reproducibility and comparability are critical in research settings, practical implementation in busy practices is rarely prioritized in published studies. Generally, the consensus of the IICAC and consultants used on this project was that validated questionnaires are critical for use as effectiveness measurements in research settings and when possible, common validated scales should be used in practice settings. More comprehensive regional and condition-specific questionnaires typically administered at 2-4 week intervals are recommended for cases when more frequent patient visits and/or longer treatment durations are expected. Routine visit-to-visit changes can be addressed with numerical scales, particularly when they are aimed at rating how the condition/symptoms interfere with a patient's ability to do particular activities (as opposed to just capturing perceived pain levels).

As this resource illustrates, there are numerous scales that have been developed and validated. Some perform better than others in psychometric tests and many scales are combinations of questions from other validated scales. Those recommended by the IICAC on the preceding pages were selected for a variety of reasons including: the scale itself and/or its elements have been shown to

meaningfully detect change in function; they are relatively straightforward to use; and they are available for use in individual practice (and inclusion here) without proprietary licensing issues and costs. Many other scales may also be validated or may be preferable for individual practice reasons and this resource does not intend to discourage the use of such scales. Additional instruments are included in the summaries below and relevant citations and websites are listed where possible. The most important consideration is to track functional improvement using some kind of patient reported scale. It is perhaps the most certain way to document if the patient is making functional improvement as care is provided.

## GENERAL HEALTH STATUS & QUALITY OF LIFE MEASUREMENT INSTRUMENTS

### General Health Status, Quality of Life Scales

General physical function scales and subscales may be useful for tracking general health and health status instruments in occupational health settings. However, other instruments with scales more specific to activities related to an injured area may be preferred. The instruments typically include elements related to physical abilities, but also capture information related to mental health and general activities.

- **Ambulatory Care Experiences Survey (ACES)** – Purposed similarly to the PCAS (below), but shorter, the ACES is designed to evaluate sustained clinical relationships. Domains addressed include the quality of doctor-patient relationship (communication, care integration, patient's understanding, health promotion) and organization of care (access, continuity, staff).<sup>47</sup>  
<http://160.109.101.132/icrhps/research/thi/questionnaires.asp>
- **Health Status Questionnaire; -36, and -12 question versions (HSQ-36, HSQ-12)** – Developed concurrently with the SF-36 and SF-12 scales (below), the HSQ-36 and HSQ-12 are similar in structure, subscales, and scoring. Additionally, they have been validated against the SF-36, and -12. HSQs have the same utility and limitations, but offer the advantage of less restrictive licensing requirements. An on-line source for the scale could not be found at time of publication, but numerous sources for the instrument and information can be found by searching on the terms HSQ-36, and HSQ-12.
- **Primary Care Assessment Survey (PCAS)** – A 51-question survey which emphasizes domains of the doctor-patient relationship in primary care (communication quality, patient trust, physician knowledge of patient, interpersonal treatment, relationship duration). Elements include linking primary care performance to important outcomes of care (e.g. patient adherence to clinical advice, health improvement, loyalty to a practice). Perhaps best suited for research settings.<sup>47, 48</sup>  
<http://160.109.101.132/icrhps/research/thi/questionnaires.asp>
- **Short Form; -36, -12, and -8 question versions (SF-36, SF-12, SF-12H, SF-8)** – The SF-36 is a general health status questionnaire that includes sections on general health and well-being, mental health, physical function and others. It is widely used and validated in research settings. It is somewhat lengthy and cumbersome to score by hand and requires licensing. Additionally, such scales are geared toward primary care practice and longer term changes in health. Although physical function and mental health subscales are responsive to change, other questionnaires and scales are preferred for routine outcomes tracking in occupational health and musculoskeletal practice settings. Overall, these scales might be most useful to establish a general health baseline once a patient's acute problem stabilizes and it is anticipated the patient will be seen in the practice over multiple episodes and disorders. <http://www.sf-36.org/>

## GENERIC MUSCULOSKELETAL SCALES

### Musculoskeletal Scales

- **Bournemouth Questionnaire (BQ)** – The full BQ includes a baseline and follow-up version to assess how pain and the patient's condition interfere with particular common activities, as well as identifying psychosocial elements. Function questions use a numerical scale approach for pain and pain interference. Theoretically the scale could apply to a variety of musculoskeletal conditions, but two versions have been tailored for neck and back conditions. The initial version (with 27 questions) and follow-up version (with 16 questions) have been validated in neck and back conditions.<sup>12, 13, 49</sup>
- **Örebro Musculoskeletal Pain Questionnaire (ÖMPQ)** – ÖMPQ is a 25-item self-administered questionnaire applicable to any musculoskeletal complaint. Elements address basic intake information (complaint location, duration) along with numerical scales for

usual work activity, pain over previous periods, psychosocial elements, and impact on ADLs. It has been validated as a predictor of failure for return-to-work and has been utilized frequently in research setting but seems to be somewhat cumbersome compared to regional alternatives.<sup>50-53</sup>

[http://www.workcover.com/site/treat\\_home/outcome\\_measures\\_and\\_risk\\_screening\\_tools/links\\_to\\_outcome\\_measures\\_and\\_screening\\_tools.aspx?str=functional%20patient%20scale%20specific](http://www.workcover.com/site/treat_home/outcome_measures_and_risk_screening_tools/links_to_outcome_measures_and_screening_tools.aspx?str=functional%20patient%20scale%20specific)

- **Pain Disability Questionnaire (PDQ)** – The PDI is a 15-item scale (derived from questions used in other scales) that primarily addresses how pain interferes or affects numerous activities of daily living. Nine questions (1, 2, 3, 4, 5, 6, 7, 12, 13) focus on functional status with six (8, 9, 10, 11, 14, 15) emphasizing psychosocial aspects. It is scored by totaling the responses. Subscales can be calculated by totaling responses for the items that make up that subscale. The instrument has been validated for chronic musculoskeletal disorders<sup>54, 55</sup> and is recommended in the American Medical Association's Guides to the Evaluation of Permanent Impairment for determining "functional history adjustment for the cervical, thoracic, and lumbar spine when rating permanent impairment."<sup>56</sup>
- **Patient Specific Functional Scale (PSFS)** – The patient self-selects activities of daily living that are most impacted by their injury or limitation. This scale has the advantage of having a single scale within a practice that can be tailored to the majority of musculoskeletal conditions and is consistently scored. Its utility and psychometric properties have been documented in moderate quality studies. It may not be meaningful for certain activities a patient may select.<sup>2, 57-59</sup>

## PSYCHOSOCIAL SCALES

### Psychosocial Scales

- **Fear-Avoidance Belief Questionnaire (FABQ)** – A 16-item questionnaire validated for chronic low back pain in an injured worker population<sup>3</sup> but may help identify acute back patients at risk of poor outcome.<sup>4</sup> May be used for other conditions by modifying Items 3 and 11 from back pain to the condition the patient has. Includes two sections: Physical Activity (PA-5 questions) and Work Activity (WA-11 questions). Each item has an agreement response scale (0 completely disagree- 3 unsure- 6 completely agree). The FABQ has a total score (sum all marked items -96 possible) and two subscales PA (Items 2, 3, 4, 5; -24 possible) and WA (Items 6,7,9,10,11,12,15; -42 possible). Higher scores reflect higher fear avoidance beliefs and has been reported to better predict 6-month outcomes with physical therapy than the ODI.<sup>60</sup> The FABQ was not designed as a tracking instrument but it has been shown to correlate with TSK-11 scores (Woby 2004). If used for tracking, it is recommended to use 30-50% improvement as meaningful.
- **Functional Recovery Questionnaire (FRQ)** – Currently under development and testing in the Department of Labor and Industries' (L&I) Centers for Occupational Health and Education (COHE) program, the FRQ is based on research specifically in Washington's injured worker population.<sup>61, 62</sup> It is 6 questions of which the first 3 have been shown to be predictive of being off of work one year post-injury. The remaining questions cover work accommodation, recovery expectation and fear-avoidance, which may help target specific interventions. It has only been used as a screening tool and has not been validated to track improvement.
- **Generalized Anxiety Disorder-7 (GAD-7)** – A brief anxiety screening tool developed after the PHQ. Includes 7 items scored 0-3 for a possible 24. A higher score indicates greater anxiety.<sup>63</sup> [www.phgscreeners.com](http://www.phgscreeners.com)
- **Patient Health Questionnaire (PHQ-9)** – A brief 9-question scale that primarily screens for depression and rates its severity. It has been validated against the DSM-IV for screening for depression and depressive episode. It is aimed for use in primary care settings. In addition to assisting in the diagnosis of depression, it may be of use in occupational health settings in slow responders as an indicator for risk of chronic pain. The central mental health orientation of questions may be off-putting to some patients in acute care for musculoskeletal complaints.<sup>64-66</sup> It has also been validated as a brief 2 question screen (PHQ-2) to flag for depressed mood in the previous 2 weeks, primarily useful if positive to target who should receive the PHQ-9.<sup>67</sup> [www.phgscreeners.com](http://www.phgscreeners.com)
- **STarT Back Screening Tool-9 (SBST-9)** – A brief 9-item questionnaire increasingly used in primary care for non-specific back pain especially where chronicity is a potential or current concern. Domains addressed include referred leg pain and comorbid pain, disability, catastrophizing, fear avoidance, anxiety and depression. Wording of psychosocial elements are particularly tolerable for acute care settings and may be used initially. As a screening tool, it has not been assessed as a progress tracking tool but its questions have been drawn from other tools validated for that purpose.<sup>5, 6, 8</sup>
- **Tampa Scale for Kinesiophobia (TSK-11)** – Motivated by the conundrum of uncomplicated back pain patients becoming chronic,

the TSK-11 scale has been developed and tested to determine the role that fear-avoidance (avoiding activities for fear of aggravation or re-injury) might play in the transition from acute injury to chronic pain behavior.<sup>9, 10, 68-70</sup>

- **World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0)** – Developed by WHO to measure function and disability in a standardized fashion. Available in multiple versions (36, 12 questions, self or interview administered) and in several languages. It has been validated and is reported to have good scaling properties across different populations along with strong sensitivity to change.<sup>71</sup> It is easily administered and has been validated for both impacts of musculoskeletal and mental disorders. It may be used without cost but requires registration including completion of a user agreement. Available at <http://www.who.int/classifications/icf/whodasii/en/>
- **Yellow Flag Questionnaire** – A 13-item numerical scale questionnaire, based on several elements (consisting of questions drawn from other validated instruments), that captures pain, self-perception of health, anxiety, depression, function, sleep, and fear avoidance. The purpose of the questionnaire is to assess and track risk factors for chronic disability.<sup>11</sup>

## REGIONAL SCALES

### Spine

#### Cervical

- **Bournemouth Questionnaire – Neck (BQ-Neck)** – Based on the function questions of the BQ, this questionnaire was modified specifically for neck complaints. The BQ-Neck includes 7 items: 1 for pain, 3 for pain-interference on ADLs, and 3 for psychosocial factors (anxiety, depression, and locus of control).<sup>13, 49</sup>
- **Functional Rating Index (FRI)** – The FRI is a 10 item scale (based on elements from the NDI and ODI) that has been validated for neck and low back conditions. Eight of the items address activities of daily living typically impacted by spinal conditions with 2 items addressing pain. Each item asks the patient to rate their perceived ability to perform a function ‘right now’ ranked on a 5 point scale anchored as 0 = full ability to function/no pain and 4 = unable to perform function at all/worst possible pain. The scale is scored by summing all items/40 x 100 to obtain a percent functional disability.<sup>72</sup>
- **Headache Disability Inventory (HDI)** – The HDI is a 25-item tool with 2 subscales including 12 emotion and 13 functional questions. There are 3 possible responses: “always” (4 points each), “sometimes” (2 points), and “never” (0 points).<sup>73</sup> An on-line source for the scale could not be found, but numerous sources for the instrument and information can be found by searching online using the term Headache Disability Inventory.
- **Neck Disability Index (NDI)** – Templated on the ODI, the NDI includes 10 questions addressing pain and pain interference on common ADLs. It is scored similarly to ODI and has been validated for common neck problems.<sup>15, 74</sup>  
**Whiplash Disability Questionnaire/Index (WDQ)** – Includes 13 numerical scale questions that address functional limitation following neck whiplash injury. The WDI addresses pain level, abilities with personal care, work, home and leisure activities, transportation, sleep, fatigue, and psychosocial factors (depression, anger, anxiety, and concentration). Each question’s numerical answer (0-10) is summed for a total of up to 130 points, with a higher score indicating greater disability. Minimal detectable change is about 15 points, but 30-50% change is considered clinically meaningful.<sup>75</sup> An on-line source for the scale could not be found, but numerous sources for the instrument and information can be found by searching online using the term Whiplash Disability Index.<sup>76</sup>

#### Thoracic/Chest

- **Bournemouth Questionnaire – Back (BQ-Back)** – ODI and RMQ scales focus primarily on low back conditions. Although not yet specifically validated for the thoracic and/or chest wall region, the BQ-back, is readily tailored to assessing problems in these areas.

#### Low Back

- **Bournemouth Questionnaire – Back (BQ-Back)** – Based on the function questions of the BQ, the BQ-Back was modified



specifically for back complaints. The BQ-Back includes 7 items: 1 for pain, 3 for pain-interference on ADLs, and 3 for psychosocial factors (anxiety, depression, and locus of control).<sup>12, 13</sup>

- **Functional Rating Index (FRI)** – The FRI is a 10-item scale (based on elements from the NDI and ODI) that has been validated for neck and low back conditions. Eight of the items address activities of daily living typically impacted by spinal conditions with 2 items addressing pain. Each item asks the patient to rate their perceived ability to perform a function ‘right now’ ranked on a 5 point scale anchored as 0 = full ability to function/no pain and 4 = unable to perform function at all/worst possible pain. The scale is scored by summing all items/40 x 100 to obtain a percent functional disability.<sup>72</sup>
- **Modified Oswestry Low Back Disability Index (ODI)** – The ODI measures disability and functional limitation related to back pain. It includes 10 questions addressing pain level, pain interference with ADL, sleep, etc. The original version includes a question on sex life which has been replaced in the modified version by a question on employment and homemaking. The ODI has been validated and is commonly used in clinical and research settings.<sup>21, 22, 77</sup>
- **Roland Morris Low Back Disability Index (RMQ)** – The RMQ has 24 statements regarding activities that are limited by the patient's low back pain. The patient marks each statement that describes their limitation. Positive statements are summed. A higher score indicates greater disability with scores over 13 points considered “high disability”. It has been validated in numerous studies, but meaningful change requires larger differences in those with higher initial scores.<sup>21, 23, 24, 78, 79</sup>

## Upper Extremity

### General

- **Disability of Arm, Shoulder, and Hand (DASH) Scale** – A 30-item, self-report scale addressing physical function and symptoms associated with common upper extremity disorders. It has good clinometric properties and includes a work component. It has been used increasingly as an outcome measure for upper limb pathology, especially in research studies. It assesses entire upper limb function including elbow and hand. Reliability and reproducibility have been demonstrated in several studies.<sup>80</sup>  
<http://www.dash.iwh.on.ca/conditions.htm>.
- **QuickDASH** – The QuickDASH is an easier-to-use, 11-question version of the full DASH that measures somewhat different content. It includes 4 additional questions on work and 4 questions on sports. The QuickDASH is a validated measure of arm function, but is reported to be less specific than the DASH in the subdomains, especially in symptoms. It has also been reported to underestimate symptoms and overestimate disabilities. The QuickDASH can be recommended to save time to obtain a summary assessment of arm symptoms and function based on the score.<sup>81</sup> The Quick DASH is available for use with registration and may be obtained online without charge at <http://www.dash.iwh.on.ca/conditions.htm>.
- **Upper Extremity Functional Index (UEFI)** – A validated, one-page form that addresses general arm function with specific incorporation of activities that involve the elbow and wrist extensors and flexors.<sup>82</sup>
- **Upper Limb Functional Index (ULFI)** – A validated, one-page form that has been compared to the UEFI as well as the DASH questionnaire and is considered by the developers to be practical in clinical settings.<sup>83</sup> <http://www.tac.vic.gov.au/upload/UE.pdf>

### Shoulder

- **Shoulder Pain and Disability Index (SPADI)** – The SPADI is a valid measure to assess pain and disability in community-based patients reporting shoulder pain due to musculoskeletal pathology. It is not useful for initial differential diagnosis but appears sensitive to change especially for range of motion with adhesive capsulitis. Therefore, like the SST, its primary utility is to measure improvement over time with care. This instrument is not validated for diagnostic purposes nor comparing severity between different individuals, rather how a patient's pain and function changes over time. SPADI has the ability to distinguish change in pain and function separately. Results for test-retest reproducibility indicated a small detectable difference of 17 points on the 1-100 scale, and on intra-class correlation coefficient of 0.89. The SPADI was generally more responsive than standard ROM testing. When compared to three other diagnostic questionnaires (Dutch Shoulder Disability Questionnaire (SDQ-NL), United Kingdom Shoulder Disability Questionnaire (SDQ-UK), Shoulder Rating Questionnaire (SRQ)), the SPADI was found to be valid with similar patient

acceptability, but most responsive to change and the quickest to complete. When compared to the Croft Index and the DASH in adhesive capsulitis patients, the SPADI was found to be valid and responsive with a slight advantage over other questionnaires. The VAS scale was found to be the best performing generic measure in terms of responsiveness in the patient group.

<http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/ShoulderPracticeResourceFinalapproved.pdf>

- **Simple Shoulder Test (SST)** - A 12-question shoulder activity scale developed at the University of Washington that has high patient utility. It is highly reliable across age groups and is sensitive to change. This instrument captures the patient's perception of how well they function. Its primary utility is to measure improvement over time with care. It also has the advantage of being free of licensing fees. <http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/ShoulderPracticeResourceFinalapproved.pdf>

## Elbow

- **Patient-Rated Elbow Evaluation (PREE)** – A 20-item questionnaire using numerical scales (0=no pain or difficulty – 10=worst pain, unable to do) to assess pain (5 items) and function (11 specific activities and 4 usual activities). The scales are scored as a pain subscale (sum the 5 items up to 50 points); a function subscale (sum the 15 function items and divide by 3 for up to 50 points). The total score can be reported as a 100 point scale. The tool has been validated in both surgical and non-surgical settings.<sup>84, 85</sup> [http://www.srs-mcmaster.ca/Portals/20/pdf/research\\_resources/PREE\\_UserManual\\_Dec2007.pdf](http://www.srs-mcmaster.ca/Portals/20/pdf/research_resources/PREE_UserManual_Dec2007.pdf)

## Wrist/Hand

- **Patient-Rated Wrist Evaluation (PRWE)** – A 15-item numerical scale (0=no pain or difficulty – 10=worst pain, unable to do) including 5 questions on pain frequency & intensity and 10 addressing function with specific and usual activities. The scales are scored as a pain subscale (sum the 5 items up to 50 points); a function subscale (sum the 10 function items and divide by 2 for up to 50 points). The total score can be reported as a 100 point scale. The tool has been validated. [http://www.srs-mcmaster.ca/Portals/20/pdf/research\\_resources/PRWE\\_PRWHEUserManual\\_Dec2007.pdf](http://www.srs-mcmaster.ca/Portals/20/pdf/research_resources/PRWE_PRWHEUserManual_Dec2007.pdf)

## Lower Extremity

### General

- **Lower Extremity Functional Scale (LEFS)** – A 20-question numerical scale (0= extreme difficulty – 4=no difficulty) addressing functional limitation of everyday activities and positions with the lower extremity. Activities include sitting, standing, walking, squatting, running, hopping, stair-climbing, moving in bed, bathing, and dressing. The indicated values of each item are summed for a total of up to 80 points (higher being less difficulty). The scale has been validated against the SF-36 with minimal detectable change reported as 9 points.<sup>40</sup> The LEFS appears to correlate with the Anterior Knee Pain Scale and the WOMAC hip osteoarthritis questionnaire.<sup>86</sup> Meaningful change may be considered similar for other instruments at 30-50%.<sup>17</sup> Each of the 20 question's 5 possible responses are scored on an ascending scale (0 = Extreme Difficulty, 4= No Difficulty). Points are summed for a maximum possible score of 80. Lower score means worse function.
- **Lower Limb Outcome Questionnaire (LLOQ)** – The LLOQ was developed by the American Academy of Orthopedic Surgeons and numerous other orthopedic organizations. It is made up of 7 items addressing symptoms and activities related to the lower extremity over the previous week. Test-retest reliability within 24 hours of re-administration has been reported as well as comparability to SF-36 measures.<sup>87</sup> The instrument and a scoring worksheet is available online: [http://www.aaos.org/research/outcomes/outcomes\\_list.asp#low](http://www.aaos.org/research/outcomes/outcomes_list.asp#low)

### Hip & Knee

- No universal disability scales appear to be validated for multiple different hip or knee conditions; however, several condition-specific

scales for each joint (see below) have been reported to have good clinometric properties.<sup>88, 89</sup>

## Foot & Ankle

**Foot and Ankle Ability Measure (FAAM)** – A revised version of the FADI, including the sports subscale, with a few questions modified or removed to improve the survey's psychometric properties.<sup>37, 90</sup> Each item is scored on a five point scale with 4 being "No Difficulty" and 0 being "Unable To Do." The lowest potential score of the Activities of Daily Living (ADL) subscale of the FAAM is 0 points, the highest 84 points. The lowest potential score of the Sports subscale of the FAAM is 0 points, the highest 32 points. Total score is converted into percentage. Higher percentage indicates higher level of physical function.<sup>39</sup>

- **Foot and Ankle Disability Index (FADI)** – A one-page scale with 26 elements of routine daily activities, each rated on a 5 point difficulty or pain level scale. In addition, an optional sport module addresses 8 elements associated with common athletic activities. The scale has been validated and appears especially useful for ankle instability.<sup>37, 90, 91</sup> <http://www.middleburg-pt.com/pdfs/fadi.pdf>
- **Foot Function Index (FFI)** - Developed to measure the impact of foot pathology on function in terms of pain, disability and activity restriction.<sup>92</sup> An on-line source for the scale could not be found, but numerous sources for the instrument and information can be found by searching online using the term Foot Function Index.

## CONDITION SPECIFIC SCALES

### Lateral Epicondylitis

**Patient-Rated Tennis Elbow Evaluation (PRTEE)** – The PRTEE was validated specifically for lateral epicondylitis and is a straightforward, one-page questionnaire easily administered in clinical settings. Refer to the IICAC Work-Related Epicondylitis Practice Resource for additional information: <http://www.lni.wa.gov/ClaimsIns/Files/OMD/LEResourceFINAL.pdf>

### Carpal Tunnel Syndrome

- **Carpal Tunnel Syndrome Assessment Questionnaire (CTSAQ); also known as the Boston Carpal Tunnel Questionnaire** – A self-administered symptom severity questionnaire that has been used in population-based research trials for which psychometric properties have been validated. It includes symptom severity and function subscales. It has demonstrated sensitivity to pre- and post-surgery changes in self-reported severity of wrist symptoms and several basic activities of daily living. It does not appear to have been correlated to NCV findings and does assess typical work tasks or durations.<sup>93, 94</sup> An on-line source for the scale could not be found, but numerous sources for the instrument and information can be found online using the search term "CTSAQ."
- **Katz Hand Diagram** – A self-administered diagram of the dorsal & palmar hand. The patient marks the locations of pain, numbness, tingling or decreased sensation. It is used primarily for diagnosis based on symptom distribution marked by the patient. Refer to the IICAC Occupational Carpal Tunnel Syndrome Practice Resource for additional information: <http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/FinalConsCTSSummary41.pdf>

### Knee Ligament Tears & Chondral Defects

**Lysholm Scale** – An 8-item questionnaire developed to evaluate patients following knee ligament reconstruction. It has been validated for ligament tears and chondral defects.<sup>95-97</sup> The 100-point scale measures knee stability (25 points), pain (25 points), locking (15 points), swelling and stair climbing (10 points each), and limping, use of support, and squatting (5 points each). Scoring: <65 Poor, 65-83 Fair, 84-90 Good, >90 Excellent.<sup>98</sup> <https://cours.etsmtl.ca/qts813/Documents/Lysholm.pdf>

### Knee Osteoarthritis

**Knee Osteoarthritis Outcome Score (KOOS)** – A 42-item scale addressing knee pain and symptoms, their impact on activities of daily living, sports & recreation, and quality of life. Various domains are scored separately and also in summation. Each section score is multiplied into a percentage and reversed so that a lower score means worse function.<sup>99-101</sup> <http://www.koos.nu/>

Anterior Knee Pain

**Anterior Knee Pain Scale (AKPS)** – A 13-item questionnaire in multiple choice format with simple topics such as walking, running, and jumping as well as more clinically sophisticated topics such as ‘atrophy of thigh’ and ‘flexion deficiency.’ Each response has a certain number of points that are summed to achieve the score. Lower scores mean worse pain and function. The AKPS has been compared to other scales such as the LEFS and although it is a valid and reliable measure, it does not appear to be superior.<sup>86, 102</sup> The LEFS may be preferable for regular use in general practice considering that it can be used for a broader range of joints and conditions. An on-line source for the scale could not be found, but numerous sources for the instrument and information can be found by searching online using the term “AKPS.”

Achilles  
Tendinopathy

**Victorian Institute of Sport Assessment - Achilles Questionnaire (VISA-A)** – An 8-question scale covering domains of pain, function, and activity validated for severity against two other clinical severity measures<sup>103</sup> and reported reliable in a well done systematic review.<sup>104</sup> The first 7 questions are numerical scales (0-10) scored by summing the values indicated by the patient. The last question is valued at 30 points and one of three different options based on the intensity of the pain as selected and filled out by the patient. <http://bjsm.bmj.com/content/35/5/335.full>

Osteoarthritis

**WOMAC Osteoarthritis Index** – The WOMAC is a disease-specific, self-administered questionnaire used with patients who have hip or knee osteoarthritis. It is most commonly used for assessing progress following total hip or knee arthroplasty. It contains a multi-dimensional scale made up of 24 items grouped into three dimensions: pain (5 items), stiffness (2 items), and physical function (17 items). Each item is scored 0-4 (none, mild, moderate, severe, extreme). Score: 0-100 (0 being best to 100 being worst).<sup>105-112</sup> <http://www.womac.org/womac/index.htm>

## ROUTINE USE ANCHORED/NUMERICAL SCALES

Pain Scales

- **Anchored Numerical Scale** – Endpoints are typically anchored to using an 11-point scale (0-10). The patient circles a number indicating their pain level with the circled number becoming the score (with higher scores reflecting more pain). Numerical scales may ask about the level of pain at the time of filling it out, or request an average over a particular time period (the past day, past week, etc).<sup>45</sup> Example:

On average, how would you rate your pain during the past week?

**No  
Pain**

0    1    2    3    4    5    6    7    8    9    10

**Worst Possible  
Pain**

- **Anchored Pain Interference Scale** - Specific attention to how a patients’ pain interferes with their ability to perform usual activities has been shown to be useful in predicting chronicity for low back and other musculoskeletal problems, particularly in injured worker populations.<sup>113-115</sup> Pain interference is combined with pain severity in the Graded Chronic Pain Scale (GCPS). Example:

On average how much does your pain interfere with your ability to do your usual daily activities?

**I can do all  
usual activities**

0    1    2    3    4    5    6    7    8    9    10

**Unable to do any  
usual activities**

- **Anchored Symptom Scale** – Similar in concept to PSFS but for routine visit-to-visit assessment, a particular symptom associated with the patient’s condition is anchored to a relevant metric or context. For example, time or distance can be an anchor and visit-to-visit change can be captured: “How many minutes can you type at a time until numbness returns?” or “How far can you walk until the

pain becomes unbearable?”

- **Graded Chronic Pain Scale – Two-Item Pain Assessment** – Essentially a combined numerical pain intensity and interference scale over the past month that has been validated in a chronic pain setting.<sup>114</sup> It has also become a standard for quickly and routinely tracking pain and function in the Washington State Agency Medical Directors Group Opioid Dosing Guideline for Chronic Non Cancer Pain: <http://www.agencymeddirectors.wa.gov/opioiddosing.asp>

**In the last month**, on average, how would you rate your pain? Use a scale from 0 to 10, where 0 is “no pain” and 10 is “pain as bad as could be” *[That is your usual pain, at times you were in pain]*

<b>No Pain</b>											<b>Pain as bad as could be</b>
0	1	2	3	4	5	6	7	8	9	10	

**In the last month**, how much has pain interfered with your daily activities? Use a scale from 0 to 10, where 0 is “no interference” and 10 is “unable to carry on any activity”

<b>No interference</b>											<b>Unable to carry on any activities</b>
0	1	2	3	4	5	6	7	8	9	10	

- **Pain in Multiple Body Sites** – Pain in multiple body sites has been shown to be a strong and consistent risk factor for chronic pain and disability.<sup>116</sup> A simple checklist of different body parts (arms, legs, etc.) and instructions to check if the person has had persistent bothersome pain in this body part in the past month (or 6 months) is an easy way to capture this information. Example:

Please check any areas where you have had persistent, bothersome pain in the past 6 months:

- |  |  |
|--|--|
| <input type="checkbox"/> Low Back                                  | <input type="checkbox"/> Shoulder(s)         |
| <input type="checkbox"/> Head                                      | <input type="checkbox"/> Neck                |
| <input type="checkbox"/> Arms/Hands                                | <input type="checkbox"/> Abdomen/Pelvic Area |
| <input type="checkbox"/> Hips/Buttocks                             | <input type="checkbox"/> Legs/Feet           |
| <input type="checkbox"/> Chest/Rib Cage                            | <input type="checkbox"/> Upper/Mid Back      |
| <input type="checkbox"/> No areas with persistent, bothersome pain |  |

- **Self-Control of Pain Scale** – Self-control (locus of control) over pain reflects the coping capacity a patient might have with their condition. Poorer coping capacity has been associated with chronicity.<sup>117</sup>

**In the past week**, on average, how much control were you able to have over your pain? Use a scale from 0 to 10, where 0 is “complete control of your pain” and 10 is “no control of your pain”

<b>Complete control of your pain</b>											<b>No control of your pain</b>
0	1	2	3	4	5	6	7	8	9	10	

- **Visual Analog Scale (VAS)** – Virtually replaced by numerical scales, analog pain scales use a fixed distance line anchored on one end by “No Pain” and the other end by a descriptor such as “Disabling Pain.” The hash mark made by the patient on the scale is measured and usually reported in millimeters.

## Function Scales

- **Anchored Function Scale** – Similar in concept to the PSFS and an Anchored Symptom Scale for routine visit-to-visit assessment. The focus of the anchors is related to activity goal setting with the patient selecting a particular activity that is impacted by the condition. A relevant anchor/context is selected by the patient capturing numbers of repetitions, minutes, or distance the activity is engaged in. This kind of scale is typically used to set incremental goals for increasing capacity, but serves to track progress as well.

## PHYSICAL CAPACITY MEASUREMENT

### Physical Performance Tests

Physical Performance Tests (PPT) typically include strength, coordination, and endurance tests that can be easily performed in office settings. The batteries and tests included here are simple to administer requiring only chairs, exam/treatment tables, some form of strapping or supportive restraint, a goniometer, and a stopwatch. Normative data is included in general terms based on published reports where available and if highly variable by age or gender, is indicated as such. This serves as a guide for what to expect, but it should be noted that even though data may be reported in fine measures like seconds or fractions of seconds, there is great variation across individuals. The most important feature of these tests is the ability to assess recovery (or lack thereof) when a patient's performance improves (or stagnates/worsens) over time. Like most clinical examination procedures, very few physical performance tests have been adequately validated, thus they should not be considered precise tools. As a rule, baseline performance testing (for outcomes tracking) might be considered if recovery is not meaningfully evident. In typical work injury situations, they should be only be considered after at least two weeks following initiation of a care program. Generally, PPTs can help identify underlying conditioning issues that not only impede recovery but may be worth addressing to facilitate injury/aggravation-free return-to-work.

This resource does not specifically address treatment issues; however, activity is important in nearly all musculoskeletal injury recovery. Active care should include incrementally increasing daily activities as soon as they can be tolerated with more emphasis on specific exercises as recovery occurs. Referral for more structured exercise/conditioning programs typically would not be considered before 4-6 weeks of home-based exercises and/or, when clinically meaningful improvement in outcomes assessment measures is not obtained. An IICAC Conservative Care Practice Resource is available for rehabilitation of work-related low back conditions which reviews and summarizes relevant evidence: <http://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/ActiveRehabWkRelatedLowBackCond.pdf>

- **Back Physical Performance Battery (BPPB)**- No formal structured “battery” of in-office physical performance tests for common work injuries has been validated in the literature. However, several individual tests for back strength and endurance have been described and are commonly used in rehabilitation settings. Assessing physical function and which basic activities associated with back strength may provoke symptoms are worth documenting and measuring, especially if higher frequency care is continuing beyond 4 weeks and/or return to work is not imminent by that time. In general, tests can be performed in-office with minimal equipment and are scored by time or repetitions as described below. Improvement may be graded and assessed by increasing capacity as measured by time and/or repetition scores, however an 85% pass-fail approach provides a simple method to document performance. As an option, consider lowering the cut-off to 70% (or less) for those >50 years old, and/or significantly debilitated / deconditioned at any age.
  - **Static Back Endurance (SBE)** – The patient lies prone, trunk extended off the edge of a bench with anterior superior iliac spines on the table edge. Arms remain at sides with ankles, thighs and buttocks strapped to the bench. The patient should hold the static, neutral, horizontal position until fatigue or 240 seconds (whichever comes first). There are several minor variations for performing SBE as well as different strategies for scoring and interpretation. The pass-fail method is recommended.
    - **Pass – Fail Method:** Based on average normative data,<sup>118</sup> middle-aged working males should be able to hold position for  $97 \pm 53$  seconds and middle-aged working females for  $87 \pm 59$  seconds. Using an 85% of norm as passing for under 50 years old and 70% for over 50 years old, the following cutoffs are recommended:



Age	Male	Female
Norm	97	87
< 50	82 secs	74 secs
> 50	68 secs	61 secs

For low back conditions, static extensor endurance tests appear to be the most useful in terms of sensitivity, specificity, and predictive value for low back conditions.<sup>119-121</sup> Poor static endurance (less than 58 seconds in both males & females) appears to be associated with increased risk of low back pain at 1 year follow-up.<sup>118</sup> Additionally, decreased extensor endurance is associated with back pain in workers and otherwise healthy individuals.<sup>122-125</sup>

- **One Leg Balance (Proprioception Test)** – The patient stands on one leg with eyes open. Time is measured in seconds for a maximum of 30 seconds or when the patient loses balance (reaches out, hops, touches floor with non-weight-bearing foot). The test is repeated with eyes closed. Based on normative data by age<sup>126</sup> and using an 85% cutoff, the test can also be scored as pass-fail:

Age	Eyes Open	Eyes Closed	Pass (85%)
20-59	30 secs	25 secs	21 secs
60-69	22 secs	10 secs	8.5 secs
70-79	14 secs	4 secs	3 secs

- **Side Bridge** – Patient lays on their side propped up on one elbow with top ankle crossed in front of bottom ankle. Hips are then lifted up and held in alignment so that the weight is supported only by the feet and elbow. The length of time the position is held is recorded. This test assesses the core stabilization strength of primarily the quadratus lumborum muscles. In normal individuals, the position should be able to be held for more than 95 seconds in men and 75 seconds in women on each side without difficulty.<sup>127</sup> Using an 85% of norm as passing for under 50 years old and 70% for over 50 years old, the following cutoffs are recommended:

Age	Male	Female
Norm	95 sec	75 sec
< 50 (85%)	81 secs	64 secs
> 50 (70%)	67 secs	53 secs

- **Trunk Stabilizer Strength** – Both squatting and abdominal strength reflect important aspects in core or trunk stability. Squatting assesses hip, knee and ankle mobility as well as strength, endurance and coordination of hip and knee extensors. Sit-ups primarily assess some basic lumbar mobility and strength and endurance of the rectus abdominal muscles. Some authorities recommend performing both squatting and sit up tests, but performing only one will provide a sense of trunk stabilizer condition and may be preferable in certain patients.

- **Repetitive Squat:** Patient stands with feet 15 cm apart and squats down until the thighs are parallel to the floor, then returning to the upright position in a 2-3 second cycle. Squats are repeated until fatigue or about 50 repetitions are achieved and the number is recorded. Fatigue may be considered reached when difficulty to complete cycle impacts quality of movements. Based on normative data,<sup>128</sup> middle aged males should be able to complete 37 repetitions and females should be able to complete 21. Using an 85% of norm as passing for under 50 years old and 70% for over 50 years old, the following cutoffs are recommended:

Age	Male	Female
Norm	37 reps	21 reps

< 50 (85%)	31 reps	18 reps
> 50 (70%)	26 reps	15 reps

- **Repetitive Sit-up:** In a supine, knees flexed 90 degrees position with ankles supported, the patient performs a partial sit-up (reaching with arms extended until the thenar pad approximates the superior pole of the patella) over a 2-3 second cycle time. The test is performed until fatigue or 50 repetitions is achieved. Based on normative data, middle aged men should be able to perform 27 and woman 19 repetitions. Using the 85% pass approach, passing for men can be considered 23 and 16 repetitions for women.

Age	Male	Female
Norm	27 reps	19 reps
< 50 (85%)	23 reps	16 reps
> 50 (70%)	19 reps	13 reps

- **Hamstring Length (Straight Leg Raising)** – This test assesses both mobility of the hip joint and flexibility (length and/or tension) of the hamstring muscles of the upper leg. Flexibility and/or of these muscles may indicate overall conditioning and stability of the lower extremity and trunk core. This test is performed passively with the examiner supporting the lower leg and raising the straight leg to the point of knee flexion (on the raised leg) or when the pelvis or opposite knee begin to move. The distance the raised leg moves is recorded in degrees (using an inclinometer zeroed out on the table and measured on the mid-tibia. Average flexion has been reported to be about 80 degrees.<sup>129</sup> Using an 85% of norm as passing for under 50 years old and 70% for over 50 years old, the following cutoffs are recommended:

Age	Male	Female
Norm	95 degrees	75 degrees
< 50 (85%)	81 degrees	64 degrees
> 50 (70%)	67 degrees	53 degrees

- **Short Physical Performance Battery (SPPB)** – Three well-validated timed tests aimed at assessing basic function in older patients. They are easily performed in office settings and include the following: Sequential Balance Tests; Gait Speed Test; Chair Stand Test. They require a stop watch, a marked 4 meter straight walking course, and a straight backed chair placed against a wall. Total score is the sum of each individual test for a maximum of 12 points.<sup>126</sup>

- **Sequential Balance Tests** – These assess the patient's three basic standing positions with the eyes open:
  - Side-by-side stand (stand with feet side by side for 10 seconds). If successful, score 1 point and move on to:
  - Semi-tandem stand (stand with inside heel of one foot next to inside of big toe on opposite foot for (10 seconds). If successful add 1 additional point and move on to:
  - Full tandem stand (heel of one foot is placed fully in front of toes of opposite foot for 10 seconds). Add 2 points for patients holding this position for 10 full seconds; 1 point for 3-9.9 seconds; no additional points for <3 seconds.
- **Gait Speed Test** – The patient is timed twice, walking at usual speed for 4 meters. The faster time is used for scoring; >8.20 seconds = 1 point; 6.21-8.20 seconds = 2 points; 4.82-6.20 seconds = 3 points; <4.82 seconds = 4 points
- **Chair Stand Test** – This assesses the patient's ability to rise from a chair with arms folded across chest. If the patient cannot perform, the score is zero. If it can be performed, the patient should perform five complete rises and reseatings as quickly as he can. Time is measured from command to stand to last rise; 16.70-60 seconds = 1 point; 13.7-16.69 seconds = 2 points; 11.20-13.69 = 3 points; < 11.20 second = 4 points.

- **Static Neck Endurance** – The patient lays supine with knees bent. Patient tucks chin towards chest and lifts head off table 1 inch holding until fatigue (dropping of head). The time in seconds is recorded. Neck flexor muscle endurance was reported to be both

Physical Capacity  
Testing Instruments/  
System

statistically and clinically greater in subjects without neck pain than those with neck pain.<sup>130</sup> Neck endurance also appears to be a predictor of future neck pain.<sup>122, 123, 131</sup> There are published variations of this test, including using sphygmomanometers to measure force of cervical flexion, however, timed static testing is simplest for routine in-office use. Based on unpublished normative data, males without neck pain should be able to sustain flexion for 85 seconds until fatigue and females should be able to hold the position for 39 seconds.<sup>132</sup> Using an 85% of norm as passing for under 50 years old and 70% for over 50 years old, the following cutoffs are recommended:

<b>Age</b>	<b>Male</b>	<b>Female</b>
Norm	85 secs	39 secs
< 50 (85%)	72 secs	33 secs
> 50 (70%)	60 secs	27 secs

Using higher technology physical capacity measurement does not appear to offer any advantages over self-administered functional scales and low-tech physical performance for assessing improvement in the early phases of injured worker care. Further, a Cochrane Library review was unable to find any studies comparing re-injury rates for workers receiving functional capacity evaluation to workers not having the intervention.<sup>133</sup>

## EVIDENCE & METHODOLOGY

### Intervention/Experimental Studies

**Randomized Controlled Trial (RCT)** – A study that randomly allocates patients to treatment groups, usually blinding patients, therapists and/or study evaluators. Typically of high quality as randomization assures similarities of subjects within treatment groups.

### Observational Studies

**Cohort Design** – Cohort (retrospective or prospective) – A study that follows patients who self-allocate to treatment groups through the course of their care for a given occurrence of a condition. Larger, well-designed cohort studies may be of good quality, but lack of randomization predisposes to heterogeneity issues within groups, some of which may be able to be adjusted for with statistical methods.

**Cross sectional** – Involves observing a population to measure disease and exposure status. It is usually thought to be a “snapshot” of the frequency and characteristics of a disease in a population at a specific given time.

**Case control** – Is a study that compares patients who have an outcome (cases) of interest with patients who do not have the disease or outcome (controls). The study may retrospectively to compare how frequently the exposure was present in a group to determine risk factors.

**Case series** – Is a study that describes a series of patients with an outcome of interest, may be of variable quality. Better designs use consecutive patients and include robust baseline and follow up outcome measures.

**Case reports** – Describes an individual case, typically only achieving publication if it represent a unique or unusual clinical experience.

### Blinding

Blinding minimizes potential bias. Typically three levels of blinding are sought: patient, treating provider and evaluator. Many conservative interventions do not allow for patient blinding (e.g. someone is likely to know if they received a splint or a pill). At a minimum, single blinding of the evaluator as to what group a subject was in is expected.

### Literature Reviews

**Quantitative systematic reviews** – Studies that review previously published clinical trials that include quantitative comparisons (e.g. meta-analyses). Systematic reviews should have rigorous and comprehensive methodology to identify relevant published research and include appraisal of study quality. Cochrane reviews frequently are of this type.

**Qualitative systematic reviews** – Similar to quantitative reviews but without systematic quantitative comparison or data pooling.

**Narrative literature reviews** – Such reviews typically do not include rigorous study selection methodology and may be subject to significant author bias

### Literature Retrieval and Review

1. **Initial systematic searches** of electronic databases (e.g. PubMed). Search terms used typically included MeSH terms for tests and interventions with conditions being addressed. Follow-up searches also included population attributes (e.g., workers compensation, occupational).
2. **Abstract screening** for relevance.
3. **Original paper retrieval** with review for relevance, quality, outcome meaningfulness, and effect magnitude.
4. **Additional studies identified** through clinical summaries (e.g., reviews, texts), citation tracking, and feedback from public.

### About Evidence for Physical Examination and Conservative Interventions

Conservative musculoskeletal care is typically care of first resort based on long standing practices. Typically ‘low tech,’ low cost, with minimal and rare side effects, it is frequently delivered in primary care settings, and by various health providers. The rigor and quality expected of high cost, higher risk, emerging, and tertiary interventions is less common for many routine physical examination procedures and conservative interventions. Much of the evidence summarized here would be considered Class “C” or “III” in ratings systems. Thus, the committee has not presented explicit *recommendations*, rather, *evidence summaries* guided by expert consensus to assist in formulating care options. Further, significant emphasis is made regarding tracking and documenting meaningful functional improvement with patients. Study attributes most likely to strengthen or limit confidence are characterized in the evidence descriptions.

### Assessing Study Methodologic Quality

Attributes of study methodology quality vary according to the clinical procedure (eg, diagnostic, therapeutic intervention) looked at, and specific research questions being studied. The American Academy of Neurology’s Clinical Practice Guideline Process Manual<sup>134</sup> offers a comprehensive guide to systematic evidence review, quality attributes and consensus process that generally serves as the approach taken by IICAC.

General attributes identified when extracting evidence from studies include identification of population, the intervention and co-interventions and outcomes being addressed in each study. The clinical questions addressed such as diagnostic accuracy, therapeutic effectiveness, or causation are determined. Studies are extracted into evidence tables including quality attributes and/or ratings which are reviewed both by department staff and committee members (usually 2 per study).

Specific quality attributes include: Diagnostic Accuracy – design, spectrum of patients, validity and relevance of outcome metric; Therapeutic Interventions – comparison groups (no treatment, placebo, comparative intervention), treatment allocation, blinding/masking (method and degree: single, double, independent), follow-up (period and completion), and analysis (statistical power, intent-to-treat). Specific attention is paid to several factors including reporting of outcomes (primary vs. secondary), relevance of outcome (eg, function vs. pain), and meaningfulness (clinically important change vs minimally detectable change).

### Synthesizing Evidence

Consideration of study quality (class), significance (statistical precision), consistency across studies, magnitude of effect, and relevance to populations and procedures were taken into account in preparing draft summaries. Special attention was given to clarifying conclusions related to the clinical questions of interest. Evidence, particularly with low tech and highly diffused examination and conservative procedures addressed here, is rarely truly “definitive,” even when multiple studies exist. Inconsistent conclusions typically reflect error (systematic, random) and/or bias in studies. Data pooling via meta-analysis is useful to reduce random error when studies are of sufficient power and methodologic strength. Larger meaningful effect size may increase confidence in findings.

## CITATIONS

1. Stratford, P., Gill, C., Westaway, M.D., and Binkley, J., *Assessing disability and change on individual patients: a report of a patient specific measure*. Physiotherapy Canada, 1995. **47**(4): p. 258-263.
2. Westaway, M.D., Stratford, P.W., and Binkley, J.M., *The patient-specific functional scale: validation of its use in persons with neck dysfunction*. J Orthop Sports Phys Ther, 1998. **27**(5): p. 331-8.
3. Waddell, G. and Newton, M., *A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability*. Pain, 1993. **52**: p. 157-168.
4. Fritz, J.M. and George, S.Z., *Identifying psychosocial variables in patients with acute work-related low back pain: the importance of fear-avoidance beliefs*. Phys Ther, 2002. **82**: p. 973-983.
5. Fritz, J.M., Beneciuk, J.M., and George, S.Z., *Relationship between categorization with the STarT Back Screening Tool and prognosis for people receiving physical therapy for low back pain*. Phys Ther, 2011. **91**(5): p. 722-32.
6. Hill, J.C., Dunn, K.M., Lewis, M., Mullis, R., Main, C.J., Foster, N.E., and Hay, E.M., *A primary care back pain screening tool: identifying patient subgroups for initial treatment*. Arthritis Rheum, 2008. **59**(5): p. 632-41.
7. Hill, J.C., Dunn, K.M., Main, C.J., and Hay, E.M., *Subgrouping low back pain: a comparison of the STarT Back Tool with the Orebro Musculoskeletal Pain Screening Questionnaire*. Eur J Pain, 2010. **14**(1): p. 83-9.
8. Kongsted, A., Johannesen, E., and Leboeuf-Yde, C., *Feasibility of the STarT back screening tool in chiropractic clinics: a cross-sectional study of patients with low back pain*. Chiropr Man Therap, 2011. **19**: p. 10.
9. Murphy, D.R. and Hurwitz, E.L., *The usefulness of clinical measures of psychologic factors in patients with spinal pain*. J Manipulative Physiol Ther, 2011. **34**(9): p. 609-13.
10. Woby, S.R., Roach, N.K., Urmston, M., and Watson, P.J., *Psychometric properties of the TSK-11: a shortened version of the Tampa Scale for Kinesiophobia*. Pain, 2005. **117**(1-2): p. 137-44.
11. Liebenson, C. and Yeomans, S., *Assessment of psychosocial risk factors of chronicity- "yellow flags"*. in *Rehabilitation of the Spine: A Practitioners Manual*, C. Liebenson, Editor 2007, Lippencott Williams & Wilkins: Baltimore. p. 183-202.
12. Bolton, J.E. and Breen, A.C., *The Bournemouth Questionnaire: a short-form comprehensive outcome measure. I. Psychometric properties in back pain patients*. J Manipulative Physiol Ther, 1999. **22**(8): p. 503-10.
13. Hurst, H. and Bolton, J., *Assessing the clinical significance of change scores recorded on subjective outcome measures*. J Manipulative Physiol Ther, 2004. **27**(1): p. 26-35.
14. Bolton, J.E., *Sensitivity and specificity of outcome measures in patients with neck pain: detecting clinically significant improvement*. Spine 2004. **29**(21): p. 2410-7; discussion 2418.
15. Vernon, H. and Mior, S., *The Neck Disability Index: a study of reliability and validity*. J Manipulative Physiol Ther, 1991. **14**(7): p. 409-15.
16. Stratford, P., Riddle, D.L., and Binkley, J., *Using the neck disability index to make decisions concerning individual patients*. Physiotherapy Canada, 1999. **51**(2): p. 107-112.
17. Ostelo, R.W., Deyo, R.A., Stratford, P., Waddell, G., Croft, P., Von Korff, M., Bouter, L.M., and de Vet, H.C., *Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change*. Spine, 2008. **33**(1): p. 90-4.
18. Davidson, M. and Keating, J.L., *A comparison of five low back disability questionnaires: reliability and responsiveness*. Phys Ther, 2002. **82**(1): p. 8-24.
19. Fairbank, J.C., Couper, J., Davies, J.B., and O'Brien, J.P., *The Oswestry low back pain disability questionnaire*. Physiotherapy, 1980. **66**(8): p. 271-3.
20. Fritz, J.M., Hebert, J., Koppenhaver, S., and Parent, E., *Beyond minimally important change: defining a successful outcome of physical therapy for patients with low back pain*. Spine 2009. **34**(25): p. 2803-9.
21. Lauridsen, H.H., Hartvigsen, J., Manniche, C., Korsholm, L., and Grunnet-Nilsson, N., *Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients*. BMC Musculoskelet Disord, 2006. **7**: p. 82.
22. Vianin, M., *Psychometric properties and clinical usefulness of the Oswestry Disability Index*. J Chiropr Med, 2008. **7**(4): p. 161-3.
23. Roland, M. and Morris, R., *A study of the natural history of low-back pain. Part II: development of guidelines for trials of treatment in primary care*. Spine, 1983. **8**(2): p. 145-50.
24. Roland, M. and 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): p. 141-4.
25. Stratford, P.W., Binkley, J.M., Riddle, D.L., and Guyatt, G.H., *Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1*. Phys Ther, 1998. **78**(11): p. 1186-96.
26. Von Korff, M. and Saunders, K., *The course of back pain in primary care*. Spine 1996. **21**(24): p. 2833-7; discussion 2838-9.
27. Beaton, D.E., Wright, J.G., and Katz, J.N., *Development of the QuickDASH: comparison of three item-reduction approaches*. J Bone Joint Surg Am, 2005. **87**(5): p. 1038-46.
28. Polson, K., Reid, D., McNair, P.J., and Larmer, P., *Responsiveness, minimal important difference and minimal detectable change of the shortened disability arm shoulder hand (QuickDASH) questionnaire*. Man Ther, 2010. **15**: p. 404-407.
29. Macdermid, J., Solomon, P., and Prkachin, K., *The shoulder pain and disability index demonstrates factor, construct and longitudinal validity*. BMC Musculoskelet Disord, 2006. **7**(12).
30. Roach, K.E., Budiman-Mak, E., Songsiridej, N., and Lertratanakul, Y., *Development of a shoulder pain and disability index*. Arthritis Care Res, 1991. **4**(4): p. 143-9.
31. Williams, J.W., Jr., Holleman, D.R., Jr., and Simel, D.L., *Measuring shoulder function with the Shoulder Pain and Disability Index*. J Rheumatol, 1995. **22**(4): p. 727-32.
32. MacDermid, J.C., Solomon, P., and Prkachin, K., *The Shoulder Pain and Disability Index demonstrates factor, construct and longitudinal validity*. BMC Musculoskelet Disord, 2006. **7**: p. 12.
33. Roddey, T.S., Olson, S.L., Cook, K.F., Gartsman, G.M., and Hanten, W., *Comparison of the University of California-Los Angeles Shoulder Scale and the Simple Shoulder Test with the shoulder pain and disability index: single-administration reliability and validity*. Phys Ther, 2000. **80**(8): p. 759-68.

34. Stratford, P., Binkley, J., and Stratford, D., *Development and initial validation of the upper extremity functional index*. Physiotherapy Canada, 2001: p. 259-266.
35. Lehman, L.A., Sindhu, B.S., Shechtman, O., Romero, S., and Velozo, C.A., *A comparison of the ability of two upper extremity assessments to measure change in function*. J Hand Ther, 2010. **23**(1): p. 31-39.
36. Carcia, C.R., Martin, R.L., and Drouin, J.M., *Validity of the Foot and Ankle Ability Measure in athletes with chronic ankle instability*. J Athl Train, 2008. **43**(2): p. 179-83.
37. Eechaute, C., Vaes, P., Van Aerschot, L., Asman, S., and Duquet, W., *The clinimetric qualities of patient-assessed instruments for measuring chronic ankle instability: a systematic review*. BMC Musculoskelet Disord, 2007. **8**: p. 6.
38. Goldstein, C.L., Schemitsch, E., Bhandari, M., Mathew, G., and Petrisor, B.A., *Comparison of different outcome instruments following foot and ankle trauma*. Foot Ankle Int, 2010. **31**(12): p. 1075-80.
39. Martin, R.L., Irrgang, J.J., Burdett, R.G., Conti, S.F., and Van Swearingen, J.M., *Evidence of validity for the Foot and Ankle Ability Measure (FAAM)*. Foot Ankle Int, 2005. **26**(11): p. 968-83.
40. Binkley, J.M., Stratford, P.W., Lott, S.A., and Riddle, D.L., *The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application*. North American Orthopaedic Rehabilitation Research Network. Phys Ther, 1999. **79**(4): p. 371-83.
41. Yeung, T.S., Wessel, J., Stratford, P., and MacDermid, J., *Reliability, validity, and responsiveness of the lower extremity functional scale for inpatients of an orthopaedic rehabilitation ward*. J Orthop Sports Phys Ther, 2009. **39**(6): p. 468-77.
42. Lin, C.W., Moseley, A.M., Refshauge, K.M., and Bundy, A.C., *The lower extremity functional scale has good clinimetric properties in people with ankle fracture*. Phys Ther, 2009. **89**(6): p. 580-8.
43. Childs, J.D., Piva, S.R., and Fritz, J.M., *Responsiveness of the numeric pain rating scale in patients with low back pain*. Spine, 2005. **30**(11): p. 1331-4.
44. van der Roer, N., Ostelo, R.W., Bekkering, G.E., van Tulder, M.W., and de Vet, H.C., *Minimal clinically important change for pain intensity, functional status, and general health status in patients with nonspecific low back pain*. Spine, 2006. **31**(5): p. 578-82.
45. de Vet, H.C., Ostelo, R.W., Terwee, C.B., van der Roer, N., Knol, D.L., Beckerman, H., Boers, M., and Bouter, L.M., *Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach*. Qual Life Res, 2007. **16**(1): p. 131-42.
46. Bolton, J.E. and Hurst, H.C., *Prognostic factors for short-term improvement in acute and persistent musculoskeletal pain consulters in primary care*. Chiropr Man Therap, 2011. **19**(1): p. 27.
47. Safran, D.G., Kosinski, M., Tarlov, A.R., Rogers, W.H., Taira, D.H., Lieberman, N., and Ware, J.E., *The Primary Care Assessment Survey: tests of data quality and measurement performance*. Med Care, 1998. **36**(5): p. 728-39.
48. Rodriguez, H.P., von Glahn, T., Rogers, W.H., Chang, H., Fanjiang, G., and Safran, D.G., *Evaluating patients' experiences with individual physicians: a randomized trial of mail, internet, and interactive voice response telephone administration of surveys*. Med Care, 2006. **44**(2): p. 167-74.
49. Bolton, J.E. and Humphreys, B.K., *The Bournemouth Questionnaire: a short-form comprehensive outcome measure. II. Psychometric properties in neck pain patients*. J Manipulative Physiol Ther, 2002. **25**(3): p. 141-8.
50. Dunstan, D.A., Covic, T., Tyson, G.A., and Lennie, I.G., *Does the Orebro Musculoskeletal Pain Questionnaire predict outcomes following a work-related compensable injury?* Int J Rehabil Res, 2005. **28**(4): p. 369-70.
51. Hockings, R.L., McAuley, J.H., and Maher, C.G., *A systematic review of the predictive ability of the Orebro Musculoskeletal Pain Questionnaire*. Spine 2008. **33**(15): p. E494-500.
52. Linton, S.J. and Boersma, K., *Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Orebro Musculoskeletal Pain Questionnaire*. Clin J Pain, 2003. **19**(2): p. 80-6.
53. Westman, A., Linton, S.J., Ohrvik, J., Wahlen, P., and Leppert, J., *Do psychosocial factors predict disability and health at a 3-year follow-up for patients with non-acute musculoskeletal pain? A validation of the Orebro Musculoskeletal Pain Screening Questionnaire*. Eur J Pain, 2008. **12**(5): p. 641-9.
54. Anagnostis C, G.R., Mayer TG. , *The pain disability questionnaire: A new psychometrically sound measure for chronic musculoskeletal disorders*. Spine 2004. **29**(20): p. 2290-2302.
55. Gatchel RJ, M.T., Theodore BR, *The pain disability questionnaire: relationship to one year functional and psychosocial rehabilitation outcomes*. J Occup Rehabil, 2006. **16**(1): p. 75-94.
56. Rondinelli, *Guides to the Evaluation of Permanent Impairment, 6<sup>th</sup> edition*, ed. R. RD2007, Chicago: American Medical Association. 632.
57. Chatman, A.B., Hyams, S.P., Neel, J.M., Binkley, J.M., Stratford, P.W., Schomberg, A., and Stabler, M., *The Patient-Specific Functional Scale: measurement properties in patients with knee dysfunction*. Phys Ther, 1997. **77**(8): p. 820-9.
58. Cleland, J.A., Fritz, J.M., Whitman, J.M., and Palmer, J.A., *The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy*. Spine 2006. **31**(5): p. 598-602.
59. Kowalchuk Horn, K., Jennings, S., Richardson, G., van Vliet, D., Hefford, C., and Abbott, J.H., *The patient-specific functional scale: psychometrics, clinimetrics, and application as a clinical outcome measure*. J Orthop Sports Phys Ther, 2012. **42**(1): p. 30-42.
60. George, S.Z., Fritz, J.M., and Childs, J.D., *Investigation of elevated fear-avoidance beliefs for patients with low back pain: a secondary analysis involving patients enrolled in physical therapy clinical trials*. J Orthop Sports Phys Ther, 2008. **38**(2): p. 50-58.
61. Turner, J.A., Franklin, G., Fulton-Kehoe, D., Sheppard, L., Stover, B., Wu, R., Gluck, J.V., and Wickizer, T.M., *ISSLS prize winner: early predictors of chronic work disability: a prospective, population-based study of workers with back injuries*. Spine 2008. **33**(25): p. 2809-18.
62. Wickizer, T.M., Franklin, G., Fulton-Kehoe, D., Gluck, J., Mootz, R., Smith-Weller, T., and Plaeger-Brockway, R., *Improving quality, preventing disability and reducing costs in workers' compensation healthcare: a population-based intervention study*. Med Care, 2011. **49**(12): p. 1105-11.



63. Spitzer, R.L., Kroenke, K., Williams, J.B., and Lowe, B., *A brief measure for assessing generalized anxiety disorder*. Arch Intern Med, 2006. **166**: p. 1092-1097.
64. Gilbody, S., Richards, D., Brealey, S., and Hewitt, C., *Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis*. J Gen Intern Med, 2007. **22**(11): p. 1596-602.
65. Kroenke, K., Spitzer, R.L., and Williams, J.B., *The PHQ-9: validity of a brief depression severity measure*. J Gen Intern Med, 2001. **16**(9): p. 606-13.
66. Wittkamp, K.A., Naeije, L., Schene, A.H., Huyser, J., and van Weert, H.C., *Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a systematic review*. Gen Hosp Psychiatry, 2007. **29**(5): p. 388-95.
67. Kroenke, K., Spitzer, R.L., and Williams, J.B., *The patient health questionnaire-2: validity of a two-item depression screener*. Med Care, 2003. **41**: p. 1284-1292.
68. Roelofs, J., Goubert, L., Peters, M., Vlaeyen, J.W., and Crombez, G., *The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia*. Eur J Pain, 2004. **8**(5): p. 495-502.
69. Swinkels-Meewisse, E.J., Swinkels, R.A., Verbeek, A.L., Vlaeyen, J.W., and Oostendorp, R.A., *Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain*. Man Ther, 2003. **8**(1): p. 29-36.
70. Swinkels-Meewisse, I.E., Roelofs, J., Verbeek, A.L., Oostendorp, R.A., and Vlaeyen, J.W., *Fear of movement/(re)injury, disability and participation in acute low back pain*. Pain, 2003. **105**(1-2): p. 371-9.
71. Ustun, T.B., Chatterji, S., Kostanjsek, N., Rehm, J., Kennedy, C., Epping-Jordan, J., Saxena, S., von Korff, M., and Pull, C., *Developing the World Health Organization Disability Assessment Schedule 2.0*. Bull World Health Organ, 2010. **88**(11): p. 815-23.
72. Feise RJ, M.J., *Functional rating index: a new valid and reliable instrument to measure the magnitude of clinical change in spinal conditions*. Spine 2001. **26**(1): p. 78-86.
73. Jacobson, G.P., Ramadan, N.M., Aggarwal, S.K., and Newman, C.W., *The Henry Ford Hospital Headache Disability Inventory (HDI)*. Neurology, 1994. **44**(5): p. 837-42.
74. MacDermid, J.C., Walton, D.M., Avery, S., Blanchard, A., Etruw, E., McAlpine, C., and Goldsmith, C.H., *Measurement properties of the neck disability index: a systematic review*. J Orthop Sports Phys Ther, 2009. **39**(5): p. 400-17.
75. Pinfold, M., Niere, K.R., O'Leary, E.F., Hoving, J.L., Green, S., and Buchbinder, R., *Validity and internal consistency of a whiplash-specific disability measure*. Spine 2004. **29**(3): p. 263-8.
76. Willis, C., Niere, K.R., Hoving, J.L., Green, S., O'Leary, E.F., and Buchbinder, R., *Reproducibility and responsiveness of the Whiplash Disability Questionnaire*. Pain, 2004. **110**(3): p. 681-8.
77. Fritz, J.M. and Irrgang, J.J., *A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale*. Phys Ther, 2001. **81**(2): p. 776-88.
78. Maughan, E.F. and Lewis, J.S., *Outcome measures in chronic low back pain*. Eur Spine J, 2010. **19**(9): p. 1484-94.
79. Stratford, P.W., Binkley, J., Solomon, P., Finch, E., Gill, C., and Moreland, J., *Defining the minimum level of detectable change for the Roland-Morris questionnaire*. Phys Ther, 1996. **76**(4): p. 359-65; discussion 366-8.
80. Angst, F., Goldhahn, J., Drerup, S., Flury, M., Schwyzer, H., and Simmen, B., *How sharp is the short QuickDASH? A refined content and validity analysis of the short form of the disabilities of the shoulder, arm, and hand questionnaire in the strata of symptoms and functions and specific joint conditions*. Qual Life Res 2009, 2009. **18**: p. 1043-1051.
81. Kocher, M., Horan, M., Briggs, K., Richardson, T., O'Holleran, J., and Hawkins, R., *Reliability, validity, and responsiveness of the American Shoulder and Elbow Surgeons subjective shoulder scale in patients with shoulder instability, rotator cuff disease, and glenohumeral arthritis*. J Bone Joint Surg Am, 2005. **87**(9): p. 2006-2011.
82. Stratford, P., Binkley, J., and Stratford, D., *Development and initial validation of the upper extremity functional index*. Physiotherapy Canada, 2001: p. 259-267.
83. Gable, P., Michener, L., Burkett, B., and Neller, A., *The upper limb functional index: development and determination of reliability, validity, and responsiveness*. J Hand Ther, 2006: p. 328-349.
84. Angst, F., John, M., Goldhahn, J., Herren, D.B., Pap, G., Aeschlimann, A., Schwyzer, H.K., and Simmen, B.R., *Comprehensive assessment of clinical outcome and quality of life after resection interposition arthroplasty of the thumb saddle joint*. Arthritis Rheum, 2005. **53**(2): p. 205-13.
85. MacDermid, J.C., *Outcome evaluation in patients with elbow pathology: issues in instrument development and evaluation*. J Hand Ther, 2001. **14**(2): p. 105-14.
86. Watson, C.J., Propps, M., Ratner, J., Zeigler, D.L., Horton, P., and Smith, S.S., *Reliability and responsiveness of the lower extremity functional scale and the anterior knee pain scale in patients with anterior knee pain*. J Orthop Sports Phys Ther, 2005. **35**(3): p. 136-46.
87. Johanson NA, L.M., Daltroy L, Rudicel S, Richmond J. May;(5):902-9., *American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change*. J Bone Joint Surg Am., 2004. **86-A**(5): p. 902-909.
88. Garratt, A.M., Brealey, S., and Gillespie, W.J., *Patient-assessed health instruments for the knee: a structured review*. Rheumatology (Oxford), 2004. **43**(11): p. 1414-23.
89. Wang, D., Jones, M.H., Khair, M.M., and Miniaci, A., *Patient-reported outcome measures for the knee*. J Knee Surg, 2010. **23**(3): p. 137-51.
90. Martin, R.L. and Irrgang, J.J., *A survey of self-reported outcome instruments for the foot and ankle*. J Orthop Sports Phys Ther, 2007. **37**(2): p. 72-84.
91. Hale, S.A. and Hertel, J., *Reliability and Sensitivity of the Foot and Ankle Disability Index in Subjects With Chronic Ankle Instability*. J Athl Train, 2005. **40**(1): p. 35-40.
92. Budiman-Mak, E., Conrad, K.J., and Roach, K.E., *The Foot Function Index: a measure of foot pain and disability*. J Clin Epidemiol, 1991. **44**(6): p. 561-70.
93. Leite, J.C., Jerosch-Herold, C., and Song, F., *A systematic review of the psychometric properties of the Boston Carpal Tunnel Questionnaire*. BMC Musculoskelet Disord, 2006. **7**: p. 78.
94. Levine, D.W., Simmons, B.P., Koris, M.J., Daltroy, L.H., Hohl, G.G., Fossel, A.H., and Katz, J.N., *A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome*. J Bone Joint Surg Am, 1993. **75**(11): p. 1585-92.
95. Briggs, K.K., Lysholm, J., Tegner, Y., Rodkey, W.G., Kocher, M.S., and Steadman, J.R., *The reliability, validity, and responsiveness of the Lysholm score and Tegner activity scale for anterior cruciate ligament injuries of the knee: 25 years later*. Am J Sports Med, 2009. **37**(5): p. 890-7.

96. Kocher, M.S., Steadman, J.R., Briggs, K.K., Sterett, W.I., and Hawkins, R.J., *Reliability, validity, and responsiveness of the Lysholm knee scale for various chondral disorders of the knee*. J Bone Joint Surg Am, 2004. **86-A**(6): p. 1139-45.
97. Marx, R.G., Jones, E.C., Allen, A.A., Altchek, D.W., O'Brien, S.J., Rodeo, S.A., Williams, R.J., Warren, R.F., and Wickiewicz, T.L., *Reliability, validity, and responsiveness of four knee outcome scales for athletic patients*. J Bone Joint Surg Am, 2001. **83-A**(10): p. 1459-69.
98. Tegner, Y. and Lysholm, J., *Rating systems in the evaluation of knee ligament injuries*. Clin Orthop Relat Res, 1985(198): p. 43-9.
99. Roos, E.M. and Lohmander, L.S., *The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis*. Health Qual Life Outcomes, 2003. **1**: p. 64.
100. Roos, E.M., Roos, H.P., Lohmander, L.S., Ekdahl, C., and Beynnon, B.D., *Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure*. J Orthop Sports Phys Ther, 1998. **28**(2): p. 88-96.
101. Roos, E.M. and Toksvig-Larsen, S., *Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement*. Health Qual Life Outcomes, 2003. **1**: p. 17.
102. Crossley, K.M., Bennell, K.L., Cowan, S.M., and Green, S., *Analysis of outcome measures for persons with patellofemoral pain: which are reliable and valid?* Arch Phys Med Rehabil, 2004. **85**(5): p. 815-22.
103. Robinson, J.M., Cook, J.L., Purdam, C., Visentini, P.J., Ross, J., Maffulli, N., Taunton, J.E., and Khan, K.M., *The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy*. Br J Sports Med, 2001. **35**(5): p. 335-41.
104. Iversen, J.V., Bartels, E.M., and Langberg, H., *The Victorian Institute of Sports Assessment - achilles questionnaire (visa-a) - a reliable tool for measuring achilles tendinopathy*. Int J Sports Phys Ther, 2012. **7**(1): p. 76-84.
105. Bellamy, N., *Pain assessment in osteoarthritis: experience with the WOMAC osteoarthritis index*. Semin Arthritis Rheum, 1989. **18**(4 Suppl 2): p. 14-7.
106. Bellamy, N., Buchanan, W.W., Goldsmith, C.H., Campbell, J., and Stitt, L.W., *Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee*. J Rheumatol, 1988. **15**(12): p. 1833-40.
107. Bellamy, N., Kean, W.F., Buchanan, W.W., Gercz-Simon, E., and Campbell, J., *Double blind randomized controlled trial of sodium meclofenamate (Meclomen) and diclofenac sodium (Voltaren): post validation reapplication of the WOMAC Osteoarthritis Index*. J Rheumatol, 1992. **19**(1): p. 153-9.
108. Hawker, G., Melfi, C., Paul, J., Green, R., and Bombardier, C., *Comparison of a generic (SF-36) and a disease specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery*. J Rheumatol, 1995. **22**(6): p. 1193-6.
109. Lequesne, M., *Indices of severity and disease activity for osteoarthritis*. Semin Arthritis Rheum, 1991. **20**(6 Suppl 2): p. 48-54.
110. McConnell, S., Kolopack, P., and Davis, A.M., *The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties*. Arthritis Rheum, 2001. **45**(5): p. 453-61.
111. Stucki, G., Sangha, O., Stucki, S., Michel, B.A., Tyndall, A., Dick, W., and Theiler, R., *Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis*. Osteoarthritis Cartilage, 1998. **6**(2): p. 79-86.
112. Whitehouse, S.L., Crawford, R.W., and Learmonth, I.D., *Validation for the reduced Western Ontario and McMaster Universities Osteoarthritis Index function scale*. J Orthop Surg (Hong Kong), 2008. **16**(1): p. 50-3.
113. Fulton-Kehoe, D., Stover, B.D., Turner, J.A., Sheppard, L., Gluck, J.V., Wickizer, T.M., and Franklin, G.M., *Development of a brief questionnaire to predict long-term disability*. J Occup Environ Med, 2008. **50**(9): p. 1042-52.
114. Von Korff, M., *Chronic pain assessment in epidemiologic and health services research: empirical bases and new directions.*, in *Handbook of Pain Assessment*, D.C. Turk and R. Melzack, Editors. In press, Guilford Press: New York.
115. Von Korff, M., Shortreed, S.M., Saunders, K.W., LeResche, L., Berlin, J.A., Stang, P., and Turner, J.A., *Comparison of back pain prognostic risk stratification item sets*. J Pain, 2014. **15**(1): p. 81-9.
116. Carnes, D., Parsons, S., Ashby, D., Breen, A.C., Foster, N.E., Pincus, T., Vogel, S., and Underwood, M., *Chronic musculoskeletal pain rarely presents in a single body site: results from a UK population study*. Rheumatology, 2007. **46**(7): p. 1168-1170.
117. Lipchik, G.L., Milles, K., and Covington, E.C., *The effects of multidisciplinary pain management treatment on locus of control and pain beliefs in chronic non-terminal pain*. Clin J Pain, 1993. **9**(1): p. 49-57.
118. Alaranta, H., Luoto, S., Heliovaara, M., and Hurri, H., *Static back endurance and the risk of low-back pain*. Clin Biomech (Bristol, Avon), 1995. **10**(6): p. 323-324.
119. Adedoyin, R.A., Mbada, C.E., Farotimi, A.O., Johnson, O.E., and Emechete, A.A., *Endurance of low back musculature: normative data for adults*. J Back Musculoskelet Rehabil, 2011. **24**(2): p. 101-9.
120. Arab, A.M., Salavati, M., Ebrahimi, I., and Ebrahim Mousavi, M., *Sensitivity, specificity and predictive value of the clinical trunk muscle endurance tests in low back pain*. Clin Rehabil, 2007. **21**(7): p. 640-7.
121. Moreau, C.E., Green, B.N., Johnson, C.D., and Moreau, S.R., *Isometric back extension endurance tests: a review of the literature*. J Manipulative Physiol Ther, 2001. **24**(2): p. 110-22.
122. Hamberg-van Reenen, H.H., Ariens, G.A., Blatter, B.M., Twisk, J.W., van Mechelen, W., and Bongers, P.M., *Physical capacity in relation to low back, neck, or shoulder pain in a working population*. Occup Environ Med, 2006. **63**(6): p. 371-7.

123. Hamberg-van Reenen, H.H., Ariens, G.A., Blatter, B.M., van der Beek, A.J., Twisk, J.W., van Mechelen, W., and Bongers, P.M., *Is an imbalance between physical capacity and exposure to work-related physical factors associated with low-back, neck or shoulder pain?* Scand J Work Environ Health, 2006. **32**(3): p. 190-7.
124. Johnson, O.E., Mbada, C.E., Akosile, C.O., and Agbeja, O.A., *Isometric endurance of the back extensors in school-aged adolescents with and without low back pain.* J Back Musculoskeletal Rehabil, 2009. **22**(4): p. 205-11.
125. Taechasubamorn, P., Nopkesorn, T., and Pannarunothai, S., *Comparison of physical fitness between rice farmers with and without chronic low back pain: a cross-sectional study.* J Med Assoc Thai, 2010. **93**(12): p. 1415-21.
126. Bohannon, R.W., Larkin, P.A., Cook, A.C., Gear, J., and Singer, J., *Decrease in timed balance test scores with aging.* Phys Ther, 1984. **64**(7): p. 1067-70.
127. McGill, S.M., Childs, A., and Liebenson, C., *Endurance times for low back stabilization exercises: clinical targets for testing and training from a normal database.* Arch Phys Med Rehabil, 1999. **80**(8): p. 941-4.
128. Alaranta, H., Hurri, H., Heliovaara, M., Soukka, A., and Harju, R., *Non-dynamometric trunk performance tests: reliability and normative data.* Scand J Rehabil Med, 1994. **26**(4): p. 211-5.
129. Ekstrand, J., Wiktorsson, M., Oberg, B., and Gillquist, J., *Lower extremity goniometric measurements: a study to determine their reliability.* Arch Phys Med Rehabil, 1982. **63**(4): p. 171-5.
130. Harris, K.D., Heer, D.M., Roy, T.C., Santos, D.M., Whitman, J.M., and Wainner, R.S., *Reliability of a measurement of neck flexor muscle endurance.* Phys Ther, 2005. **85**(12): p. 1349-55.
131. Olson, L.E., Millar, A.L., Dunker, J., Hicks, J., and Glanz, D., *Reliability of a clinical test for deep cervical flexor endurance.* J Manipulative Physiol Ther, 2006. **29**(2): p. 134-8.
132. Yeomans, S., *The clinical application of outcomes assessment.* 2000, Norwalk, CT: Appelton Lange.
133. Mahmud, N., Schonstein, E., Schaafsma, F., Lehtola, M.M., Fassier, J.B., Verbeek, J.H., and Reneman, M.F., *Functional capacity evaluations for the prevention of occupational re-injuries in injured workers.* Cochrane Database Syst Rev, 2010(7): p. CD007290.
134. American Academy of Neurology. *Clinical Practice Guideline Process Manual* 2011.