Title
Development and Initial Validation of the PEG, a Three-item Scale Assessing Pain Intensity and Interference

Authors
Krebs, Erin E.
Lorenz, Karl A.
Bair, Matthew J.
et al.

Peer reviewed

BRIEF REPORT

Development and Initial Validation of the PEG, a Three-item Scale Assessing Pain Intensity and Interference

Erin E. Krebs, MD, MPH[1,2,3], Karl A. Lorenz, MD, MSHS[5,6,7], Matthew J. Bair, MD, MS[1,2,3], Teresa M. Damush, PhD[1,2,3], Jingwei Wu, MS[3], Jason M. Sutherland, PhD[4], Steven M. Asch, MD, MPH[5,6,7], and Kurt Kroenke, MD[1,2,3]

[1]Center on Implementing Evidence-Based Practice, Roudebush VA Medical Center, Indianapolis, IN, USA; [2]Regenstrief Institute, Inc., Indianapolis, IN, USA; [3]Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA; [4]The Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth Medical School, Hanover, NH, USA; [5]VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA; [6]RAND Corporation, Santa Monica, CA, USA; [7]Geffen School of Medicine at UCLA, Los Angeles, CA, USA.

BACKGROUND: Inadequate pain assessment is a barrier to appropriate pain management, but single-item "pain screening" provides limited information about chronic pain. Multidimensional pain measures such as the Brief Pain Inventory (BPI) are widely used in pain specialty and research settings, but are impractical for primary care. A brief and straightforward multidimensional pain measure could potentially improve initial assessment and follow-up of chronic pain in primary care.

OBJECTIVES: To develop an ultra-brief pain measure derived from the BPI.

DESIGN: Development of a shortened three-item pain measure and initial assessment of its reliability, validity, and responsiveness.

PARTICIPANTS: We used data from 1) a longitudinal study of 500 primary care patients with chronic pain and 2) a cross-sectional study of 646 veterans recruited from ambulatory care.

RESULTS: Selected items assess average pain intensity (P), interference with enjoyment of life (E), and interference with general activity (G). Reliability of the three-item scale (PEG) was $\alpha=0.73$ and 0.89 in the two study samples. Overall, construct validity of the PEG was good for various pain-specific measures (r=0.60–0.89 in Study 1 and r=0.77–0.95 in Study 2), and comparable to that of the BPI. The PEG was sensitive to change and differentiated between patients with and without pain improvement at 6 months.

DISCUSSION: We provide strong initial evidence for reliability, construct validity, and responsiveness of the PEG among primary care and other ambulatory clinic patients. The PEG may be a practical and useful tool to improve assessment and monitoring of chronic pain in primary care.

KEY WORDS: pain; measurement; primary care.

BACKGROUND

Inadequate pain assessment has been identified as a key barrier to appropriate pain management.[1–3] Recently, important initiatives have aimed to increase awareness of pain as a clinical problem by promoting better pain assessment.[4–6] These initiatives have led to widespread adoption of pain screening through measurement of current pain intensity.

In chronic pain, the most common type of pain seen in primary care, assessment of pain intensity alone is inadequate. Guidelines encourage comprehensive assessment that includes measurement of pain-related functioning, which may be even more relevant to patients' overall quality of life than intensity.[7,8] To facilitate chronic pain assessment, numerous multidimensional patient-reported measures have been developed;[9–12] however, none of these have been widely adopted in the general medical settings where most chronic pain treatment is delivered.

In primary care, use of multidimensional pain measures is limited by factors such as instrument length and scoring complexity; however, a brief and straightforward multidimensional measure could potentially improve assessment of chronic pain. We sought to develop a very brief measure that would be feasible, valid, and sensitive to change in primary care. We started with the Brief Pain Inventory (BPI) because it is relatively easy to administer, score, and interpret; includes items assessing pain intensity and functional interference; and has been validated in many pain conditions.[10,13–17] As its name implies, the BPI is shorter than other multidimensional pain measures, but it is still too lengthy for implementation in primary care practice. We hypothesized that a shortened scale based on the BPI could be developed that would be more feasible, but just as useful, for assessing chronic pain in primary care. Our objectives were to develop an ultra-brief scale derived from the BPI and to initially assess its reliability, validity, and responsiveness.

Items	Alpha	Item-total correlation	Alpha (deleted variable)*
Preferred scale (PEG)	**0.73**		
Average intensity		0.49	0.74
General activity		0.65	0.52
Enjoyment of life		0.59	0.62
Alternate scale 1	**0.70**		
Average intensity		0.48	0.68
General activity		0.60	0.49
Mood		0.53	0.62
Alternate scale 2	**0.66**		
Average intensity		0.50	0.59
General activity		0.53	0.49
Sleep		0.48	0.62
Alternate scale 3	**0.77**		
Average intensity		0.50	0.76
General activity		0.63	0.68
Sleep		0.56	0.73
Enjoyment of life		0.65	0.66

*Cronbach coefficient alpha for the scale with that item deleted.

METHODS

Participants

We used data from two sources: 1) Stepped Care for Affective Disorders and Musculoskeletal Pain (SCAMP), a longitudinal study that enrolled a total of 500 patients with chronic musculoskeletal pain, and 2) Helping Veterans Experience Less Pain (HELP-vets), a cross-sectional study of 646 veterans receiving care at VA clinics. We used data from Study 1 to develop and initially validate the ultra-brief measure and data from Study 2 to confirm reliability and validity in an independent patient population.

Study 1 (SCAMP) enrolled 500 primary care patients with persistent back, hip, or knee pain of at least moderate severity, 250 of whom had concurrent depression.[18] Participants were recruited from university (n=300) and VA-affiliated (n=200) internal medicine clinics in Indianapolis. Patients with concurrent depression were enrolled in a trial of depression and pain treatment vs. usual care (n=250). Those without depression were followed in a parallel observational study (n=250). The mean age of SCAMP participants was 59 years; 52% were women, 58% were white, and 38% were black. The mean numeric rating of current pain (on a 0–10 scale) was 6.1 (SD 1.9) at baseline.

Study 2 (HELP-vets) enrolled a random visit-based sample of 646 veterans from ambulatory care clinics at two VA hospitals and six affiliated community sites in three urban California counties. Patients with chronic illness were over-sampled by design. The mean age was 63 years and 95% were male. Self-reported race/ethnicity was 54% white, 30% black, and 10% Latino. Sixty-one percent of participants reported pain at the time of enrollment and 63% had one or more pain diagnoses (33% back pain, 45% other musculoskeletal pain, 12% neuropathic pain, 5% headache). The mean rating of current pain (on a 0–10 scale) was 3.1 (SD 3.2) overall and 5.1 (SD 2.6) among those with pain.

Measures

Study 1 participants completed the BPI, Chronic Pain Grade questionnaire (CPG), Roland disability scale, and SF-36 bodily pain scale at baseline; they completed the BPI, CPG, and pain global rating of change at 6 months. Study 2 was cross-

sectional; participants completed the BPI, Functional Morbidity Index, and a single-item rating of overall pain-related distress.

- The *Brief Pain Inventory (BPI)* includes two scales that assess pain intensity and pain-related functional impairment (physical and emotional).[13,15] The four items of the BPI severity scale assess the intensity of current pain and pain at its least, worst, and average during the past week on scales from 0 ("no pain") to 10 ("pain as bad as you can imagine"). The BPI interference scale assesses pain-related functional interference with seven items assessing different domains (general activity, mood, walking ability, normal work, relations with other people, sleep, and enjoyment of life) rated from 0 ("does not interfere") to 10 ("interferes completely").

- The *Chronic Pain Grade questionnaire (CPG)* includes two three-item scales (intensity and disability) that are transformed into 0–100 scores.[19] An algorithm classifies pain into four graded categories: 1) low disability-low intensity, 2) low disability-high intensity, 3) high disability-moderately limiting, and 4) high disability-severely limiting. The CPG has been validated in primary care, chronic pain, and general populations.[20–22]

- The *Roland Disability questionnaire* is a pain-specific measure of physical disability validated in patients with back pain and other chronic pain conditions.[23,24] It includes a checklist of 24 statements about pain effects on function; the score is the number of items endorsed.

- The *Short-Form 36-item questionnaire (SF-36) Bodily Pain Scale* is a two item scale assessing pain severity and interference.[25,26] Responses are transformed into a 0–100 score.

- The *Pain Global Rating of Change* is a single item assessing patients' overall impression of change in their pain. Study 1 participants were asked whether their pain was worse, about the same, or better since the start of the study. Those who reported that pain was better were asked to rate the magnitude of improvement (a little, somewhat, moderately, a lot, or completely better). Global ratings of change may be more sensitive to improvement and better correlated with patient satisfaction than serial measures.[27]

- The *Functional Morbidity Index* was developed to assess general functional status in older adults.[28] Patients indicate whether they are able perform four different activities independently, and if not, whether the impairment is due to a health problem.

1. What number best describes your pain on average in the past week:

0   1   2   3   4   5   6   7   8   9   10

No pain   Pain as bad as you can imagine

2. What number best describes how, during the past week, pain has interfered with your enjoyment of life?

0   1   2   3   4   5   6   7   8   9   10

Does not interfere   Completely interferes

3. What number best describes how, during the past week, pain has interfered with your general activity?

0   1   2   3   4   5   6   7   8   9   10

Does not interfere   Completely interferes

Figure 1. The PEG three-item scale. *Items from the Brief Pain Inventory reproduced with permission from Dr. Charles Cleeland.

Table 2. PEG and Individual Item Statistics at Baseline in Study 1 and Study 2

	Study 1 (n=500)	Study 2 (n=638)
	Mean (SD)	Mean (SD)
PEG scale	6.1 (2.2)	4.1 (3.1)
Item 1: average pain intensity	6.1 (1.9)	3.9 (2.8)
Item 2: interference with general activity	6.4 (2.7)	4.4 (3.6)
Item 3: interference with enjoyment of life	5.9 (3.3)	4.0 (3.9)

Study 1 included primary care patients with chronic musculoskeletal pain; Study 2 included ambulatory VA patients, 61% of whom had current pain. For each item, the response range is 0–10; the PEG scale score is the mean of the three individual item scores. Higher scores represent worse pain.

- *Overall Pain Distress* is a single item: "How much did overall pain distress or bother you during the past week?" Response options are not at all, a little bit, somewhat, quite a bit, and very much.

Item Selection

We used a consensus-based process, drawing on a literature review, expert opinion, and statistical data, to develop a shortened scale.[29] Pre-specified criteria guided initial item selection. First, we decided to include at least one item representing each of three domains included in the BPI: pain intensity, physical functioning, and emotional functioning. We then selected items with the following characteristics: 1) easy to understand and applicable to patients with all types of pain; 2) good statistical characteristics (e.g., high response variability, high item-remainder correlation); 3) similar performance in depressed and non-depressed patients.

We chose "pain average" for the intensity item because it had a good distribution of responses, lacking the ceiling and floor effects seen with "pain worst" and "pain least," respectively. We did not select "pain now" because we wanted to avoid duplicating information provided by the "fifth vital sign," and capture intermittent pain. Although the ideal reporting period for pain assessment is debated, recalled average pain over one week is a valid measure of pain intensity.[30,31]

BPI interference items include those assessing physical status (general activity, walking, normal work), emotional status (mood, relations with others, enjoyment of life), and sleep. For physical interference, we chose "interference with

general activity" because it applies equally to all patients, as opposed to "interference with work" (which may be affected by occupation, employment status, etc.) and "interference with walking" (which may not apply to non-ambulatory patients or those with upper body pain).

For emotional interference, we considered both "interference with mood" and "interference with enjoyment of life." In our experience, "interference with relations with other people" is more difficult than other items for patients to answer. We wanted a scale that would discriminate between chronic pain and depression, which commonly co-occur.[32] In our sample of patients with and without comorbid depression, we found that "interference with enjoyment of life" was more independent of depression than "interference with mood." We also considered "interference with sleep" in place of the emotional interference items.

We reached consensus on a preferred three-item scale ("pain average," "interference with enjoyment of life," and "interference with general activity") and alternative three-item and four-item scales, which we then evaluated statistically.

Reliability and Validity

We assessed reliability (internal consistency) by calculating Cronbach's coefficient alpha. To assess construct validity, we compared the PEG with measures of pain and function using Pearson correlation coefficients. We used multiple measures for construct validity assessment, including the BPI, because no criterion standard exists for pain. We hypothesized that coefficients would be slightly higher for comparisons with pain-specific functional measures than for those with pain severity measures (because two of the three PEG items assess function). We also expected that coefficients would be larger for comparisons with pain-specific functional measures than for comparisons with generic functional measures.

Responsiveness

Assessment of responsiveness, or sensitivity to change, requires an independent standard to define change.[33] We used two different measurements to define the presence or absence of patient improvement: 1) global rating of change and 2) serial CPG grade. We categorized patients according to their pain trajectory as assessed by each of the two measures. Global rating of change categories were defined by the patient's retrospective assessment at 6 months of the change in their

Table 3. Correlation between the PEG, BPI Scales, and other Measures at Baseline

Measure						
Correlation coefficient						
Study 1 (n=500)						
	BPI severity	BPI interference	CPG intensity	CPG disability	Roland disability	SF-36 bodily pain*
PEG	0.69	0.89	0.64	0.67	0.60	−0.61
BPI severity	–	0.58	0.82	0.47	0.41	−0.46
BPI interference	–	–	0.62	0.71	0.70	−0.65
Study 2 (n=638)						
	BPI severity	BPI interference	Functional morbidity	Overall pain distress		
PEG	0.84	0.95	0.54	0.77		
BPI severity	–	0.75	0.47	0.78		
BPI interference	–	–	0.55	0.72		

* For SF-36, low score represents worse pain. For all other measures, a high score is worse.

Table 4. Responsiveness of PEG among Patients Classified by Pain Global Rating of Change and Serial CPG Grade at 6 Months (n=210)

	PEG baseline	PEG 6 months	PEG change*	ES†	SRM‡
Global rating of change					
Improved (n=66)	6.35 (2.32)	3.37 (2.33)	2.99 (2.49)	1.29	1.20 (0.96, 1.44)
Unchanged (n=83)	6.95 (1.75)	6.49 (1.91)	0.46 (1.60)	0.26	0.29 (0.07, 0.51)
Worse (n=61)	6.99 (1.73)	7.10 (1.75)	−0.10 (1.69)	−0.06	−0.06 (-0.31, 0.19)
CPG grade					
Improved (n=62)	6.83 (1.73)	4.20 (2.37)	2.62 (2.66)	1.51	0.99 (0.74, 1.24)
Unchanged (n=115)	6.92 (2.11)	6.40 (2.35)	0.53 (1.80)	0.25	0.29 (0.11, 0.47)
Worse (n=32)	6.16 (1.66)	6.07 (2.43)	0.08 (1.92)	0.05	0.04 (-0.31, 0.39)

*Change score = baseline score – 6 month score
†ES = change score/SD of baseline score
‡SRM = change score/SD of change score

pain since the trial began: 1) improved ("better"), 2) unchanged ("about the same"), and 3) worse ("worse"). CPG categories were defined by the change in CPG grade from baseline to 6 months: 1) improved (pain grade decreased by ≥1 level), 2) unchanged (pain grade at baseline = pain grade at follow-up), and 3) worse (pain grade increased by ≥1 level).

Using data from Study 1, we assessed responsiveness by calculating the following three metrics: 1) change score (difference between mean score at baseline and follow-up), 2) effect size (ES; change score divided by the standard deviation of the baseline score), and 3) standardized response mean (SRM; change score divided by the standard deviation of the change score). These calculations were performed for patients in the improved, unchanged, and worse categories. Confidence intervals for SRM were calculated as + /- 1.96 divided by the square root of the sample size.[34] We assessed responsiveness using all three methods because they can produce differing results and because agreement is lacking on the preferred method.[34,35] We compared responsiveness of the PEG, BPI severity and BPI interference scales by comparing ES and SRM for each measure among patients in the improved category. Finally, we assessed responsiveness to varying degrees of improvement by comparing change scores for PEG and BPI scales to degree of improvement by global rating of change.

RESULTS

Item Selection

Using baseline data from the Study 1 sample, we assessed the preferred and alternate scales. Results were similar for all scales evaluated (Table 1). Our preferred scale demonstrated initial characteristics similar to or better than the alternatives, so we chose it as our final scale. From here on, we refer to this final abbreviated scale as the PEG, an acronym representing the three items: "Pain average," "interference with Enjoyment of life," and "interference with General activity." Principal components analysis of the PEG in both samples demonstrated a single factor, accounting for 66% of the variance in Study 1 and 81% in Study 2.

The PEG comprises 1 intensity item and 2 interference items (Fig. 1). Consistent with BPI scoring, we calculated the average of individual item scores to get an overall PEG score (potential range 0–10). Table 2 shows means and standard deviations (SD) for each item and the full three-item scale in both populations.

Reliability and Validity

Reliability of the PEG was 0.73 in Study 1 and 0.89 in Study 2. Table 3 shows correlation matrices for the PEG, BPI scales, and other pain and function measures in both study populations. Overall, construct validity of the PEG was good (r=0.60–0.89 in Study 1 and r=0.77–0.95 for pain-specific measures in Study 2), with correlations comparable to those of the BPI scales. As expected, PEG correlations were slightly higher for BPI interference (r=0.89 and 0.95) than for BPI severity (r=0.69 and 0.84) and for CPG disability (r=0.67) than for CPG intensity (r=0.64). Correlations were higher for the pain-specific function measures than for the Functional Morbidity Index (r=0.54), a generic measure of function.

Responsiveness

Six-month follow-up data was available for 210 Study 1 clinical trial participants. The proportion with pain improvement was approximately the same according to global rating of change (31.4%) and serial CPG pain grade (29.5%); however, more patients were classified as worse by global rating (29.0%) than by CPG grade (15.0%). Table 4 shows PEG scores and measures of responsiveness for patients classified as improved, unchanged, and worse at 6 months. Confidence intervals for the improved and unchanged groups did not overlap, but the unchanged and worse groups were not statistically different from each other. The improved group according to global rating of change had a mean improvement of 3.0 points (SD 2.5) on the PEG. Similarly, the
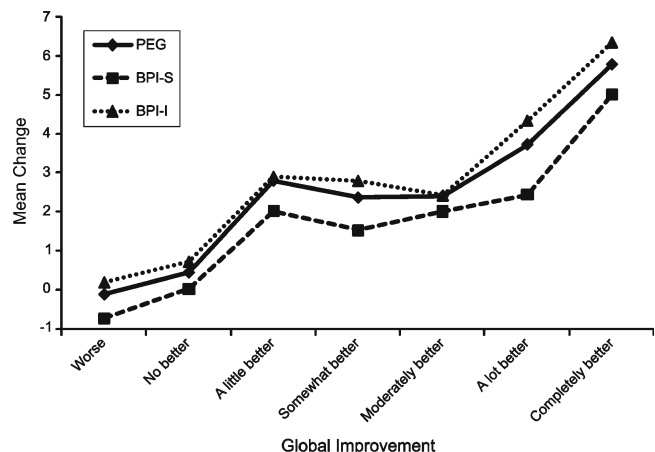

Figure 2. Mean change in PEG and BPI scales compared with global rating of change at 6 months (n=210).

improved group according to serial CPG grade had a mean change of 2.6 points (SD 2.7).

The SRM among participants who improved at 6 months according to global rating of change were similar for the PEG (1.20, 95% CI 0.96, 1.44), BPI severity (1.04, 95% CI 0.80, 1.28), and BPI interference (1.13, 95% CI 0.89, 1.37). Results were similar for improvement according to serial CPG grade (data not shown). For all measures of improvement, ES and SRM were consistent with a large effect. Figure 2 shows 6-month change scores for the PEG, BPI severity scale, and BPI interference scale plotted against global rating of change.

DISCUSSION

We demonstrated that the PEG, an ultra-brief three-item scale derived from the BPI, was a reliable and valid measure of pain among primary care patients with chronic musculoskeletal pain and diverse VA ambulatory patients. The PEG appears comparable to the BPI in terms of responsiveness to change. These findings support our hypothesis that an abbreviated scale derived from the BPI may be both useful and practical for chronic pain assessment in primary care and other ambulatory care settings, such as medical and surgical specialty clinics.

Strengths of this study include the confirmation of reliability and validity in an independent patient population, the diversity of the study populations, and the availability of multiple pain and functional measures with which to assess construct validity. Our choice of the BPI as the basis for our abbreviated scale development is another strength. The BPI is a widely used instrument that has been validated in numerous patient populations, clinical settings, and languages. BPI items are rated from 0–10, a format that has become familiar to patients and clinicians since assessment of pain with numeric scales has been broadly implemented in US health care settings. We took advantage of our collective experience with the BPI in observational and interventional research by employing a consensus-based process for scale shortening, consistent with recommendations to avoid over-reliance on statistical techniques.[29]

We believe our use of two different ambulatory study populations is a strength; however, each study has its own limitations. Study 1 was a sample of patients with chronic back and lower extremity musculoskeletal pain and included an over-representation of patients with depression (50% by design). The more clinically diverse patient population of Study 2, including ambulatory VA patients with and without chronic pain, enhances the generalizability of our findings. However, Study 2 included fewer pain measures with which to assess construct validity and was cross-sectional; therefore, we were able to assess responsiveness only in the first sample. Forty percent of Study 1 patients and 100% of Study 2 patients were recruited from VA clinics, so our findings may be less generalizable to non-VA settings.

We found that the PEG differentiated well between patients who improved and those who did not. According to responsiveness metrics, patients in the improved category had a large improvement in PEG score, whereas those in the unchanged category had a minimal change. Proper interpretation of the magnitude of change according to SRM and one group pre-post ES is not entirely clear, although authors have suggested that Cohen's definition of small (0.2), moderate (0.5), and large (0.8) effects can be applied to interpretation of both responsiveness

measures.[36,37] We did not find evidence of PEG responsiveness in the worse direction (i.e., change scores between those who were unchanged and those who were worse did not significantly differ). We are limited in our ability to adequately assess sensitivity to worsening because we evaluated responsiveness in a single study population that likely had a ceiling effect for worsening due to high baseline pain severity.

The competing demands of primary care, in which visits are short and pain is only one of several problems warranting attention, make efficiency of assessment a paramount concern.[38,39] A balance must be found between feasibility and key characteristics such as reliability, validity, and responsiveness. For example, ultra-brief depression measures containing two to three items perform better than single item depression measures.[40] We also evaluated a four-item abbreviated scale, but found that adding an item contributed little. An abbreviated version of the BPI that eliminated a few items has been previously published,[41] but the PEG is the first ultra-brief scale based on the BPI.

New assessment strategies are needed to support improved chronic pain management in primary care. We believe the PEG, which includes items assessing pain intensity, emotional function, and physical function, is an important step forward. However, further studies are needed to confirm our findings and validate the PEG in additional patient populations. Prospective research should determine whether serial pain measurement can improve the quality of clinical decision-making and pain outcomes in primary care. Given the huge clinical and societal burden associated with pain, developing efficient and effective strategies to enhance care is an important priority.

Corresponding Author: *Erin E. Krebs, MD, MPH, Roudebush VAMC, 11H, 1481 W. 10th Street, Indianapolis, IN 46202, USA (e-mail: krebse@iupui.edu).*

REFERENCES

1. Improving the Quality of Pain Management Through Measurement and Action. 2003. Oakbrook Terrace, IL, Joint Commission on Accreditation of Healthcare Organizations and National Pharmaceutical Council.
2. **Carr DB, Jacox AK, Chapman CR, Fillingim RB, Heidrich G, Hester NK.** Acute pain management: Operative or Medical Procedures and Trauma. Clinical Practice Guideline (AHCPR Pub. No. 92–0032). 1992. Rockville, MD, Agency for Health Care Policy and Research, Public Health Service, US Department of Health and Human Services.
3. **Jacox A, Carr DB, Payne R.** Management of Cancer Pain. Clinical Practice Guideline (AHCPR Publication No. 94–0592). 1994. Rockville,

MD, Agency for Health Care Policy and Research, U.S. Department of Health and Human Services, Public Health Service. 1994.

4. Pain as the 5th vital sign toolkit, revised edition. National Pain Management Coordinating Committee, Veterans Health Administration 2000; Available at: http://www1.va.gov/pain_management/docs/TOOLKIT.pdf. Accessed March 27, 2009.

5. Comprehensive Accreditation Manual for Ambulatory Care. Oakbrook Terrace, IL: Joint Commission on Accreditation of Healthcare Organizations, 2006.

6. **Dahl JL, Gordon DB.** Joint Commission pain standards: a progress report. APS Bulletin 2002;12.

7. VHA pain outcomes toolkit. National Pain Management Coordinating Committee, Veterans Health Administration 2003; Available at: http://www1.va.gov/Pain_Management/page.cfm?pg=41. Accessed March 27, 2009.

8. **Simon L, Lipman A, Jacox A, et al.** Guideline for the Management of Arthritis Pain in Osteoarthritis, Rheumatoid Arthritis, and Juvenile Chronic Arthritis. Glenville, IL: American Pain Society; 2002.

9. **Turk DC, Melzack R.** The measurement of pain and the assessment of people experiencing pain. In: Turk DC, Melzack R, eds. Handbook of Pain Assessment. 2nd ed. New York, NY: Guilford Press; 2001.

10. **Dworkin RH, Turk DC, Farrar JT, et al.** Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005;113:9–19.

11. **Turk DC, Dworkin RH, Allen RR, et al.** Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain. 2003; 106: 337–45.

12. **Dworkin RH, Turk DC, Wyrwich KW, et al.** Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain. 2008;9:105–21.

13. **Cleeland CS.** Pain assessment in cancer. In: Osoba D, ed. Effect of Cancer on Quality of Life. Boca Raton, FL: CRC Press; 1991:293–305.

14. **Cleeland CS, Ryan KM.** Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singapore. 1994;23:129–38.

15. **Cleeland CS, Nakamura Y, Mendoza TR, Edwards KR, Douglas J, Serlin RC.** Dimensions of the impact of cancer pain in a four country sample: new information from multidimensional scaling. Pain. 1996;67:267–73.

16. **Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS.** Validity of the Brief Pain Inventory for use in documenting the outcomes of patients with noncancer pain. Clin J Pain. 2004;20:309–18.

17. **Tan G, Jensen MP, Thornby JI, Shanti BF.** Validation of the Brief Pain Inventory for chronic nonmalignant pain. J Pain. 2004;5:133–7.

18. **Kroenke K, Bair M, Damush T, et al.** Stepped Care for Affective Disorders and Musculoskeletal Pain (SCAMP) study: design and practical implications of an intervention for comorbid pain and depression. Gen Hosp Psychiatry. 2007;29:506–17.

19. **Von Korff M, Ormel J, Keefe FJ, Dworkin SF.** Grading the severity of chronic pain. Pain. 1992;50:133–49.

20. **Elliott AM, Smith BH, Smith WC, Chambers WA.** Changes in chronic pain severity over time: the Chronic Pain Grade as a valid measure. Pain. 2000;88:303–8.

21. **Smith BH, Penny KI, Purves AM, et al.** The Chronic Pain Grade questionnaire: validation and reliability in postal research. Pain. 1997;71:141–7.

22. **Von Korff M, Deyo RA, Cherkin D, Barlow W.** Back pain in primary care. Outcomes at 1 year. Spine. 1993;18:855–62.

23. **Roland M, Morris R.** A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine. 1983;8:141–4.

24. **Jensen MP, Strom SE, Turner JA, Romano JM.** Validity of the Sickness Impact Profile Roland scale as a measure of dysfunction in chronic pain patients. Pain. 1992;50:157–62.

25. **McHorney CA, Ware JE, Raczek AE.** The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care. 1993;31:247–63.

26. **Ware JE, Gandek B.** The SF-36 Health Survey: development and use in mental health research and the IQOLA Project. Int J Ment Health. 1994;23:49–73.

27. **Fischer D, Stewart AL, Bloch DA, Lorig K, Laurent D, Holman H.** Capturing the patient's view of change as a clinical outcome measure. JAMA. 1999;282:1157–62.

28. **Carey EC, Walter LC, Lindquist K, Covinsky KE.** Development and validation of a functional morbidity index to predict mortality in community-dwelling elders. J Gen Intern Med. 2004;19:1027–1U31.

29. **Coste J, Guillemin F, Pouchot J, Fermanian J.** Methodological approaches to shortening composite measurement scales. J Clin Epidemiol. 1997;50:247–52.

30. **Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA.** The accuracy of pain and fatigue items across different reporting periods. Pain, In Press, Corrected Proof.

31. **Jamison RN, Raymond SA, Slawsby EA, McHugo GJ, Baird JC.** Pain assessment in patients with low back pain: comparison of weekly recall and momentary electronic data. J Pain. 2006;7:192–9.

32. **Bair MJ, Robinson RL, Katon W, Kroenke K.** Depression and pain comorbidity: a literature review. Arch Intern Med. 2003;163:2433–45.

33. **Revicki D, Hays RD, Cella D, Sloan J.** Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008;61:102–9.

34. **Beaton DE, Hogg-Johnson S, Bombardier C.** Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. J Clin Epidemiol. 1997;50:79–93.

35. **Wright JG, Young NL.** A comparison of different indices of responsiveness. J Clin Epidemiol. 1997;50:239–46.

36. **Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH.** Comparative measurement sensitivity of short and longer health status instruments. Med Care. 1992;30:917–25.

37. **Kazis LE, Anderson JJ, Meenan RF.** Effect sizes for interpreting changes in health status. Med Care. 1989;27:S178–89.

38. **Klinkman MS.** Competing demands in psychosocial care. A model for the identification and treatment of depressive disorders in primary care. Gen Hosp Psychiatry. 1997;19:98–111.

39. **Ostbye T, Yarnall KS, Krause KM, Pollak KI, Gradison M, Michener JL.** Is there time for management of patients with chronic diseases in primary care? Ann Fam Med. 2005;3:209–14.

40. **Mitchell AJ, Coyne JC.** Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. Br J Gen Pract. 2007;57:144–51.

41. **Mendoza T, Mayne T, Rublee D, Cleeland C.** Reliability and validity of a modified Brief Pain Inventory short form in patients with osteoarthritis. Eur J Pain. 2006;10:353–61.