



Published in final edited form as:

*J Pain Symptom Manage.* 2009 May ; 37(5): 848–862. doi:10.1016/j.jpainsymman.2008.05.013.

## Development of a Novel Location-Based Assessment of Sensory Symptoms in Cancer Patients: Preliminary Reliability and Validity Assessment

**Adam R. Burkey, MD, MSCE and Peter A. Kanetsky, PhD, MPH**

*Departments of Neurology and Anesthesiology & Critical Care (A.R.B.), University of Pennsylvania School of Medicine, Philadelphia Pennsylvania, USA Center for Clinical Epidemiology and Biostatistics (P.A.K.), University of Pennsylvania School of Medicine, Philadelphia Pennsylvania, USA*

### Abstract

We report on the development of a novel location-based assessment of sensory symptoms in cancer (L-BASIC) instrument, and its initial estimates of reliability and validity. L-BASIC is structured so that patients provide a numeric score and an adjectival description for any sensory symptom, including both pain and neuropathic sensations, present in each of ten predefined body areas. Ninety-seven patients completed the baseline questionnaire; 39 completed the questionnaire on two occasions. A mean of 3.5 body parts was scored per patient. On average, 2.7 (of 11) descriptor categories were used per body part. There was good internal consistency (Cronbach's  $\alpha = 0.74$ ) for a 4-item scale that combined location-specific metrics. Temporal stability was adequate ( $\kappa > 0.50$  and  $r > 0.60$  for categorical and continuous variables, respectively) among patients without observed or reported subjective change in clinical status between L-BASIC administrations. We compared our 4-item scale against scores obtained from validated pain and quality-of-life (QOL) scales, and as expected, correlations were higher for pain-related than for QOL-related items. We detected differences in L-BASIC responses among patients with cancer-related head or neck pain, chemotherapy-related neuropathy and breast cancer-related lymphedema. We conclude that L-BASIC provides internally consistent and temporally stable responses, while acknowledging that further refinement and testing of this novel instrument are necessary. We anticipate that future versions of L-BASIC will provide reliable and valid syndrome-specific measurement of defined clinical pain and symptom constructs in the cancer population, which may be of particular value in assessing treatment response in patients with multiple such complaints.

### Keywords

Cancer pain; reliability; quality of life; location; neuropathy; description

---

Address correspondence to: Adam R. Burkey, MD, MSCE, 2nd Floor Tuttleman Building, 1840 South Street, Philadelphia, PA 19104, USA, E-mail: burkeya@uphs.upenn.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Introduction

Location is considered an important dimension of pain in cancer patients along with other elements such as intensity, temporal pattern, treatment modality, exacerbating/relieving factors, and interference with health-related quality of life (QOL) [1-3]. Often, patients with cancer suffer from pain or other sensory disorders, such as numbness or paresthesias, in more than one body site [4-7]. We are unaware of estimates that detail the contribution of multiple distinct pain or neuropathic problems to the overall pain burden when several occur, or their individual roles as determinants of QOL. While two widely used and validated pain assessment tools, the Brief Pain Inventory (BPI) [8,9] and the McGill Pain Questionnaire (MPQ) [10], allow patients to mark worst pain site on a figure, this does not capture the full symptom distribution or unrelated, secondary complaints.

Several attempts have been made to categorize cancer pain syndromes in order to understand their particular contributions to cancer-related morbidity [11-13]. Typically, investigators schematize empirical syndromes such as visceral distension, pathologic fracture or chemotherapy-related neuropathy, within a biological framework that involves the constructs of nociceptive, neuropathic and “mixed” pain [14-17]. Empirical research, however, has demonstrated difficulty in distinguishing the neuropathic and nociceptive constructs in cancer patients [14,17], which lessens the therapeutic guidance that these constructs could provide. The ability to obtain values specific to clinical syndromes could aid with management decisions and clarify the relationship of cancer pain to QOL, which is poorly understood [18,19]. Some studies indicate no improvement in QOL with reduction of pain severity [20-22], while others suggest a relationship between certain components of the pain experience and QOL [23,24].

As an approach to the measurement of sensory symptoms, including painful nociceptive and painful or non-painful neuropathic symptoms, we piloted L-BASIC, a novel instrument that uses location-specific ratings of sensory symptoms in the cancer population. Prior work with pain drawings indicates that test-retest reliability should be high for both location and total body area in pain [25]. We assessed the temporal stability (reliability) of L-BASIC by correlating body part-specific scores collected approximately two weeks apart in a test-retest paradigm of the same subjects. We also calculated the internal consistency (reliability) of responses for numeric scales and adjectival descriptors completed for each symptomatic body part. We assessed criterion validity by relating L-BASIC scores to scores from well-studied and validated scales that measure pain, pain-related interference with function and general cancer-related QOL. Construct validity was assessed by comparing L-BASIC scores to validated scales among patients within three distinct clinical constructs. Our hypotheses were that scores on L-BASIC would correlate better with pain-related than with QOL measures, and that unrelated clinical syndromes would differ significantly in their location-specific numeric scores and adjectival descriptions, reflecting true differences in their severities and sensory characteristics.

## Methods

### Instrument Development

Patients reported symptoms at specific body parts that were defined both pictorially and by name. This approach differs from previous research methodologies that rely on pain drawings or a pain site checklist to assess symptom location; only the former has been used to assess pain in cancer patients [2]. For pain drawings, respondents indicate specific areas of pain or other sensations by freehand marking or shading within an outline of a human figure [26]. Free-hand shading makes quantifying body area and the assessment of overlap of shaded areas at different timepoints difficult. We avoided this complexity by defining body areas for the patient. We believed that a systematic presentation of individual body parts might provoke a

more detailed expression of sensory symptoms particular to that body area. Terms assigned to the pictorial representation included joint locations (i.e., thigh and knee) and dorsal/ventral distinctions (e.g., chest and upper back) to clarify body part boundaries (Appendix).

The numeric and adjectival descriptor components used in L-BASIC derive from two widely accepted pain questionnaires: the BPI [27,28] and the MPQ [29,30]. The BPI utilizes an 11-point numeric rating scale (NRS) scored from 0-10 with anchors of “no pain” and “pain as bad as you can imagine”. These prompts have shown good reliability and validity in many studies [28,31,32]. L-BASIC provides an NRS for symptom “intensity” and symptom “unpleasantness” that tap the sensory and affective dimensions described for pain experience. Our anchors, however, were written as “No symptoms” and “Symptoms as bad as you can imagine.” We used both terms “pain” and “sensory symptoms” when presenting the study to patients because we wanted them to score abnormal sensations they experienced which might not otherwise be reported if told only to score “pain.” These symptoms predominantly would be due to neuropathy with numbness and non-painful paresthesias.

Adjectival descriptors of sensory symptoms can help distinguish between neuropathic and non-neuropathic pain syndromes [33,34] and whether pains are “more or less” neuropathic [35]. Verbal descriptors on the MPQ have been used to distinguish between diabetic neuropathy and leg pain of other causes [36]; trigeminal neuralgia, temporomandular joint dysfunction and atypical facial pain [37-39]; and reversible and non-reversible dental pain [40]. Our list of sensory descriptors derives from the work of Fernandez and Towery [41], who refined and categorized the sensory descriptors of the original MPQ. The adjective descriptors in L-BASIC are provided weights (mild, moderate or severe) similar to the MPQ short-form [42]. L-BASIC also includes a temporal dimension of constant or occasional as modifiers of the adjectives, as this dimension is considered important to pain measurement in cancer [1,2].

### Study Setting and Population

Our study was conducted at the Rena Rowan Breast Cancer Center, the general oncology clinics and the Penn Pain Medicine Center of the University of Pennsylvania. Patients were enrolled at a clinic follow-up visit, during a chemotherapy session, or at presentation to the Penn Pain Medicine Center. Patients with pain or sensory complaints were identified by their treating physician or by review of clinic charts. Patients were excluded from participation if they were non-native English speaking or had known pre-existing cognitive impairment or intracranial involvement of their disease. Informed consent was obtained from each participant, and the study was approved by the Institutional Review Board of the University of Pennsylvania.

### Data Collection

**Questionnaire Administration**—Study participants were shown the pages of L-BASIC and asked to provide responses on two 0-10 NRS representing the intensity and unpleasantness of any pain or sensory symptom they experienced in the body part corresponding to that page (Appendix). The first NRS asks “How *intense* have the sensory symptoms in this part of your body been on average over the past week?” and the second asks “How *unpleasant* have the sensory symptoms in this part of your body been on average over the past week?” If patients had no symptoms for a given body part, they were instructed to mark both scales zero and go to the next page (i.e., to the next body part). For any symptomatic body part, they were instructed to choose adjective descriptors from a list on the reverse side of the body part page to best describe the symptom (Appendix). The 34 adjectives were presented to them in 11 boxed categories. Patients could write in up to two additional adjective descriptors to articulate their symptom. It was emphasized that they need not linger over every descriptor but rather select the words to which they had an immediate response. Participants were told to qualify their adjective description with a modifier for the severity (mild, moderate or severe) and timing

(constant or occasional). Verbal examples were given of how to think of the adjective scoring, such as “mild aching occasionally” or “severe burning constantly.” No reference was made to any specific complaint the participant may have already volunteered. After instructions, patients were left to complete the questionnaire on their own. Demographic information including visit date, date of birth, sex, ethnicity (Caucasian, African-American, Pacific Islander, Asian-American, Other), and clinical information including diagnosis and stage [43] were collected at the initial visit.

**Follow-Up for Test-Retest Stability Assessment**—Every attempt was made to contact the patient at a follow-up visit approximately two weeks after completion of L-BASIC for re-administration of the instrument. At retest, patients also completed the Patient Global Impression of Change scale (PGIC) [44]: a 7-point balanced Likert scale to indicate their overall subjective impression of change. The PGIC has been recommended for use in chronic clinical pain trials [45] to indicate the participants' assessment of worsening or improvement at successive time points [46,47]. We considered patients indicating “no” or “a little” worsening or improvement as not reporting subjective change and those indicating “significant” or “a lot” of worsening or improvement as reporting subjective change. We abstracted information on interval hospitalizations, new metastases or changes in therapy by review of medical documentation and considered each as evidence of clinical change.

**Criterion Validity Assessment**—Patients completed the BPI and Functional Assessment of Cancer Therapy (FACT-G) scales at each visit. The BPI is the recommended scale for capturing pain-related data in cancer patients [48]. This questionnaire gathers information on worst, least, average, and current pain and pain-related functional interference with certain activities of daily living using 11-point NRS. The BPI provides a global pain score without discriminating among single or multiple complaints. The FACT-G instrument is a 27-item questionnaire composed of four subscales that measure the dimensions of emotional, social, physical and functional well-being. The FACT-G is a widely utilized general quality-of-life instrument for the cancer population.

**Construct Validity Assessment**—We assessed whether L-BASIC provided scores that were specific to three clinically distinct cancer-related pain and sensory syndromes: painful head and neck cancer, breast-cancer related upper extremity lymphedema, and chemotherapy-related neuropathy. The latter two syndromes, which can be considered either painful or nonpainful by individual patients, differed in their nociceptive (lymphedema) and neuropathic (chemoneuropathy) mechanisms. Our hypothesis was that the reported sensory descriptions would differ significantly between these three constructs, and that each construct would show a different pattern of correlation with our criterion scales (BPI and FACT-G).

Head or neck symptoms were described by patients with primary head and neck cancer diagnosed by surgical biopsy. For these patients data on received radiotherapy dosage were collected. Upper extremity edema in breast cancer patients was confirmed by comparison of tissue swelling and fluctuance in the arm ipsilateral to the described symptoms versus the contralateral limb. In most cases, the arm was in pressure wraps and had serial measurements recorded in the charts as part of standard lymphedema care at the University of Pennsylvania under the direction of the Rehabilitation Medicine department. Chemotherapy-related neuropathy was determined by medical documentation of foot symptoms during the course of treatment with platinum-based therapy and/or taxanes [49-54]; and the absence of diabetes, renal insufficiency or other common metabolic cause of distal polyneuropathy. Sensory loss was confirmed in each case by exam with light touch (cotton swab) and sterile pinprick.

## Statistical Analysis

**Reliability, Internal Consistency**—The internal consistency of the L-BASIC was evaluated using Cronbach's  $\alpha$  [55]. All instrument items that could be used to form a scale were used to calculate  $\alpha$ . The best scale was determined by iteratively calculating  $\alpha$  after the removal of one scale item that maximally increased the  $\alpha$  statistic. Factors tested included the numeric values for intensity and unpleasantness; the number of adjectives selected to describe the symptom; the mean of the weights used to modify the adjectival descriptors (scored 1, 2 or 3 for mild, moderate or severe, respectively); and timing (scored 1 or 0 for constant or occasional, respectively).

**Reliability, Temporal Stability**—We used Pearson's pairwise correlation to estimate test-retest stability of the NRS values, the mean adjective severity weights, and the number of descriptors used per body part at the two time-points. For each body part scored at test and retest, a distinct correlation coefficient was calculated. We used the kappa statistic [56] to evaluate the agreement of location and descriptor categories used by a given patient at the two time-points. For each patient, one 2×2 table was constructed with a marginal total of 10 pairs. Two primary analyses were performed. First we analyzed only patients without evidence of clinical change over the test-retest interval. Second, we further removed patients reporting subjective change on the PGIC and repeated all analyses. For each analysis, 2×2 tables were summed across patients to give a statistic for location agreement over time.

We used three coding schemes to determine cell counts for analysis of location data. First, strict agreement was determined. Body parts with a non-zero score for either the intensity or unpleasantness NRS were considered to be symptomatic, and body parts with a zero marked for both of these scales or where the page was left blank were considered asymptomatic. We also determined “relaxed agreement” with the intention to offset the potential difficulty for patients reporting mild symptoms, which may be intrinsically unreliable or fluctuate over time. Here, a score of 0 or 1 on either the intensity or unpleasantness NRS was considered to be non-symptomatic and scores of two or greater were coded as symptomatic. Finally, for “adjacent agreement,” kappa was recalculated using the strict coding scheme, but allowing for agreement in those instances where the same or an adjacent body part was scored at test and retest.

Temporal stability of the descriptor categories used for a particular body part was assessed by the kappa statistic. We also assessed agreement of descriptor category usage on a per-patient basis irrespective of body part recognizing that a particular pain or sensory symptom could span several body parts, and therefore descriptors might reasonably be applied to any or all of the involved body parts at a given administration of the L-BASIC. For this analysis each adjectival descriptor category used at either administration was counted once.

**Criterion Validity**—Using the 4-item scale with good internal consistency (see Results), we defined a global body score as the average of location-specific scale scores across all body parts scored, and worst body part score as the maximum score calculated for one body part per patient. We calculated Pearson's pairwise correlation between these scores and items on the BPI and FACT-G. We used multiple regression analysis to assess the independent association of each of the 4-item scale components with items on the BPI and FACT-G.

**Construct Validity**—All analyses were repeated within patients with head and neck cancer, breast-cancer related upper extremity lymphedema, and chemotherapy-related neuropathy using only location-specific scores and component values from relevant body parts, i.e., from head and neck only, from arm only, and from foot only. *P*-values are adjusted for the false discovery rate [57]. In order to determine if sensory qualities represented by the 11 descriptive qualities differed among the three clinical constructs, a 3×11 contingency table Chi-square

analysis ( $df = 20$ ) for homogeneity was undertaken. We also assessed criterion validity as described above within each construct to determine differences in their relationships to our validating scales.

## Results

### Descriptive Statistics

Ninety-seven patients completed the study questionnaire at first administration. All but 15 patients (four lung cancer; six breast; five head and neck) had stage IIIb disease or worse indicating that their disease was unresectable and/or metastatic. Seven patients with breast cancer, four with lung cancer and one patient with ovarian cancer were African-American; all others were Caucasian. Our study sample consisted of 50 women with breast cancer; seven men and seven women with lung cancer; 25 men and four women with head and neck cancer; one man with prostate cancer; one man with colon cancer; one woman with pancreatic cancer; and one woman with ovarian cancer. The mean age of patients was 54.1 years. The youngest patient was 21 years old and the eldest 85 years old.

A total of 39 patients repeated the questionnaire at a follow-up between 10 and 21 days after the first administration. The median test-retest interval in our patients was 14 days (interquartile range: 14-19.25 days). Seven patients had medical evidence of interval clinical all of whom reported subjective change on the PGIC. Eight additional patients without clinical evidence of change reported interval subjective change on the PGIC.

**L-BASIC Location, Severity, and Adjective Descriptors**—Table 1 gives the frequency with which each body part was scored in our study sample at first questionnaire administration. Sensory symptoms were noted in 343 body parts for an average of 3.5 body parts per patient. At first questionnaire administration, the distribution of intensity and unpleasantness NRS values ranged from 0 to 10. The modal score was eight. The intensity and unpleasantness scores were identical for 223 of the 343 scored body parts. Utilization of adjectival descriptor categories was not uniform in our population (Table 2). “Aching,” “dull,” or “sore” (category I) were the most often cited descriptors, accounting for 28% of the total category usage. Categories with less than 5% of the total usage were “cutting,” “sharp,” or “lacerating” (category VII); “cool,” “freezing,” or “cold” (category IX); and “other” (category XII: patient-specified adjectives).

### Reliability

**Internal Consistency**—Using values on the intensity NRS (0-10), the unpleasantness NRS (0-10), the number of descriptors selected (0-36) and the average descriptor weight (mild=1, moderate=2, severe=3) reported for all 343 body parts, we determined a 4-item scale with good internal consistency (Cronbach's  $\alpha = 0.74$ ). The scale score was computed as the average of the NRS scales multiplied by one-tenth the product of descriptor number and the average descriptor weight.

**Temporal Stability Among Patients Without Interval Clinical Change**—Among the 32 patients without clinical change between L-BASIC administrations, strict and relaxed agreement in the location of reported symptoms was good ( $\kappa=0.76$ ; 95% confidence interval [CI] = 0.66-0.86; and  $\kappa=0.80$ ; 95% CI = 0.70-0.89, respectively; Table 3). Strict agreement improved notably considering adjacent body parts as concordant ( $\kappa=0.88$ , 95% CI 0.80-0.96).

For the 109 concordant body parts reported at test and retest, the intensity NRS values correlated slightly better ( $r = 0.72$ ) than the unpleasantness NRS values ( $r = 0.66$ ). 661 adjectives were selected at test and 636 adjectives at retest to describe the sensory symptoms

in these body parts. In total, 1297 adjective descriptors were distributed among 587 categories for an average of 2.7 categories per body part, with 2.2 adjectives chosen per category. The correlations for the number of adjectival descriptors used per body part ( $r = 0.48$ ) and the mean adjectival descriptor severity weights ( $r = 0.31$ ) were modest between test and retest. Strict body part-to-body part agreement of adjectival descriptor categories used to score the 109 matched regions between time points was only fair ( $\kappa = 0.56$ ; 95% CI 0.50-0.62). Adjectival descriptor category agreement improved considerably when they were analyzed on a per-patient basis irrespective of the body part or number of body parts to which they were applied on a given administration ( $\kappa = 0.68$ ; 95% CI 0.57 – 0.79).

**Temporal Stability Among Patients Without Clinical or Reported Subjective Impression of Change**—For the 24 patients with no clinical or reported interval change between L-BASIC administrations, strict and relaxed agreement was good for the location of reported symptoms ( $\kappa=0.71$ ; 95% CI = 0.58-0.84;  $\kappa=0.73$ , 95% CI 0.60-0.86, respectively; Table 3). Considering adjacent body parts as concordant, agreement improved ( $\kappa=0.86$ , 95% CI 0.74-0.98). Correlations for the intensity and unpleasantness NRS scores for the 80 strictly-concordant body parts were 0.65 and 0.61, respectively (Table 3). Agreement between adjectival descriptor categories used in concordant body parts at test and retest was moderate ( $\kappa= 0.51$ ; 95% CI = 0.44-0.58). Calculating the agreement of categories used at test and retest used by each patient regardless of single or multiple occurrences across body parts improved agreement ( $\kappa=0.64$ , 95% CI 0.57-0.71; Table 3).

## Validity

**Criterion Validity**—Reliability (Cronbach's  $\alpha$ ) of the pain severity and functional interference subscales of the BPI and the physical, social/family, emotional and functional subscales of the FACT-G ranged from 0.72 to 0.93 when these instruments were administered at the same time as the L-BASIC. This indicated the expected, good internal consistency of the items that compose each of these previously validated scales[55]. Responses on these two questionnaires were used to assess the convergent validity with pain scores and divergent validity from QOL scores of responses on the L-BASIC.

**Global Body Score**—Global body score, the average of body part-specific scale scores ranged from 0-22.9 (median 3; interquartile range 1.2-7.1). We detected a significant correlation between the global body score and every item on the BPI pain (Pearson's  $r$  range: 0.47-0.61) and functional interference subscales ( $r$  range = 0.22-0.49).

**Worst Body Part Score**—Scores for worst body part ranged from 0 to 32.4 (median 4.9, interquartile range 2.4-12). Correlations were significant between worst body part score and every BPI item pain subscale item ( $r$  range = 0.54-0.60) and functional interference subscale item except pain-related interference with appetite ( $r$  range = 0.35-0.53). Worst body part score was only associated significantly with the pain item on the physical well-being subscale ( $r = 0.33$ ) of the FACT-G.

**Construct Validity**—The relative frequency of descriptor category utilization differed among the three constructs ( $\chi^2=223$ ;  $P<0.001$ ) suggesting an overall difference in the sensory qualities of the three subgroups. For patients with head and neck cancer ( $n=11$  head;  $n=14$  neck), the L-BASIC score associated significantly with the worst, least, average, and right now pain items on the BPI ( $r$  range = 0.59-0.67). For patients with neuropathy due to chemotherapy for breast cancer ( $n=27$ ), lung cancer ( $n=4$ ), or head and neck cancer ( $n=1$ ), foot scores did not correlate significantly with any item on the BPI or FACT-G. For patients with upper extremity lymphedema secondary to breast cancer treatment ( $n=27$ ), we also found no significant associations with items on the BPI or FACT-G.

## Discussion

Our data suggest that L-BASIC can provide internally consistent and temporally stable measurements. We discuss the results of our test-retest analyses in terms of how the questionnaire's reliability might be improved for future testing. A score calculated from four location-specific metrics correlated well against BPI items and much less so against FACT-G items, supporting convergent validity with a pain construct and divergent validity from a QOL construct. Finally, differences observed in our subgroup analyses of three clinical constructs suggest how a location-specific measurement might be validated in larger, clinically homogenous cohorts.

### Reliability of Location, Severity and Descriptive Metrics

**Location**—We calculated good internal consistency for a four-item body part-specific scale ( $\alpha=0.74$ ); a scale  $\alpha$  between 0.70-0.90 is generally desired, since an  $\alpha$  greater than 0.90 implies redundancy of the items and less than 0.70 indicates a lack of consistency to the construct [58]. Test-retest stability of our instrument was substantial for location ( $\kappa=0.76$ ). An inherent difficulty with test-retest is determining if interval change has occurred within the test-retest period. Often there is no gold-standard objective test for change in the construct being measured. Regarding empirical evidence on retest reliability, one expects a coefficient from high 0.60s to low 0.70s for variable pathological states such as anxiety, or, in our case, pain/sensory symptoms [55].

On average, patients in our study scored between three and four body parts to describe their symptoms. Our data suggest that most patients were describing one or two sensory syndromes across multiple adjacent body parts, and that this phenomenon was detrimental to our test-retest reliability. Allowing adjacent body parts to count as a match at test and retest significantly increased our kappa values in patients with no evidence of interval change, implying that much of the inconsistency of location was due to a shift in the body part scored at retest to a nearest neighboring body part. Furthermore, when descriptive categories were compared at test and retest for patients irrespective of body part, kappa increased to 0.64 (Table 3). Importantly, this improvement represents a shift toward a higher standard category of agreement, from “moderate” to “substantial” [56].

**Location-Specific Severity**—In our study, test-retest correlations for intensity and unpleasantness did not indicate very high temporal stability. The highest correlation was obtained for intensity NRS values in concordant body parts scored by patients with no evidence of interval clinical change ( $r=0.72$ ). When we explored test-retest correlations among patients who reported change on the PGIC,  $r$  was 0.73 for the intensity NRS values, but only reached 0.58 for the unpleasantness NRS (Table 3). This suggests that labile symptom-related unpleasantness may underlie some of the overall impression of change in patients with pain. These considerations may be important to include in the evaluation process if a questionnaire of this type ultimately will be used in longitudinal studies where PGIC scores are obtained [59].

**Location-Specific Description**—Test-retest stability of category utilization per body part was moderate ( $\kappa=0.56$ ) [56]. These descriptive categories reflect the sensory quality of the syndrome and were much more consistently applied by patients considered as a whole rather than to a specific body part. This temporal instability likely reflects the reporting of single clinical syndromes associated with a particular sensory experience that spans several of our predefined body parts. The diffusion of description across adjacent body parts likely also contributes to the poor test-retest correlation for mean adjective severity weights and number of descriptors used per body part.



An overabundance of descriptor choices may also contribute to test-retest instability. Some members of each category might be eliminated without sacrificing the usefulness of that descriptive category, for instance the adjective “beating,” which was selected on only 6% of the occasions that category III was utilized. Some entire categories, like “blinding or smarting” (category V), which accounted for less than 2% of overall category usage, might be omitted entirely. Elimination of infrequently selected adjectives and categories would potentially enhance the clarity of the instrument and reduce the burden to the subjects.

## Validity

**Criterion Validity**—The 4-item pain score for global body and worst body part each associated significantly with nearly every item on the BPI, and with only two items each on the FACT-G, when all three questionnaires were administered together. The strength of the correlations was also uniformly much higher against the BPI items (often  $r > 0.70$ ), while even statistically significant correlations with FACT-G items were generally weak. This suggests that our questionnaire shows a good convergent validity with pain-related data, but not for cancer-related QOL data. This finding is consistent with the several studies that failed to find a close relationship between pain and QOL in cancer patients [18,19]. If we decomposed the 4-item scale into its components, the highest correlations were found to be between the intensity metric and the worst pain item on the BPI for our global, worst body part and head and neck analyses (results not shown). This observation supports prior work that has proposed that worst pain is the most reliable single metric on the BPI for cancer pain, and that has used this item to establish cutpoints for mild, moderate or severe pain based on its relationship with functional interference [60,61].

**Construct Validity**—We chose to examine the clinical constructs of head or neck symptoms with head and neck cancer, arm symptoms with breast cancer-related upper extremity lymphedema, and foot symptoms with chemotherapy-related neuropathy due to their unambiguous clinical identities and high relative frequency within our sample. In direct contrast to head or neck symptoms with head and neck cancer, neither chemotherapy-neuropathy foot scores nor breast-cancer related lymphedema arm scores demonstrated a single significant correlation with any item on the BPI. This supports the notion of a distinction between painful and non-painful symptom capture by L-BASIC. In fact, the individual correlations of head/neck intensity or unpleasantness component values with certain items on the BPI in the head and neck sample were the highest found in our study, which may reflect the specificity of head and neck symptoms for this subgroup that influence BPI scores (results not shown). Finally, the three clinical constructs differed in their use of descriptive categories to convey the sensation of their syndrome in words. Interestingly, this finding held despite the fact that fourteen breast cancer patients contributed scores to both the lymphedema and neuropathy subgroups, implying that subjects can rate the sensory qualities of different syndromes independently.

## Summary and Future Considerations

Given the good reliability of the BPI scales and the NRS scaling method in general [9], we attribute much of the temporal instability in the L-BASIC to the scoring of single symptoms across multiple body parts. We do not suggest that a single sensory disturbance as reported on L-BASIC implies a single clinical etiology, but merely that the patient's experience of a symptom cuts across physical boundaries that our questionnaire defined for them. The temporal stability of all three metrics-- location, severity and description-- might be improved by allowing patients to score each scale once for a cluster of body parts, endorsed by the patient as representing one symptom. Our interpretation of results for construct validity is limited by relatively small sample sizes for our three clinical constructs; however, their distinct correlations with criterion questionnaires and patterns of adjective descriptor usage suggest

that, in a study of adequate power, construct validity could be firmly established for clinically-defined cancer pain or sensory symptoms. A criterion validation procedure in such a study should establish benchmarks against scales specific to that syndrome, as well as a general pain questionnaire; for instance, patients with chemotherapy-related neuropathy might be administered the BPI as well as the Total Neuropathy Scale [62] or FACT-Neurotoxicity [63] to correlate with responses on the L-BASIC. Data from several well-defined oncology populations of adequate sample size will be required before general conclusions about region-specific symptom scores and cancer-related morbidity can be drawn.

## Acknowledgments

The authors would like to acknowledge J. Richard Landis, PhD, John T. Farrar, MD, PhD, and Edmund Weisberg, MS, for their assistance in the planning and write-up of this project. They would also like to thank Drs. Kevin Fox, Tracey Evans, Eric Sherman, John Glick, David Vaughn and the other physicians and staff of the University of Pennsylvania Division of Oncology who allowed us to visit with their patients.

## Appendix: The Questionnaire

---

Head

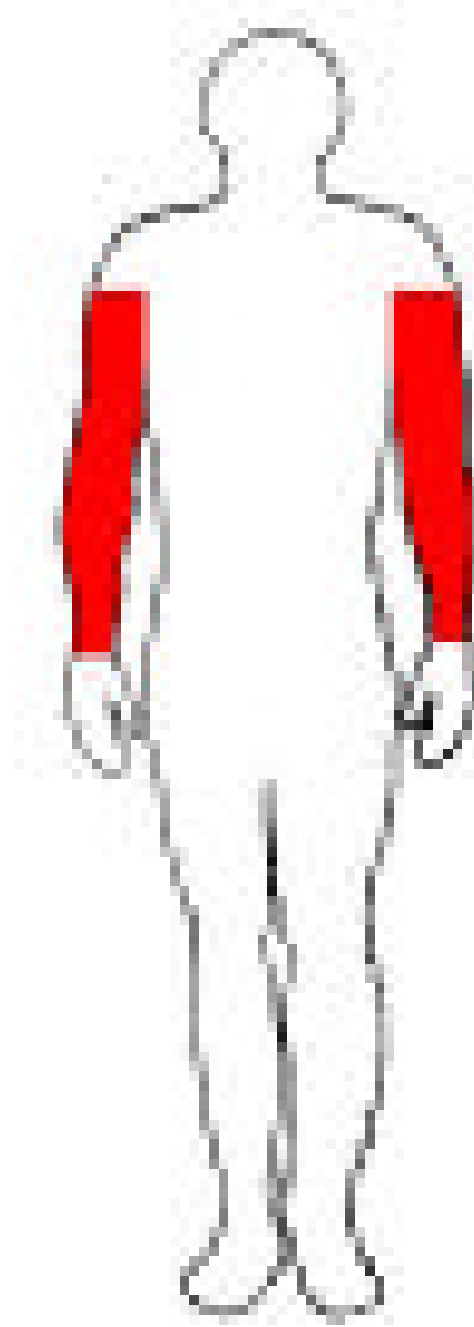


Every page of the questionnaire was identical to the first ("Head") with the exception of the heading and the image, each of which is given below as presented on the other nine pages. The adjective list is provided last as it appeared on the back of each page.

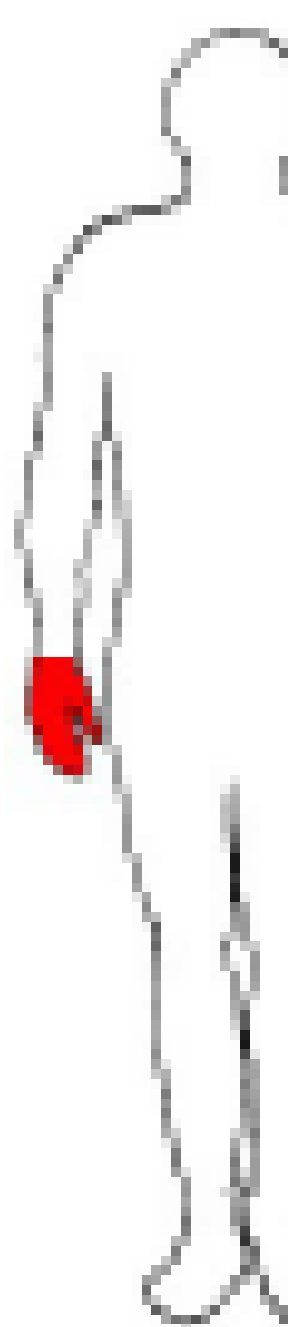
Neck and Shoulders



Arms



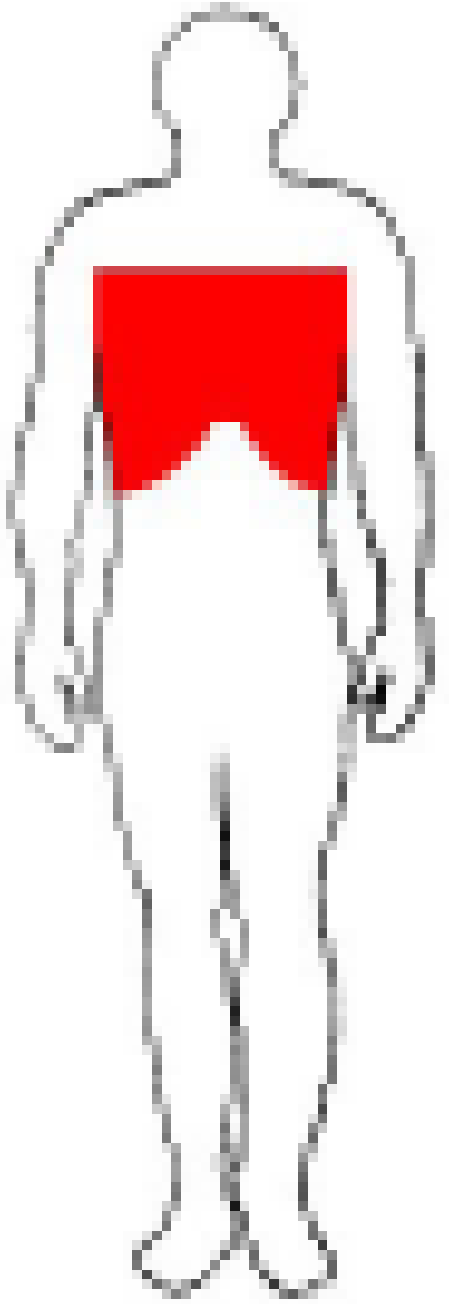
Hands



Chest and Upper Back

Abdomen and Lower Back

Hips and Pelvis

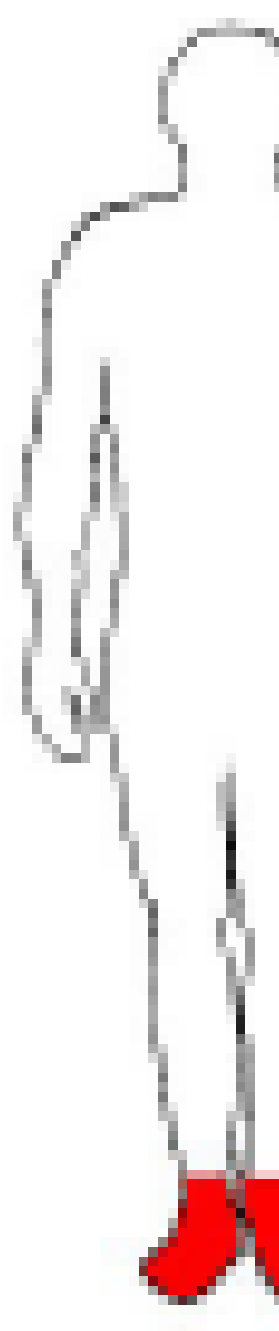


Thighs and Knees



Calves

Feet



---

---

Description of Symptom	Severity of Symptom			Timing of Symptom		
	None	Mild	Moderate	Severe	Constant	Occasional
Aching						

---

---

Description of Symptom	Severity of Symptom				Timing of Symptom	
	None	Mild	Moderate	Severe	Constant	Occasional
Dull						
Sore						
None of the above?						
Burning						
Hot						
Scalding						
Searing						
None of the above?						
Pulsing						
Throbbing						
Pounding						
Beating						
None of the above?						
Numbness						
Tingling						
None of the above?						
Smarting						
Blinding						
None of the above?						
Tugging						
Pulling						
Drawing						
None of the above?						
Cutting						
Lacerating						
None of the above?						
Radiating						
Spreading						
None of the above?						
Cool						
Freezing						
Cold						

Description of Symptom	Severity of Symptom				Timing of Symptom	
	None	Mild	Moderate	Severe	Constant	Occasional
None of the above?						
Pressing						
Crushing						
Squeezing						
Tight						
None of the above?						
Drilling						
Penetrating						
Stabbing						
Piercing						
Pricking						
None of the above?						
Other: _____						
Other: _____						

## References

- Holen JC, Hjermstad MJ, Loge JH, et al. Pain assessment tools: is the content appropriate for use in palliative care? *J Pain Symptom Manage* 2006;32(6):567–580. [PubMed: 17157759]
- Jensen MP. The validity and reliability of pain measures in adults with cancer. *J Pain* 2003;4(1):2–21. [PubMed: 14622723]
- Jensen MP, Dworkin RH, Gammaitoni AR, et al. Do pain qualities and spatial characteristics make independent contributions to interference with physical and emotional functioning? *J Pain* 2006;7(9):644–653. [PubMed: 16942950]
- Grond S, Zech D, Diefenbach C, Radbruch L, Lehmann KA. Assessment of cancer pain: a prospective evaluation in 2266 cancer patients referred to a pain service. *Pain* 1996;64(1):107–114. [PubMed: 8867252]
- Portenoy RK, Miransky J, Thaler HT, et al. Pain in ambulatory patients with lung or colon cancer. Prevalence, characteristics, and effect. *Cancer* 1992;70(6):1616–1624. [PubMed: 1516015]
- Twycross R, Harcourt J, Bergl S. A survey of pain in patients with advanced cancer. *J Pain Symptom Manage* 1996;12(5):273–282. [PubMed: 8942122]
- Twycross RG, Fairfield S. Pain in far-advanced cancer. *Pain* 1982;14(3):303–310. [PubMed: 6218464]
- Cleeland CS, Gonin R, Hatfield AK, et al. Pain and its treatment in outpatients with metastatic cancer. *N Engl J Med* 1994;330(9):592–596. [PubMed: 7508092]
- Daut RL, Cleeland CS, Flanery RC. Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. *Pain* 1983;17(2):197–210. [PubMed: 6646795]
- Melzack R. The short-form McGill Pain Questionnaire. *Pain* 1987;30(2):191–197. [PubMed: 3670870]
- Caraceni A, Portenoy RK. An international survey of cancer pain characteristics and syndromes. IASP Task Force on Cancer Pain. International Association for the Study of Pain. *Pain* 1999;82(3):263–274. [PubMed: 10488677]



12. Caraceni A, Weinstein SM. Classification of cancer pain syndromes. *Oncology (Huntingt)* 2001;15(12):1627–1642. [PubMed: 11780704]discussion 1642-1643, 1646-1647
13. Portenoy RK. Cancer pain: pathophysiology and syndromes. *Lancet* 1992;339(8800):1026–1031. [PubMed: 1349060]
14. Cherny NI, Thaler HT, Friedlander-Klar H, et al. Opioid responsiveness of cancer pain syndromes caused by neuropathic or nociceptive mechanisms: a combined analysis of controlled, single-dose studies. *Neurology* 1994;44(5):857–861. [PubMed: 7514771]
15. Martin LA, Hagen NA. Neuropathic pain in cancer patients: mechanisms, syndromes, and clinical controversies. *J Pain Symptom Manage* 1997;14(2):99–117. [PubMed: 9262040]
16. Stute P, Soukup J, Menzel M, Sabatowski R, Grond S. Analysis and treatment of different types of neuropathic cancer pain. *J Pain Symptom Manage* 2003;26(6):1123–1131. [PubMed: 14654264]
17. Wilkie DJ, Huang HY, Reilly N, Cain KC. Nociceptive and neuropathic pain in patients with lung cancer: a comparison of pain quality descriptors. *J Pain Symptom Manage* 2001;22(5):899–910. [PubMed: 11728793]
18. Portenoy RK. Pain and quality of life: clinical issues and implications for research. *Oncology (Williston Park)* 1990;4(5):172–178. [PubMed: 2143401]discussion 194
19. Strang P, Qvarner H. Cancer-related pain and its influence on quality of life. *Anticancer Res* 1990;10(1):109–112. [PubMed: 2334115]
20. Du Pen SL, Du Pen AR, Polissar N, et al. Implementing guidelines for cancer pain management: results of a randomized controlled clinical trial. *J Clin Oncol* 1999;17(1):361–370. [PubMed: 10458255]
21. Klepstad P, Borchgrevink PC, Kaasa S. Effects on cancer patients' health-related quality of life after the start of morphine therapy. *J Pain Symptom Manage* 2000;20(1):19–26. [PubMed: 10946165]
22. Wong GY, Schroeder DR, Carns PE, et al. Effect of neurolytic celiac plexus block on pain relief, quality of life, and survival in patients with unresectable pancreatic cancer: a randomized controlled trial. *JAMA* 2004;291(9):1092–1099. [PubMed: 14996778]
23. Hwang SS, Chang VT, Fairclough DL, Kasimis B. Development of a cancer pain prognostic scale. *J Pain Symptom Manage* 2002;24(4):366–378. [PubMed: 12505205]
24. Rustøen T, Moum T, Padilla G, Paul S, Miaskowski C. Predictors of quality of life in oncology outpatients with pain from bone metastasis. *J Pain Symptom Manage* 2005;30(3):234–242. [PubMed: 16183007]
25. Margolis RB, Chibnall JT, Tait RC. Test-retest reliability of the pain drawing instrument. *Pain* 1988;33(1):49–51. [PubMed: 3380550]
26. Margolis RB, Tait RC, Krause SJ. A rating system for use with patient pain drawings. *Pain* 1986;24(1):57–65. [PubMed: 2937007]
27. Shi Q, Wang XS, Mendoza TR, Pandya KJ, Cleeland CS. Assessing persistent cancer pain: a comparison of current-pain ratings and pain recalled from the past week. *J Pain Symptom Manage*. 2008;pub Aug 1
28. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med Singapore* 1994;23(2):129–138. [PubMed: 8080219]
29. Graham C, Bond SS, Gerkovich MM, Cook MR. Use of the McGill pain questionnaire in the assessment of cancer pain: replicability and consistency. *Pain* 1980;8(3):377–387. [PubMed: 7402695]
30. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1(3):277–299. [PubMed: 1235985]
31. Chang VT, Hwang SS, Kasimis B. Longitudinal documentation of cancer pain management outcomes: a pilot study at a VA medical center. *J Pain Symptom Manage* 2002;24(5):494–505. [PubMed: 12547049]
32. Portenoy RK, Payne D, Jacobsen P. Breakthrough pain: characteristics and impact in patients with cancer pain. *Pain* 1999;81(12):129–134. [PubMed: 10353500]
33. Bennett MI, Attal N, Backonja MM, et al. Using screening tools to identify neuropathic pain. *Pain* 2007;127(3):199–203. [PubMed: 17182186]

34. Galer BS, Jensen MP. Development and preliminary validation of a pain measure specific to neuropathic pain: the Neuropathic Pain Scale. *Neurology* 1997;48(2):332–338. [PubMed: 9040716]
35. Bennett MI, Smith BH, Torrance N, Lee AJ. Can pain be more or less neuropathic? Comparison of symptom assessment tools with ratings of certainty by clinicians. *Pain* 2006;122(3):289–294. [PubMed: 16540249]
36. Masson EA, Hunt L, Gem JM, Boulton AJ. A novel approach to the diagnosis and assessment of symptomatic diabetic neuropathy. *Pain* 1989;38(1):25–28. [PubMed: 2780059]
37. Melzack R, Terrence C, Fromm G, Amsel R. Trigeminal neuralgia and atypical facial pain: use of the McGill Pain Questionnaire for discrimination and diagnosis. *Pain* 1986;27(3):297–302. [PubMed: 3808740]
38. Mongini F, Italiano M. TMJ disorders and myogenic facial pain: a discriminative analysis using the McGill Pain Questionnaire. *Pain* 2001;91(3):323–330. [PubMed: 11275390]
39. Mongini F, Italiano M, Raviola F, Mossolov A. The McGill Pain Questionnaire in patients with TMJ pain and with facial pain as a somatoform disorder. *Cranio* 2000;18(4):249–256. [PubMed: 11202844]
40. Grushka M, Sessle BJ. Applicability of the McGill Pain Questionnaire to the differentiation of ‘toothache’ pain. *Pain* 1984;19(1):49–57. [PubMed: 6739113]
41. Fernandez E, Towery S. A parsimonious set of verbal descriptors of pain sensation derived from the McGill Pain Questionnaire. *Pain* 1996;66(1):31–37. [PubMed: 8857629]
42. Melzack R. The short-form McGill Pain Questionnaire. *Pain* 1987;30(2):191–197. [PubMed: 3670870]
43. Greene, FL.; Page, DL.; Fleming, ID. *AJCC cancer staging manual*. Vol. 6th. New York: Springer-Verlag; 2002.
44. Guy, W. *ECDEU assessment manual for psychopharmacology* (DHEW Publication No ADM 76-338). Washington, DC: US Government Printing Office; 1976.
45. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;113(12):9–19. [PubMed: 15621359]
46. Dunkl PR, Taylor AG, McConnell GG, Alfano AP, Conaway MR. Responsiveness of fibromyalgia clinical trial outcome measures. *J Rheumatol* 2000;27(11):2683–2691. [PubMed: 11093454]
47. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001;94(2):149–158. [PubMed: 11690728]
48. Caraceni A, Cherny N, Fainsinger R, et al. Pain measurement tools and methods in clinical research in palliative care: recommendations of an Expert Working Group of the European Association of Palliative Care. *J Pain Symptom Manage* 2002;23(3):239–255. [PubMed: 11888722]
49. Hensing TA, Peterman AH, Schell MJ, Lee JH, Socinski MA. The impact of age on toxicity, response rate, quality of life, and survival in patients with advanced, Stage IIIB or IV nonsmall cell lung carcinoma treated with carboplatin and paclitaxel. *Cancer* 2003;98(4):779–788. [PubMed: 12910523]
50. Cavaletti G, Bogliun G, Marzorati L, et al. Peripheral neurotoxicity of taxol in patients previously treated with cisplatin. *Cancer* 1995;75(5):1141–1150. [PubMed: 7850713]
51. Hilkens PH, van den Bent MJ. Chemotherapy-induced peripheral neuropathy. *J Peripher Nerv Syst* 1997;2(4):350–361. [PubMed: 10975744]
52. Quasthoff S, Hartung HP. Chemotherapy-induced peripheral neuropathy. *J Neurol* 2002;249(1):9–17. [PubMed: 11954874]
53. Lipton RB, Apfel SC, Dutcher JP, et al. Taxol produces a predominantly sensory neuropathy. *Neurology* 1989;39(3):368–373. [PubMed: 2564647]
54. Siegal T, Haim N. Cisplatin-induced peripheral neuropathy. Frequent off-therapy deterioration, demyelinating syndromes, and muscle cramps. *Cancer* 1990;66(6):1117–1123. [PubMed: 2169332]
55. Streiner, DLL.; Norman, GR.; Norman, G. *Health measurement scales: A practical guide to their development and use*. New York: Oxford Medical Publications; 2003.
56. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159–174. [PubMed: 843571]

57. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med* 1990;9(7):811–818. [PubMed: 2218183]
58. Bland JM, Altman DG. Cronbach's alpha. *BMJ* 1997;314(7080):572. [PubMed: 9055718]
59. Turk DC, Dworkin RH, Burke LB, et al. Developing patient-reported outcome measures for pain clinical trials: IMMPACT recommendations. *Pain* 2006;125(3):208–215. [PubMed: 17069973]
60. Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. *Pain* 1995;61(2):277–284. [PubMed: 7659438]
61. Paul SM, Zelman DC, Smith M, Miaskowski C. Categorizing the severity of cancer pain: further exploration of the establishment of cutpoints. *Pain* 2005;113(12):37–44. [PubMed: 15621362]
62. Cavaletti G, Bogliun G, Marzorati L, et al. Grading of chemotherapy-induced peripheral neurotoxicity using the Total Neuropathy Scale. *Neurology* 2003;61(9):1297–1300. [PubMed: 14610145]
63. Cella D, Peterman A, Hudgens S, Webster K, Socinski MA. Measuring the side effects of taxane therapy in oncology: the functional assesment of cancer therapy-taxane (FACT-taxane). *Cancer* 2003;98(4):822–831. [PubMed: 12910528]

**Table 1**  
Body Part Counts in Which Pain Was Reported on the L-BASIC in 97 Patients

Page number on L-BASIC	Body part	Counts
1	Head	31
2	Neck/Shoulders	56
3	Arm	31
4	Hand	42
5	Chest/Upper Back	40
6	Abdomen/Lower Back	31
7	Hips/Pelvis	30
8	Thigh/Knee	31
9	Calf	18
10	Foot	33
	Total	343

**Table 2**  
Counts of Adjective Descriptor Selection on the L-BASIC for 343 Body Parts in 97 Patients

Category	Category Count <sup>a</sup>	Adjective Descriptor	Adjective Descriptor Count
<b>I</b>	264	Aching	220
		Dull	96
		Sore	128
		Searing	9
		Scalding	3
<b>II</b>	56	Hot	17
		Burning	47
		Warm	1
		Pulsing	45
<b>III</b>	95	Throbbing	74
		Beating	6
		Pounding	24
<b>IV</b>	122	Tingling	81
		Numb	104
<b>V</b>	16	Blinding	4
		Smarting	14
<b>VI</b>	48	Pulling	35
		Drawing	9
		Tugging	25
<b>VII</b>	12	Cutting	9
		Sharp	1
		Lacerating	3
<b>VIII</b>	90	Radiating	77
		Spreading	46
		Cool	11
<b>IX</b>	31	Freezing	8
		Cold	20
		Crushing	11
<b>X</b>	96	Squeezing	22
		Pressing	18
		Tightness	74
		Drilling	4
		Penetrating	47
<b>XI</b>	106	Pricking	12
		Stabbing	64
		Piercing	45
<b>XII</b>	12	Other	13
<b>Total Count</b>	948		1427

<sup>a</sup> A category count is one if one or more adjective descriptor is chosen from that category by a patient for a given body part.

**Table 3**  
Measures of Agreement for Temporal Stability in Three Patient Groups Defined by Objective and Subjective Change

Group	Group Size	Location <sup>a</sup>		Adjacent criterion	Severity <sup>b</sup>		Descriptor Category <sup>c</sup>	
		Strict criterion	Relaxed criterion		Intensity NRS	Unpleasantness NRS	Site-specific	Per patient
No objective change	32	0.76	0.80	0.88	0.72	0.66	0.56	0.68
No subjective or objective change	24	0.71	0.73	0.86	0.65	0.61	0.51	0.64
Subjective change, regardless of objective change	15	0.54	0.65	0.56	0.73	0.58	0.67	0.69

<sup>a</sup> Agreement of pain location is measured by the kappa statistic. Strict indicates that any non-zero severity score for a body part is coded as symptomatic; relaxed indicates that a severity score of 2 or greater severity for a body part is coded as symptomatic; adjacent indicates that neighboring body parts each scored similarly using the strict criteria is coded as concordant.

<sup>b</sup> Agreement of pain severity is measured by Pearson's  $r^2$ . All severity correlations are determined using NRS scores obtained under the strict location criterion.

<sup>c</sup> Agreement of descriptor categories is measured by the kappa statistic. All descriptor category kappas are determined for concordant body parts and on a per patient basis using each category counted once if used to describe any body part.