# SELECTED SECTIONS OF PRO-CTCAE TECHNICAL PROPOSAL
## ~TABLE OF CONTENTS~

A. STATEMENT OF WORK

1. Abstract .....

2. Objectives .....

2.a. Overall Objective .....
2.b. Specific Tasks to Achieve Objective .....
2.c. Organizational Model and Leadership to Achieve Tasks .....
2.d. Discrete Deliverables .....
2.e. Unique Qualifications of Leadership Team .....
2.f. Areas Beyond Scope of Proposal .....

3. Background and Significance .....

4. Rationale for Approach .....

4.a. Limitations of Current CTCAE and Proposed Strategies for PRO-CTCAE .....
4.b. Statement of Need for PRO-CTCAE .....
4.c. Integration with Related Projects and Infrastructure .....
4.d. Conceptual Framework .....

5. Substantive Approach and Methods .....

5.a. Substantive Approach to Task 1: Prepare Background Report on Barriers/Strategies .....
5.b. Substantive Approach to Task 2: Develop PRO-CTCAE Items .....
5.c. Substantive Approach to Task 3: Assure Cultural Literacy and Lack of Bias .....
5.d. Substantive Approach to Task 4: Conduct Cognitive Interviews .....
5.e. Substantive Approach to Task 5: Develop PRO-CTCAE Software System .....
5.f. Substantive Approach to Task 6: Conduct Usability Testing .....
5.g. Substantive Approach to Task 7: Evaluate Measurement Properties of PRO-CTCAE .....
5.h. Substantive Approach to Task 8: Design Feasibility Study in Cooperative Groups .....
5.i. Substantive Approach to Task 9: Create Training and Educational Materials .....

6. Example Approaches to Fatigue and Neuropathy .....

6.a. Example Substantive Approach to Fatigue .....
6.b. Example Substantive Approach to Neuropathy .....

7. Schedule: Timeline of Tasks and Deliverables .....

E. LITERATURE CITED .....

G. APPENDIX

12.b. Appendix 2: Optional Studies .....

# A. STATEMENT OF WORK

## 1. ABSTRACT

The current approach to monitoring adverse symptom events in NCI-sponsored treatment trials is clinical staff reporting, using items from the Common Terminology Criteria for Adverse Events (CTCAE). Of the 1,059 discrete items in the CTCAE, approximately 77 contain a subjective component sufficient enough to be amenable to patient self-reporting. Prior and ongoing research on patient-reported outcome (PRO) approaches to the CTCAE, most of it conducted by the offeror, has been prompted by a recognition of methodologic problems with the current collection of CTCAE symptom data; increased availability of technologies that facilitate patient self-reporting; and issuance of an FDA guidance that recommends standards for PROs in the regulatory setting. In response to the NCI's RFP, the offeror proposes to work collaboratively with the NCI and other stakeholders to create a PRO version of the CTCAE (PRO-CTCAE). The offeror will work with an experienced commercial software vendor and a multidisciplinary team of investigators who have expertise in informatics, psychometrics, health services research, clinical trials, and clinical oncology to create a PRO-CTCAE platform.

To achieve this work and result, the offeror proposes a series of interrelated tasks with clear benchmarks and discrete deliverables. These tasks include cognitive interviewing and item development, assessment of the psychometric properties of items once developed, usability testing, and, ultimately, feasibility testing in a range of treatment settings among diverse groups of patients. Ready access to large and diverse populations of patients and well-established collaborations with NCI cooperative groups enables the offeror to accomplish the specified tasks and develop a valid, reliable, sensitive, and feasible PRO-CTCAE system.

## 2. Objectives

### *2.a. Overall Objective*

To create a mechanism for patient self-reporting of adverse symptom events in cancer treatment trials. This mechanism, to be called the PRO-CTCAE,* will consist of an integrated item bank and electronic platform that is validated, reliable, sensitive, generalizable, feasible, accepted by clinicians and regulators, and interoperable with NCI and industry databases.

### *2.b. Specific Tasks to Achieve Objective*

To achieve this objective, over the 2 years of the project, we will perform the following Tasks:

**Specific Task 1 (Year 1)—Background Report.** Produce a report titled *Ideal Conditions, Potential Barriers, and Strategies to Implement the PRO-CTCAE.* The report will be based on the literature review and content included in this Technical Proposal, and key informant interviews and will serve as the blueprint for Tasks 2 through 9.

**Specific Task 2 (Year 1)—Item Development.** Identify CTCAE items that are amenable to patient self-reporting, determine which items require additional clinician judgment after patient self-report, determine what concepts are appropriate for measurement by PRO-CTCAE items, and create draft items based on analysis of existing instruments, systematic literature review, mining of existing data sets, and key stakeholder feedback.

**Specific Task 3 (Years 1 and 2)—Cultural Literacy.** Use techniques to ensure that PRO-CTCAE items are developed at an appropriate educational level, are designed and evaluated to avoid bias (eg, age, sex, ethnicity, education, geography), and are structured to be optimally amenable to language translation.

**Specific Task 4 (Year 1)—Cognitive Interviews.** Design and conduct cognitive interviews of diverse patient populations and clinical staff members, to ensure comprehension of draft PRO-CTCAE items, and to ensure that terms map to intended concepts.

**Specific Task 5 (Year 1)—Platform Development.** Develop an open-source software system for patient self-reporting of adverse symptom events that capitalizes on prior evaluations of similar platforms, adheres to standards and federal regulations for data security and privacy, is mapped to standardized lexicons, and is interoperable with current and emerging platforms for adverse event and data management in cancer trials (CDUS, AdEERS, MedDRA, caAERS/CTCAE v4, caBIG/NCICB, CDISC).

---

*Key terms, abbreviations, and acronyms used in this proposal are defined in section G. Glossary.

**Specific Task 6 (Year 1)—Usability Testing**. Conduct usability testing of the PRO-CTCAE platform (electronic and paper) in diverse patient populations and with clinical staff members to ensure intuitive navigation/usability and evaluate for respondent fatigue toward platform refinement.

**Specific Task 7 (Year 2)—Measurement Properties.** Systematically evaluate in diverse patient populations the validity, reliability, and sensitivity of the PRO-CTCAE items, the clinical significance of score changes for each item, and appropriate recall period.

**Specific Task 8 (Year 2)—Feasibility Design.** Develop a protocol in the NCI cooperative group setting to evaluate the feasibility of PRO-CTCAE items nested within treatment trials in patient populations with specific expected adverse symptom events and administered at cancer centers and Community Clinical Oncology Program (CCOP) sites.

**Specific Task 9 (Year 2)—Training Materials**. Create paper and electronic training manuals and demonstrations for the PRO-CTCAE platform for both patients and clinical staff.

### *2.c. Organizational Model and Leadership to Achieve Tasks*

Figure A.2-1 shows organizational relationships relevant to completion of the objective and specific Tasks. Timely completion of all Tasks and deliverables is overseen by the principal investigator, Dr. Ethan Basch, and his project staff at the Coordinating Center at Memorial Sloan-Kettering Cancer Center (MSKCC). This includes administrative oversight/support (project manager, data manager, budget manager, contracts manager, medical editor) and in-house methodologic expertise to ensure the scientific integrity of the Tasks. For each specific Task, a leader is assigned from one of the six participating sites (Duke Cancer Center, Dana-Farber/Harvard Cancer Center, The University of Texas M.D. Anderson Cancer Center, Mayo Clinic, and University of Pennsylvania) to direct and coordinate Tasks and ensure timely deliverables to designated NCI staff. Leaders at all participating sites (except University of Pennsylvania) will coordinate IRB approval of proposed protocols and facilitate access to patients for proposed studies at main hospital and community satellite clinic sites. All designated leaders are subcontractors on this proposal (see Letters of Commitment). Advisory leadership from key stakeholder groups, including representatives of cooperative groups, the FDA, patient advocacy, pharmaceutical industry, community oncology practices, CDISC, and ePRO vendors will provide planned structured input (see Letters of Support). Regular interaction with NCI staff is incorporated into the design and the implementation of specific Tasks, with particular focus on eliciting feedback from Dr. Gordon Willis regarding design/analysis of cognitive interviews, Ann Setser and Alice Chen regarding design of PRO-CTCAE items and development of CTCAE v4/caAERS, John Speakman regarding informatics design and data standards, Lori Minasian and Ann O'Mara regarding integration of PRO-CTCAE evaluation into cooperative group research and CCOPs, and Bryce Reeve regarding item design and evaluation of measurement properties (as well as overall project conceptualization and progress).

**Figure A.2-1. Organizational Model and Leadership to Achieve Tasks**

**Abbreviations:** caBIG, Cancer Biomedical Informatics Grid; CALGB, Cancer and Leukemia Group B; CDISC, Clinical Data Interchange Standards Consortium; CTEP, Cancer Therapy Evaluation Program at NCI; DCP, Division of Cancer Prevention at NCI; DFCI Core, Dana-Farber/Harvard Cancer Center Core Research Facility; ePRO, electronic patient-reported outcomes; MDACC, M.D. Anderson Cancer Center; MSKCC, Memorial Sloan-Kettering Cancer Center; NCCTG, North Central Cancer Treatment Group; NCI, National Cancer Institute; NCICB, National Cancer Institute Center for Bioinformatics; RTOG, Radiation Therapy Oncology Group; SEALD, Study Endpoints and Label Design team at FDA; UPENN, University of Pennsylvania.

### *2.d. Discrete Deliverables*

For each specific Task, the designated leader is responsible for ensuring timely deliverables (Table A.2-1). Progress of each specific Task and discrete deliverable is tracked by the Coordinating Center administrative team. A detailed timeline for required deliverables to NCI staff with delivery months is provided in Section 7 of this proposal.

**Table A.2-1. Discrete Deliverables (See Section 7 for Timeline)**

| TASK | BRIEF DESCRIPTION | DISCRETE DELIVERABLES TO NCI |
|---|---|---|
| 1 | Background Report | • Summary of key informant interviews<br>• Adaptation of this Technical Proposal into draft, revised, and finalized background report<br>• Plan for best practices |
| 2 | Item Development | • Identification/prioritization of CTCAE items for adaptation to PRO-CTCAE<br>• Categorization of items by degree of subjective/objective content to determine level of clinician involvement in evaluation<br>• Categorization of items as core (required), disease specific, or optional<br>• Generic structure for of PRO-CTCAE items<br>• Review of existing source instruments for terms in PRO-CTCAE items<br>• Draft PRO-CTCAE items<br>• List of ideal measurement properties of PRO-CTCAE items<br>• List of target patient populations for evaluating PRO-CTCAE items |
| 3 | Cultural Literacy | • Report of methods integrated into Tasks 2, 4, 5, 6, 7, 8, and 9 to ensure appropriate educational level of items, evaluation of items in diverse patient populations, lack of bias (age, sex, ethnicity, education, geography), and adaptability for language translation |
| 4 | Cognitive Interviews | • Summary reports of cognitive interviews with patient demographics<br>• Modified draft of PRO-CTCAE items |
| 5 | Platform Development | • System security plan<br>• Blueprint for PRO-CTCAE electronic platform data flow among patients, clinicians, data managers, investigators, and NCI databases<br>• Design of interface graphics and functionality for patient input pages, clinical staff pages, and reporting functions<br>• Functional open-source PRO-CTCAE platform integrating publishing tool for generating electronic and paper CRFs, patient interface, clinical staff interface, query/reporting engine, rules-based bridge to AdEERS/caAERS for expedited reporting, interoperability with key databases/reporting standards, and |

| | | |
|---|---|---|
| | | adaptability for future handheld device and telephone-based administration<br>• Source code/object code/system delivery |
| 6 | Usability Testing | • Summary reports of usability testing<br>• Documentation of resulting system modifications<br>• Refined PRO-CTCAE platform based on results |
| 7 | Measurement Properties | • List of target patient populations for evaluating measurement properties<br>• Draft and final study designs (validity, reliability, sensitivity, etc)<br>• Multicenter protocol for assessing measurement properties<br>• Status reports<br>• Report of results of multicenter study of measurement properties<br>• Refined PRO-CTCAE items based on results |
| 8 | Feasibility Design | • List of target patient populations for feasibility study<br>• Identification of CALGB and RTOG upcoming trials in which to nest feasibility study<br>• Identification of cancer center and community sites for feasibility study<br>• Feasibility protocol: Amendment of current CALGB study of patient self-reported adverse symptom events (CALGB 70501) to incorporate PRO-CTCAE items and platform and initiation of RTOG protocol review (see Appendix) |
| 9 | Training Materials | • Draft, revised, and final paper and electronic training manuals for patients<br>• Draft, revised, and final paper and electronic training manuals for clinical staff |

### *2.e. Unique Qualifications of Leadership Team*

The leadership team is uniquely qualified and positioned to perform the specific Tasks and produce the deliverables within the allotted period. Each member brings specific skills and expertise that are complementary and that correspond to essential components of the proposed work. Each has a track record of high-quality investigation, consistent adherence to deadlines, deep relationships with key stakeholders, and a collaborative ethos. Team members have previously conducted work similar or identical to each step in the specific Tasks. Therefore, the proposed work is familiar territory and applicable research questions and methods are in place. This will ensure that each Task is performed appropriately and efficiently to successfully achieve the overall objective, yielding a usable PRO-CTCAE at the end of the allotted project period, consistent with the specifications of the RFP.

Dr. Basch is the only investigator known to have developed and evaluated a patient adaptation of CTCAE symptom items, called “STAR” (Symptom Tracking and Reporting) items. His programming team at MSKCC developed an online platform (the “STAR” platform) for patient self-reporting of these items, which since 2003 has undergone usability testing, feasibility evaluation, and comparisons with clinician CTCAE reporting in more than 1000 patients treated at a tertiary cancer center. He is national study chair of a multicenter feasibility study of patient self-reporting of CTCAE symptom items in the NCI cooperative group, the Cancer and Leukemia Group B (CALGB). He chairs an ongoing working group to develop a PRO model for adverse symptom event reporting, consisting of representatives from CTEP, the FDA, patient advocates, symptom researchers, and health literacy experts. Dr. Schrag is a medical oncologist and health services researcher with deep relationships in the cooperative groups and ASCO. She has served as Dr. Basch’s close collaborator in all aspects of his PRO research program. Drs. Cleeland and Sloan are international leaders in the development and evaluation of symptom instruments for use in cancer research. Drs. Bruner and Trotti have longstanding involvements with the development and evolution of the CTCAE. Dr. Abernethy, also a medical oncologist, has extensive experience conducting usability testing and instrument evaluations via e/Tablets in cancer outpatients. She focuses on how to make these systems provide the highest-quality meaningful data and thus to provide value in the clinical environment. Dr. Viswanath is an expert in cultural literacy and development of educational materials for patients with cancer. Dr. Hay and Mr. Shiman bring expertise in cognitive interviewing techniques. SemanticBits, as the software company that holds the caBIG/NCICB contract to develop caAERS (cancer-specific Adverse Events Reporting System), is familiar with the CTCAE and its regulatory framework. The company is thus uniquely positioned to integrate the PRO-CTCAE with CTCAE v4 and other applicable databases.

### *2.f. Areas Beyond Scope of Proposal*

There are additional areas of evaluation and development work applicable to the PRO-CTCAE that are beyond the scope of this 2-year proposal. These areas include: adaptation/evaluation of PRO-CTCAE items for reporting by children and their proxies (although we have included experts in this area to advise us during item development, Drs. Steve Joffe and Joanne Wolfe; see Letters of Support); adaptation of PRO-CTCAE items for reporting by patients with cognitive impairments (Dr. Viswanath will provide advice on techniques to create PRO-CTCAE items amenable to adaptation for this population); language translation of PRO-CTCAE items into Spanish and French (linguistics expertise from the Dana-Farber Core is included to assure that verbiage and structure in developed items are optimized towards future translation); development of handheld (PDA) and telephony (interactive voice response) versions of the PRO-CTCAE; development of guidelines for clinical trial reporting of adverse symptom events; and design of a drug label structure that includes both PROs and clinical staff reports for adverse symptom events.

## 3. Background and Significance

As described in the RFP "Background Statement," in NCI-sponsored treatment trials, clinicians are required to report patients' symptom using items in the CTCAE (Trotti et al, Semin Radiat Oncol 2003). The CTCAE includes 1,059 items, of which approximately 77 contain sufficient subjective content so as to be amenable to patient self-reporting. Clinical staff reporting of patient symptoms via the CTCAE in this context requires a complex sequence of data transfer (Figure A.3-1) and has been criticized as being error-prone, inefficient, and not reflective of the true patient experience during treatment (Fairclough, *Qual Life Stud Clin Trials* 2002; Bentzen et al, Semin Radiat Oncol 2003).

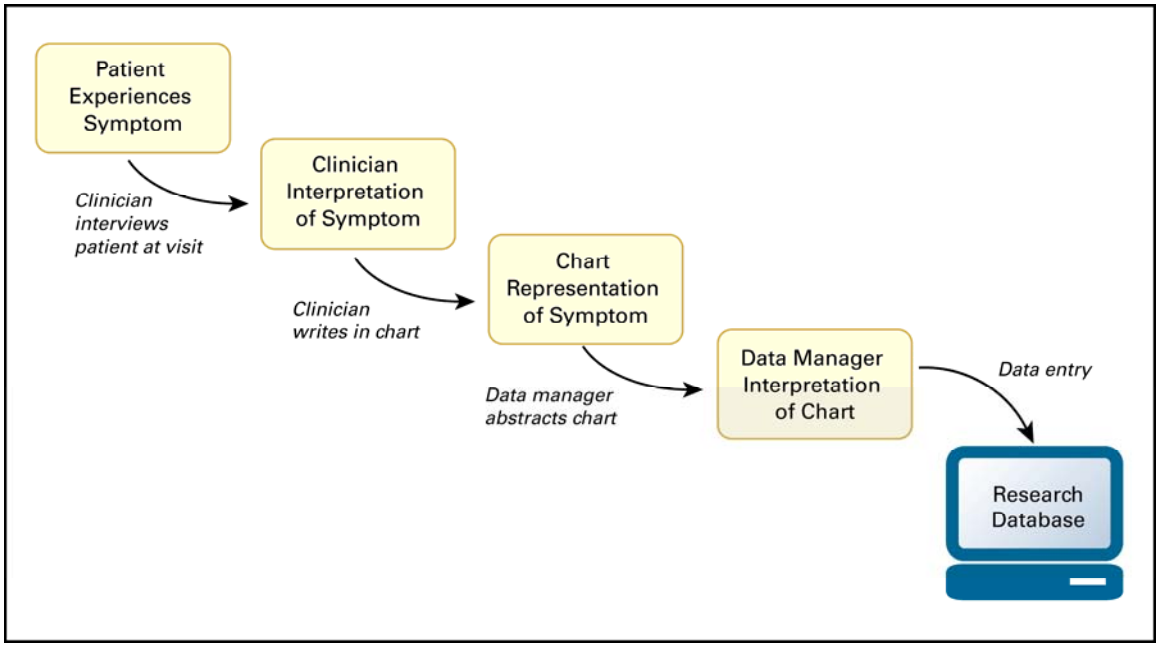



**Figure A.3-1. Multiple Steps of Data Transfer in Current Adverse Symptom Event Reporting in NCI-Sponsored Treatment Trials**

As an alternative, patient self-reporting may provide a more comprehensive picture of patients' lives, foster patient-clinician communication, and improve data completeness, administrative efficiency, and clinical outcomes (Velikova et al, J Clin Oncol 2004; Science Panel Inter Comm and Health, April 1999; Robinson et al, JAMA 1998; Fowler et al, Urology 1993; Sloan et al, J Clin Epidemio 2005). Indeed, patient-reporting is considered the gold standard for data collection in health-related quality-of-life and symptom research (Sneeuw et al, J Clin Epidemio 2002; Patrick et al, J Natl Cancer Inst 2003; Fromme et al, J Clin Oncol 2004). However, the CTCAE was not created for patient comprehension or administration (Gwede et al, J Oncol Manag 2002). Furthermore, items in the clinician version of the CTCAE itself were not created with patient input, have not undergone cognitive debriefing for refinement, and have not been evaluated for validity, reliability, sensitivity, the clinical significance of score changes over time, or recall period. Therefore, any effort to create a patient version of the

CTCAE must turn back to first principles. The work proposed in this application is intended to do just that. We begin by questioning fundamental structural characteristics of existing CTCAE items, and capitalize on several years' experience by Dr. Basch's team at MSKCC evaluating patient adaptations of the CTCAE (ie, "STAR" items). The tasks outlined in the following sections of this proposal each encompass an essential step towards creating a PRO-CTCAE item bank and assessment system which is valid, reliable, sensitive, interoperable with relevant information systems and versions of the CTCAE (including v4.0), secure, and feasible in cancer treatment trials across a diverse patient population. This work is intended to have a widespread positive impact by expediting the detection of potentially serious adverse events, improving the quality and comprehensiveness of adverse event data in publications and drug labels, and ultimately improving the patient experience during cancer care.

## 4. Rationale for Approach

### *4.a. Limitations of Current CTCAE Symptom Reporting Model and Proposed Strategies for the PRO-CTCAE*

The current approach to adverse symptom reporting in NCI-sponsored clinical trials via the CTCAE is subject to multiple limitations related to the process by which the items were developed historically (by consensus); the format of response options (which is aimed to be similar to nonsubjective CTCAE items rather than to follow more common formats used for symptom assessment); a lack of evaluation of measurement properties of items; reporting of items by clinical staff rather than by patients; and inconsistencies in how data are collected and reported among trials and across institutions (Trotti et al, J Clin Oncol 2007; Trotti, Semin Radiat Oncol 2002; Gwede et al, J Oncol Manag 2002; Trotti et al, Semin Radiat Oncol 2003; Trotti et al, J Clin Oncol 2004). Details regarding the nature of these limitations and rationales for the proposed strategies, and the specific Tasks in which the limitations will be addressed, are described below in Table A.4-1.

**Table A.4-1. Limitations of the Current Approach to Adverse Symptom Event Reporting in CTCAE v3 and Proposed Strategies**

| LIMITATIONS OF CURRENT APPROACH | PROPOSED STRATEGIES FOR PRO-CTCAE DEVELOPMENT | PERTINENT TASK(S) |
|---|---|---|
| **1.** Lack of patient input when items were developed/refined | Review existing items/instruments that were developed with patient input (and patient interview data sets/terminology preferences used to develop those items/instruments). Include patient advocate input throughout PRO-CTCAE system development. Conduct formal cognitive interviews during PRO-CTCAE item refinement across diverse patient populations (age, sex, ethnicity, geography, educational level). | Task 2<br>Task 4 |
| **2.** Lack of systematic literature reviews when items were developed | Review published literature and existing instruments during PRO-CTCAE item development. | Task 2 |
| **3.** Double-barreling of concepts in response options (ie, intensity, interference, frequency) and inconsistent combining of concepts in response options among items | Determine through consensus with NCI staff and advisors early in contract period, and literature review, whether double-barreling of concepts in current CTCAE symptom item responses should be dispensed with in PRO-CTCAE items. | Task 2 |
| **4.** Ordinal response scale is limited in range | Determine through consensus with NCI staff and advisors early in contract period, and literature review, whether it is preferable to keep a 0-4 scale for the PRO-CTCAE, or adopt a 0-10 scale (or verbal descriptor scale (none, mild, moderate, severe, very severe). | Task 2 |

| | | |
|---|---|---|
| **5.** Response options are anchored to descriptions of discrete phenomena/ interventions that are not always mutually exclusive or do not encompass all patient situations | Consult with health literacy expertise regarding feasibility of adapting such item constructions towards future language translations, and administration to children/proxies or to patients with cognitive impairment. | Task 2<br>Task 3 |
| **6.** Symptom items are reported by clinical staff not by patients | Develop patient version of PRO-CTCAE items that contain subjective components. Expand current CALGB protocol assessing the feasibility of this approach. Develop PRO-CTCAE training materials for patients. | Task 2<br>Task 8<br>Task 9 |
| **7.** Selection of which symptoms to report is not systematic; often ad hoc | Identify mandatory "core" and "diagnosis-specific" CTCAE symptom items for inclusion in all trials based on frequency in prior trials and literature review. Develop PRO-CTCAE platform that allows patients to add additional "optional" items in real-time, which are then required for reporting at subsequent visits. | Task 2<br>Task 5 |
| **8.** Unclear distinction between an <u>unreported</u> symptom grade (missing data) and <u>lack</u> of a symptom (asymptomatic) at some institutions | Require inclusion of grade 0 in reporting of all PRO-CTCAE items. Include mandatory "core" and "disease-specific" items in CRFs to ensure representation of such symptoms in reporting, thus limiting concerns that symptoms may be present but underreported. | Task 2<br>Task 5 |
| **9.** Multiple steps in data transcription/transfer | Develop electronic platform for direct data entry into a database, the PRO-CTCAE platform, and refine functionality via usability testing. | Task 5<br>Task 6 |
| **10.** Measurement properties of items not established | Measure validity, reliability, and sensitivity of items. Assess clinical significance of score changes over time. | Task 7 |
| **11.** Variable recall periods among trials; recall bias | Review literature and existing data sets regarding adequate symptom recall periods. Preliminary study of recall period in Task 7. | Task 7 |
| **12.** No reporting between visits | Depending on results of recall assessments, between-visit reporting for the PRO-CTCAE may be warranted. Study of such a model is beyond the scope of this proposal, but the PRO-CTCAE platform will be designed for future integration of portable data entry devices. | Future |
| **13.** No automated alert system for high-grade | Include rules-based notifications in the PRO-CTCAE assessment system, with automated alerts triggered by | Task 5 |

| | | |
|---|---|---|
| adverse symptoms | specific grade thresholds. Determine cut-point scores for required notifications of clinicians based on mining of existing data sets and a planned study to establish cut points in Task 7. | Task 7 |
| **14.** No automated integration with expedited reporting systems | Integrate PRO-CTCAE platform with caAERS and AdEERS and with CDUS. Determine score cut points for required expedited reporting, based on mining of existing data sets, and a planned study in Task 7. | Task 5<br>Task 7 |
| **15.** No assurance that items are generalizable to diverse patient populations or minimize bias | Conduct cognitive interviews across diverse patient populations (age, sex, ethnicity, geography, educational level) during item development and refinement. During design, consider necessity of future language translations and administration in diverse cultural contexts, and to children and patients with cognitive impairment. | Task 2<br>Task 4 |
| **16.** Lack of harmonization or linkage of adverse symptom event items with clinical trial symptom endpoint items/instruments | Include input from FDA SEALD team during item development and consider methodologic guidelines in FDA Draft Guidance on PROs during item development and evaluation in Tasks 2 and 7. Develop PRO-CTCAE system that can use PRO-CTCAE items to "screen" for symptoms and thus trigger other items or questionnaires or PROMIS tools when reports pass a prespecified severity threshold. | Task 2<br>Task 7 |

### 4.a.1. Lack of patient input when current CTCAE symptom items were developed and refined: content validity

The current CTCAE items were developed and refined through a committee consensus process (Trotti et al, J Clin 2007; Trotti, Semin Radiat Oncol 2002; Trotti et al, Semin Radiat Oncol 2003; Trotti et al, J Clin Oncol 2004). This process did not involved explicit patient interviews or focus groups to ensure that the terms used in questions or that the design or concepts included in response options are acceptable to patients. In the FDA Draft Guidance on PROs (and in the final guidance document, which is currently in final edits at the FDA), it is emphasized that patient input is essential to the development of items that assess patient subjective phenomena to ensure "content validity" (U.S. Department of Health and Human Services, 2006). Therefore, any future iteration of CTCAE items that contain a subjective component should include patient input during the design and item development process to ensure the perceived legitimacy of these items and more importantly to ensure ultimate patient acceptance, comprehension, and feasibility of administration.

As noted in Table A.4-1, our proposed Tasks include specific steps that integrate the patient perspective: inclusion of patient advocate advisers throughout the project (including structured feedback from the Research Communications Subcommittee at CALGB, which largely consists of patient advocates and of which Dr. Basch is a

standing member), and ongoing feedback from patient advocate advisers Cindy Georghegan and Deb Collyar.

**4.a.2. Lack of systematic literature reviews when items were developed**

Current CTCAE symptom items were not developed on the basis of systematic reviews of scientific literature or reviews of existing instruments or items. As a result, these items do not capitalize on the rich existing knowledge base from past symptom research (Kirkova et al, J Clin Oncol 2006) or established processes for instrument development (U.S. Department of Health and Human Services, 2006). As noted in Table A.4-1, our proposed approach includes literature reviews and reviews of existing instruments/item sets. We will thus ensure that PRO-CTCAE items are developed with the benefit of this past work. Multiple investigators on this proposal are familiar with this body of literature, as well as with systematic review techniques.

**4.a.3. Double-barreling of concepts in response options (ie, intensity, interference, frequency) and inconsistent combining of concepts in response options among items**

Examination of current CTCAE v3 items demonstrates that response options in most of the symptom items contain more than one concept. Table A.4-2 includes several CTCAE v3 items and details the concepts contained in response options for each. For example, the “fatigue” item asks respondents to consider both severity and interference with function when assigning a grade: fatigue grade 2 is defined as “moderate or causing difficulty performing some activities of daily living (ADL),” whereas grade 3 is defined as “severe fatigue interfering with activities of daily living (ADL).” This *double-barreling* of concepts within single response options can create a quandary for respondents if a patient does not meet the requirements for *both* concepts: eg, if a patient is believed to have “severe” fatigue but there is no perceived interference with ADL or the patient has “mild” fatigue that nonetheless is considered to be “disabling” (the grade 4 criterion), then that patient will defy the available response options and will be forced into a category that may not reflect the true patient experience with treatment.

**Table A.4-2. Example CTCAE v3 Item Response Options and Double-Barreling of Concepts in Each**

| CTCAE TERM | GRADE 1 | GRADE 2 | GRADE 3 | GRADE 4 | *CONCEPTS IN ITEM* |
|---|---|---|---|---|---|
| **Fatigue** | Mild fatigue over baseline | Moderate or causing difficulty performing some ADL | Severe fatigue interfering with ADL | Disabling | *Severity*<br>*Interference* |
| **Diarrhea** | Increase of <4 stools per day over baseline; mild increase in ostomy output compared with baseline | Increase of 4-6 stools per day over baseline; IV fluids indicated <24 hours; moderate increase in ostomy output compared with baseline; not interfering with ADL | Increase of ≥7 stools per day over baseline; incontinence; IV fluids ≥24 hours; hospitalization; severe increase in ostomy output compared with baseline; interfering with ADL | Life-threatening consequences (eg, hemodynamic collapse) | *Frequency*<br>*Interference*<br>*Severity* |
| **Neuropathy-sensory** | Asymptomatic; loss of deep tendon reflexes or paresthesia (including tingling) but not interfering with function | Sensory alteration or paresthesia (including tingling), interfering with function but not interfering with ADL | Sensory alteration or paresthesia interfering with ADL | Disabling | *Severity*<br>*Interference* |
| **Nausea** | Loss of appetite without alteration in eating habits | Oral intake decreased without significant weight loss, dehydration or malnutrition; IV fluids indicated <24 hours | Inadequate oral caloric or fluid intake; IV fluids, tube feedings, or TPN indicated ≥24 hours | Life-threatening consequences | *Severity*<br>*Interference*<br>*Duration* |
| **Dizziness** | With head movements or nystagmus only; not interfering with function | Interfering with function but not interfering with ADL | Interfering with ADL | Disabling | *Severity*<br>*Interference* |

Furthermore, assessment of interference is inconsistent and unclear across CTCAE items. For example, grade 2 "dizziness" is defined as "interfering with function but not with ADL." The distinction between interference with *function* and interference with *ADL* may be lost on respondents (guidelines provided by CTEP state that "*function* refers to the purpose/role of a human body part or organ," whereas "*activities of daily living* refer to the performance of the basic activities of self care"; such distinctions are likely artificial for many symptoms). In addition, the application of this distinction is inconsistent across items, with grade 2 for some items specifying "no interference with ADL" (eg, diarrhea item) and others specifying "difficulty performing some ADL" (eg, fatigue item). Although it may ultimately be reasonable for levels of severity to differ among items, there is no systematic approach or rationale for these differences, which might be applied when creating new items or developing a PRO version of the CTCAE.

*"Severity" vs. "Intensity:"* Notably, an emerging issue in the symptom research community is whether there is a distinction between "intensity" and "severity." Methodologic issues related to this possible distinction are discussed in detail in Section 5.b.6 of this proposal. In general, in this application, when referring to the results of symptom research, we will use the term "intensity." However, in our discussion of creating PRO-CTCAE items which ask patients to globally rate the magnitude of an experience without specific reference to a concept (such as "interference" or level of bother), we will use the term "severity." Unfortunately, due to ambiguity regarding the distinction between these terms at this time, in cases when PRO-CTCAE development is being discussed in the context of prior research or publications, we will use both terms or include one of the terms in parenthesis following the other.

In previous development work, Dr. Basch's team at MSKCC created initial pilot patient versions of CTCAE symptom items ("STAR" Items), which mirrored the concepts in response options of the existing CTCAE items. This approach was adopted at the time because it was believed most likely to be accepted by clinical investigators and regulators (the PRO approach to adverse event reporting was novel at the time) and to allow for comparisons of patient vs. clinician reporting of items with response options containing analogous concepts (Basch et al, J Clin Oncol 2005; Basch et al, J Clin Oncol 2007; Basch et al, Lancet Oncol 2006). However, subsequently, health literacy experts and cognitive interviewing experts participating in Dr. Basch's monthly Working Group for the Development of a PRO Version of the CTCAE (including healthy literacy expert, Wendy Mettger, patient advocate, Cindy Geoghegan, and cognitive interviewing expert, Dr. Jennifer Hay) believed that inclusion of multiple concepts would present a problem in patient-reported items (ie, would not be accepted or clearly comprehended by many patients) and would present challenges when translating items into other languages or adapting items for administration to children or patients with cognitive disabilities.

This specific issue was addressed by the investigators before the issuance of this RFP during meetings of Dr. Basch's Working Group, which has met monthly by teleconference since 2006. During these discussions, members of the FDA SEALD team and the FDA Oncology Division have expressed concerns about combining more than one concept in a response option and have stated a clear preference to limit item response options to a single concept.

Therefore, from an item development perspective, there are three different options for future PRO-CTCAE item response formats:

- Response Format Option 1: Keep the current format of the MSKCC initial pilot versions of patient CTCAE items ("STAR" Items), which mirror existing CTCAE items and contain multiple concepts per item.

- Response Format Option 2: Develop a PRO-CTCAE model in which more than a single item is administered for each symptom. For example, for nausea an item might first be administered to measure severity/intensity, and if that item screens positive (ie, grade >0), then additional item(s) would be triggered to assess interference and/or frequency.

- Response Format Option 3: Develop an initial PRO-CTCAE model which only includes items that measure severity without any additional concepts (such as interference, frequency, duration, location, bother, alleviation with interventions, or attribution). But include a built in function in the electronic PRO-CTCAE assessment system to trigger additional items or questions (such as PROMIS tools) based on a positive "screen" towards future development in this direction.

On the basis of preliminary work, the investigators are inclined toward Response Format Option 3.

- The investigators are inclined against Response Format Option 1 for the reasons outlined above: double-barreling may leave some patients without a clear response option that describes their experience and may make items challenging to reliably translate into other languages or adapt for children or patients with cognitive disabilities. Furthermore, complex response options tend to contain more verbiage and hence would necessitate longer items, which take respondents longer to complete, hence likely limiting the total number of screening PRO-CTCAE items that could be feasibly administered in a single CRF.

- The investigators are inclined against Response Format Option 3 because initial research suggests that the incremental additional information provided by

measurement of concepts beyond severity or interference is not sufficient enough to justify the additional questionnaire burden on patients for the purpose of adverse event monitoring. For example, in pilot work measuring the correlation between intensity/severity items and interference items for pain and fatigue in 120 cancer patients with bone metastases at MSKCC, the Kendall tau rank correlation coefficient was 0.76 ($p = 0.00003$) (Basch, unpublished data), similar to findings in published evaluations (Hadi et al, J Palliati Med 2008) and findings by Dr. Cleeland through analysis of data sets from the Brief Pain Inventory and Brief Fatigue Inventory. Furthermore, in an analysis in 60 patients at MSKCC who responded to double-barreled "STAR" Items (ie, adaptations of current CTCAE symptom items with inclusion of multiple concepts in response options) and single "severity" items with grading via a 0-10 NRS, Kendall tau values ranged from 0.72 through 0.9, suggesting that the incremental added value of concepts beyond intensity is limited. Although the extent to which level of correlation should be considered sufficient to dispense with a particular concept from screening questions is a matter of judgment, it is the consensus of the investigators that a correlation >0.7 (variable depending on method of measurement, but the principle holds across methods) is sufficient to reasonably dispense with a concept for the purpose of adverse event monitoring to ensure that questionnaires remain brief enough to be feasible for widespread use. Nonetheless, a technology platform is envisioned that in the future would allow for additional items to be triggered by positive severity items to characterize specific symptoms in greater depth if desired.

- The investigators are inclined toward Response Format Option 3, to design a PRO-CTCAE item bank with a single item per symptom, which measures severity only, with the intention of *screening* rather than in-depth symptom characterization. The investigators have extensive experience with the development and assessment both of multi-item symptom questionnaires (particularly Dr. Cleeland) and single-item measures (particularly Dr. Sloan) and thus are in a strong position to evaluate the options and design scientifically appropriate PRO-CTCAE items (Sloan et al, Mayo Clin Proc 2002; Huschka et al, Cancer 2007).

The investigators propose that a consensus process be employed early in the contract period including the study team, NCI staff, and selected advisors/key stakeholders, to determine the response option format to be used for the PRO-CTCAE, based on best available evidence.

**4.a.4. Ordinal response scale is limited in range**

As indicated in Table A.4-2, CTCAE items are limited to grades 0-5 (grade 5 generally is defined as "death" and therefore is not relevant for patient self-administered items). It is not clear to the investigators whether a 0-4 NRS provides sufficient granularity to characterize severity of symptoms in *screening* questions. Assuming that Response Format Option 3 (above) is selected for the PRO-CTCAE items (ie, severity only is

assessed), then the investigators concur that there are three potential approaches to the scale:

- Scale Approach 1: 0-4 NRS (none → worst I can imagine)
- Scale Approach 2: Verbal Descriptor Scale: none, mild, moderate, severe, very severe
- Scale Approach 3: 0-10 NRS (none → worst I can imagine)

The investigators are inclined toward Scale Approach 3 because it provides more granular severity data and because the authors have access to abundant prior cancer patient-reported data sets using 0-10 scales for myriad symptoms via the Brief Pain Inventory, Brief Fatigue Inventory, and MDASI (all which were developed by Dr. Cleeland's group at MDACC). Drs. Sloan and Basch also have access to large data sets using 0-10 response scales. There is also a general preference in the symptom assessment literature and community for use of 0-10 NRS. Furthermore, FDA SEALD has expressed a clear preference for 0-10 NRS for single symptom items over scales with fewer response options, and thus, a 0-10 NRS approach would harmonize the PRO-CTCAE with preferences at the FDA in the context of labeling. However, an advantage of Scale Approach 1 (0-4 NRS) is consistency with the scale used to measure other CTCAE items, which therefore would avoid any potential confusion because some CTCAE items are being graded on a scale of 0-4 (or 0-5) and others are being graded on a scale of 0-10. An advantage of Scale Approach 2 (verbal descriptor scales) for severity measurement is clarity of reporting: reporting the proportion of patients with "mild" or "moderate" symptoms may be more easily understandable to patients considering use of a therapy, whereas grade 2 may have less face validity to a lay reader (although 0-10 NRS is widely familiar and accepted by patients; for example, it is the standard recommended approach to pain assessment for inpatients). The investigators propose that a consensus process be employed early in the contract period including the study team, NCI staff, and selected advisors/key stakeholders, to determine which response scale will be used for the PRO-CTCAE, based on best available evidence.

**4.a.5. Response options are anchored to descriptions of discrete phenomena/ interventions that are not always mutually exclusive or do not encompass all patient situations**

Described above in 4.1.4.

**4.a.6. Symptom items are reported by clinical staff not by patients**

As demonstrated in prior work by Dr. Basch, clinician CTCAE grading is discrepant from patient CTCAE reporting, with patients tending to report higher levels of symptom

severity more frequently and earlier during care than paired clinicians (Basch et al, Lancet Oncol 2006; Basch, ASCO 2008), consistent with prior published work (Vogelzang et al, Semin Hematol 1997; Talcott et al, J Natl Cancer Inst 1997; Fromme et al, J Clin Oncol 2004; Litwin et al, J Urol 1998). If the patient perspective is taken to be the gold standard for subjective experiences such as symptoms (U.S. Department of Health and Human Services, 2006), then by comparison, clinicians are systematically underreporting the severity and frequency of patients' symptoms in clinical research. To meet the emerging standard of patient self-reporting for subjective phenomena in clinical research, a PRO approach for CTCAE items with subjective components is essential. Notably, a part of this emerging standard is the expectation that items will be developed with scientific rigor, as planned in the Tasks of this proposal.

**4.a.7. Selection of which symptoms to report is not systematic; often ad hoc**
In any given treatment trial, it is up to the discretion of the clinical investigator or sponsor to determine which CTCAE items will be required in a CRF at planned visits. There exists no "core" group of adverse events that is required for assessment across trials. As a result, it is difficult to compare the frequency of specific adverse symptom events across trials: when a particular symptom is not reported, it is not clear if that symptom was truly not present (ie, grade = 0) or if it simply wasn't assessed. This problem is compounded by the way in which CTCAE assessment is operationalized in real-world settings: a common scenario is that a clinician interviews an enrolled patient during a visit with no specific adverse events in mind, and subsequently a data manager abstracts the medical chart, seeking verbiage specific to the particular adverse events in a study CRF. However, symptoms not included in a CRF may be neglected, or discussed symptoms may never make it into the patient's chart.

Furthermore, discussion of specific adverse symptom events is not mandated during visits in treatment trials. Rather, discussion of symptoms is left to the discretion of the clinician and patient, who may have other topics on their minds, such as imaging results or financial issues, so that symptoms that are present fail to be discussed during visits. This ad hoc approach further exacerbates underreporting of symptoms in trials.

Patient self-reporting with mandatory inclusion of specific items in all administrations addresses these limitations. In Task 2, we will utilize literature search, data mining of prior clinical trial adverse event reporting databases, and expert consultation to identify "core" PRO-CTCAE items to be used in all clinical trials. We will also identify "diagnosis-specific" PRO-CTCAE items, which will be mandatory based on disease diagnosis or drug class. Other PRO-CTCAE items will be classified as "optional," for clinicians to add to PRO-CTCAE CRFs at the time of questionnaire creation or for patients to select from a checklist/dropdown menu at the time of CRF administration. Development of the technology infrastructure for operationalizing this approach is incorporated into Task 5.

**4.a.8. Unclear distinction between an unreported symptom grade (missing data) and lack of a symptom (asymptomatic) at some institutions**

As noted above in section 4.1.7., many treatment trial CRFs and institutional data interfaces do not include a CTCAE response category of grade 0. It is often assumed that lack of reporting of a CTCAE grade for an adverse symptom event signifies that the symptom is not present. However, it may simply reflect failure to elicit that symptom from a study subject or failure of a data manager to abstract that item from the medical chart. To address this limitation, PRO-CTCAE CRFs will include grade 0 (none) as a response option in all items to distinguish between missing data and lack of a symptom.

**4.a.9. Multiple steps in data transcription/transfer**

As noted above, in Section 3 ("Background and Significance"), the current sequence by which adverse symptoms events are reported in treatment trials is complex, with multiple steps of data transfer and vulnerability to errors of misinterpretation or omission (Gwede et al, J Oncol Manag 2002). An underlying hypothesis of this proposal is that having patients report directly into a clinical trials database will reduce data entry errors and improve efficiency by eliminating several of these steps.

**4.a.10. Measurement properties of items not established**

The measurement properties of CTCAE symptom items have not been assessed formally to ensure content or construct validity, reliability, or sensitivity. These and other measurement properties will be assessed in Task 7 of this proposal.

**4.a.11. Variable recall periods among trials; recall bias**

The periodicity of CTCAE assessments is currently variable among trials, with no specific requirements for frequency of reporting. Therefore, one treatment trial may require adverse event reporting at visits every 3 weeks, whereas a different trial may require reporting at visits every 6 weeks. The timing of reporting generally corresponds to the time points of required visits in a trial (often determined by treatment schedules). The sometimes long durations between CTCAE assessments is germane to adverse symptom reporting in particular, because respondents are generally instructed to retrospectively document the worst severity grade for any given item during the period since the last reporting time or treatment, which may be a several week period. However, there is increasing evidence that patients have difficulty remembering symptom levels beyond several days, with the number of days of relative accuracy of retrospective symptom reports likely variable among symptoms and within the same symptom depending on the instrument used (Broderick et al, Pain 2008). In Task 7 of this proposal, we will conduct preliminary evaluations of the recall period for the developed PRO-CTCAE items (comparing 1-month vs. 2-week vs. 24-hour recall). A formal in-depth study of recall periods using between-visit reporting of PRO-CTCAE items is envisioned in the future.

**4.a.12. No reporting between visits**

One strategy to limit degradation of symptom reporting accuracy due to long recall periods is to allow patient self-reporting between visits. Dr. Basch has previously conducted feasibility studies of between-visit reporting via home computers and found variable rates of voluntary self-reporting adherence without reminders, ranging from 15% to 75% (Basch et al, J Clin Oncol 2005; Basch et al, J Clin Oncol 2007). Ongoing research at MSKCC is assessing if adherence increases with implementation of email reminders and backup telephone surveys. More likely, future iterations of the PRO-CTCAE will involve the use of portable handheld media such as PDAs or IVR via cell phones. The planned PRO-CTCAE platform in Task 5 has been designed for smooth migration of CRFs to portable hardware for CRF administration. In addition, multiple investigators on this proposal have past experiences working with ePRO vendors and momentary reporting via PDA and IVR (Drs. Abernethy, Basch, and Cleeland), and this proposal includes a plan for advisory input from ePRO vendors, who bring a rich experience with these approaches (led by adviser Greg Sadowski), and industry advisers with experience using portable media to measure symptoms and HRQL as endpoints in treatment trials (led by Joe Cappelleri). Development of the PDA and IVR versions of the PRO-CTCAE is beyond the scope of this 2-year project but is envisioned by the investigators in future work.

**4.a.13. No automated alert system for high-grade adverse symptoms**

There is currently no automated system in the CTCAE to alert investigators (or other clinicians) when a patient is found to have a serious adverse symptom event. The importance of pharmacovigilance and attention to prospective capture and dissemination of adverse drug reaction data has been emphasized and demonstrated in other contexts (Bennett et al, Arch Intern Med 2007; Goldberg et al, J Clin Oncol 2002; Sargent et al, J Natl Cancer Inst 2000). Integration of a rules-based notification system, which will be similar in design to the approach developed by SemanticBits for caAERS, is planned for integration into the PRO-CTCAE platform in Task 5. This system will allow for integration of rules to send emails or text messages to designated investigators or clinicians when patients self-report symptom severities above a predetermined threshold.

Determinations of threshold cut points for PRO-CTCAE item alerts will be based on a specific study in Task 2 and data obtained in Task 7, as well as literature review and the mining of rich data sets at M.D. Anderson, Mayo, and MSKCC (Palos et al, J Pain 2006; Serlin et al, Pain 1995; Anderson, Pain 2005; Jensen et al, Pain 2001; Paul et al, Pain 2005; Zelman et al, Pain 2003; Mendoza et al, Clin J Pain 2004; McConahay et al, Acad Emerg Med 2006; Chow et al, J Palliat Med 2006; Dihle et al, Clin J Pain 2006; Jensen et al, Pain 2001). Such an alert system for patient-reported CTCAE items has been piloted in the MSKCC “STAR” system since 2003, with a high level of clinician acceptance and satisfaction with the email-based notifications system and approach (Basch et al, J Am Med Inform Assoc 2007; Basch et al, J Clin Oncol 2005). Dr.

Cleeland's group has also conducted extensive work on the determination of symptom cut points for numerous symptoms in the MDASI, Brief Pain Inventory, and Brief Fatigue Inventory (also with IVRs and alert setting).

**4.a.14. No automated integration with expedited reporting systems**

The caERS platform developed by SemanticBits integrates rules-based notifications that automatically trigger and populate online forms for expedited reporting if specific predetermined adverse event severities are reported. This same approach will be used for the PRO-CTCAE platform, which will be fully integrated with caAERS and AdEERS during Task 5 of this proposal. As such, if specific grade thresholds are reported that meet criteria for expedited reporting in the rules for a particular trial, an email alert will be sent to the designated investigator with a link to a required expedited reporting form. If forms are not completed in a timely fashion, repeat alerts to the investigator and supervisory personnel will be triggered.

**4.a.15. No assurance that items are generalizable to diverse patient populations or minimize bias**

Comprehension of concepts and measurement properties of current CTCAE symptom items have not been evaluated in diverse populations to assess for generalizability or systematic bias based on age, sex, ethnicity, educational level, geographic location, or other characteristics. Such an assessment is particularly relevant if these items will be self-reported by patients in the future via the PRO-CTCAE system. Therefore, we consider it vital to consider cultural differences and to use plain language during creation of items (Task 3) and to conduct cognitive interviews (Task 4), usability testing (Task 6), and assessment of measurement properties (Task 7) in a diverse, widely representative population. Dr. Viswanath and his team will participate in multiple Tasks of this proposal to ensure consideration of cultural diversity and health literacy, to minimize bias, and to design items amenable to future language translation and adaptation for administration to children and patients with cognitive disabilities.

**4.a.16. Lack of harmonization or linkage of adverse symptom event items with clinical trial symptom endpoint items/instruments**

In cancer treatment trials that include symptom PRO or health-related quality of life (HRQL) endpoints, there is effectively double-reporting of symptom severity information. For example, a study of a novel drug regimen in patients with metastatic prostate cancer to bone may include a "pain palliation" endpoint (eg, measured via the Brief Pain Inventory at each visit) and a "fatigue improvement" endpoint (eg, measured via the Brief Fatigue Inventory), but simultaneously clinicians will be rating these same symptoms as potential adverse events via the CTCAE. Although in theory these phenomena may differ in terms of their attribution (etiology) to disease effect vs. toxic effect of a drug vs. some combination of both, making such distinctions about the cause of the symptoms is known to be challenging to clinicians, with likely discrepant reporting of attribution between clinicians and patients (Golomb et al, Drug Saf 2007; Rogers, Drug Intell Clin Pharm 1987). This is an area of active research.

Regardless of the cause of specific symptoms, however, a particular CTCAE item that asks for the worst a symptom has been may correspond to some extent with similar items in a symptom endpoint questionnaire. Therefore, in the future, if patients become the respondents for both the symptom endpoints and adverse symptom events, it would be prudent to harmonize the approaches through the use of adverse symptom event items in the PRO-CTCAE that adhere to the same standards of item format, scale, and measurement properties that the FDA Draft Guidance on PROs recommends for measures of symptom/HRQL endpoints (U.S. Department of Health and Human Services, 2006). Such an approach does not require resolution of the debate about whether adverse event items can be designed to isolate attribution due to toxic drug effects. Rather, this line of thinking provides an additional rationale for adopting a 0-10 NRS for the PRO-CTCAE items because it is the consensus of the investigators that this approach would increase the likelihood of such information being accepted by the FDA for labeling purposes (although this issue will be explored with representatives of the FDA during Task 1 of this project and through patient-based research studies in Task 2 during item development).

***4.b. Statement of Need for PRO-CTCAE***

Based on limitations of the current CTCAE approach to adverse symptom event reporting, there is clearly a need for a revised approach that includes patient input during item development; capitalizes on prior instrument development research; ensures patient comprehension and acceptance of terms in item questions and response options across diverse populations; involves self-reporting by patients; includes mandatory items for completion at scheduled reporting time points; includes an electronic reporting platform that allows for direct data entry into a database that can trigger rules-based notifications and expedited reporting; ensures acceptable measurement properties of items; is designed to be administered via the Web, handheld devices, paper, or telephone; is parsimonious with minimal patient or staff burden; and is designed to

integrate or automatically trigger longer symptom questionnaires or PROMIS tools. All of these design considerations are included in the proposed Tasks.

### *4.c. Integration With Related Projects and Infrastructure*

The investigators consider it essential that any developed PRO-CTCAE system be fully integrated with related projects and infrastructure.

- caAERS, MedDRA, CTCAE v4, caBIG Study Calendar, CDUS: The PRO-CTCAE platform will be developed in Task 5 to be fully interoperable with caAERS, CDUS, and the caBIG Study Calendar, fully integrated with CTCAE v4, and mapped to MedDRA. The vendor SemanticBits was selected for this subcontract because it currently holds the NCI contract to create caAERS, the platform through which CTCAE v4 is being developed for electronic administration and interoperability with MedDRA. Therefore, the PRO-CTCAE platform will be created as a stand-alone "plug-in" to caAERS.

- PROMIS Item Banks and Tools: The investigators recognize the importance of developing a technology platform that can also house non–PRO-CTCAE items and instruments, which can be required at designated visits (per the caBIG Study Calendar) or be triggered by PRO-CTCAE screening items. As such, the PRO-CTCAE platform is being designed to accommodate technologies necessary for administration of other items or questionnaires or linkage to other systems. Investigators on this proposal are familiar with PROMIS approaches and items and would look forward to collaborating with PROMIS investigators toward such systems integration. If awarded this contract, we will work with NCI staff to identify members of the PROMIS network/team to serve as advisors/liaisons to our investigative team to facilitate this collaborative effort (and vice versa).

- Other Symptom, HRQL, and Functional Status Instruments: The planned PRO-CTCAE platform will include the capacity for clinical investigators to load other symptom questionnaires besides PRO-CTCAE items into the system, which can be required at designated visits (per the caBIG Study Calendar) or triggered by PRO-CTCAE screening items.

### *4.d. Conceptual Framework*

Although this proposal focuses on the use of adverse symptom reporting for clinical trial data collection and regulatory documentation, patient self-reporting of this information is intended to also serve a broader purpose of providing clinicians with a more comprehensive picture of the patient experience when making management decisions, thus improving communication between patients and practitioners, improving patient satisfaction with care, enhancing the efficiency of clinical operations, and ultimately improving the quality of cancer care delivery and patient experience overall (Figure A.4-

2) (Velikova et al, J Clin Oncol 2004; Fowler et al, Urology 1993; Sloan et al, J Clin Epidemiol 2005). The inclusion of novel features such as between-visit reporting and automated alerts (rules-based notifications) provides an opportunity to modernize the current approach to adverse symptom event monitoring in cancer care by capitalizing on technology and following the inevitable trajectory of oncology practice orientation toward outpatient-centered paradigms that incorporate bidirectional electronic communication with patients.

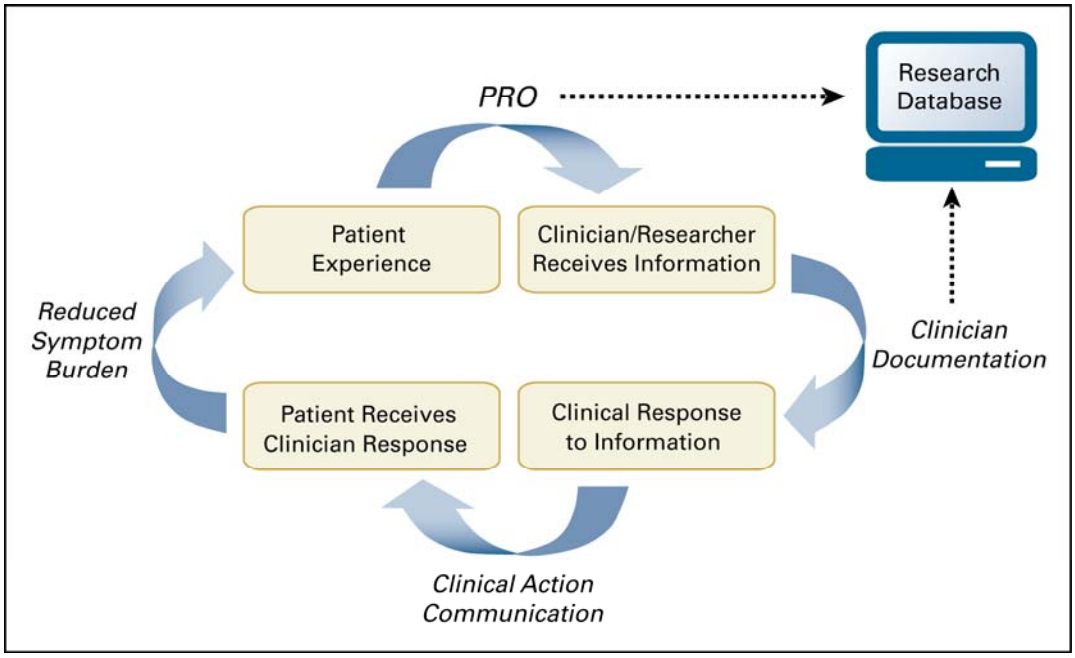



**Figure A.4-2. Conceptual Framework for the PRO-CTCAE in Clinical Oncology Research and Practice**

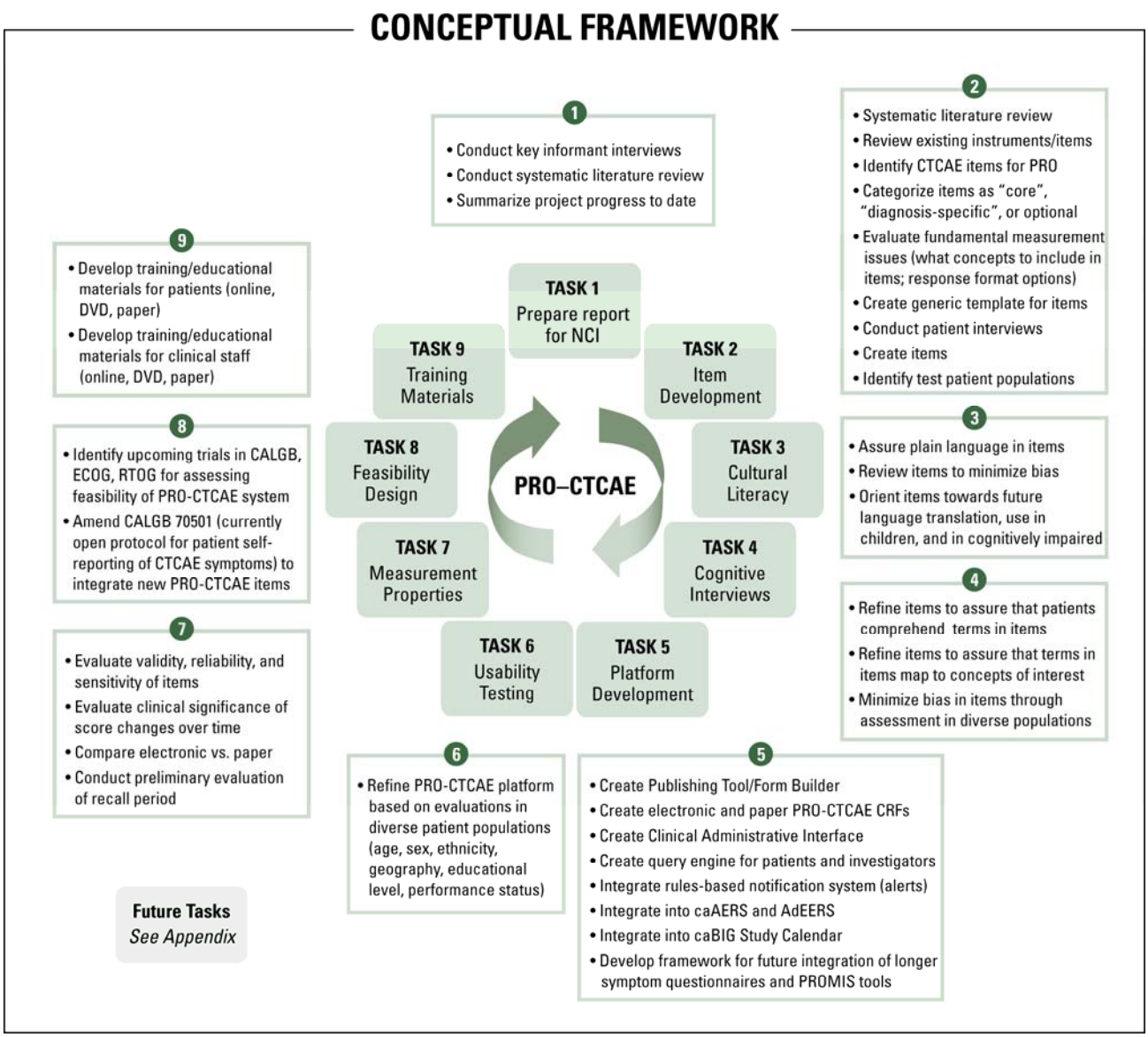
The investigators have designed a sequential series of interrelated Tasks to ensure that the developed PRO-CTCAE system will be immediately usable and useful in both the clinical trial and routine cancer care settings. The accomplished multidisciplinary team members assembled for this project bring complementary skills and perspectives that will ensure that both the scientific and humanistic dimensions of this project are adequately addressed at every step.

## 5. Substantive Approach and Methods

### Introduction

As described previously in Section A.1 of this application, "Objectives," we propose to undertake and accomplish a series of discrete but interrelated tasks, each with multiple steps and deliverables, in order to complete all work specified in the RFP (see Figure A.5-1A). Each project task is assigned a leader and deputy leaders, who may fill in for the leader if necessary to ensure timely completion of tasks and delivery of products to NCI staff on schedule.

**Figure A.5-1A. Proposed Tasks and Specific Steps for Developing and Evaluating the PRO-CTCAE Items and Assessment System**



### *5.a. Substantive Approach to Task 1: Prepare Background Reports*

| | |
|---|---|
| Task 1: | Produce report identifying ideal conditions, barriers, and strategies for implementing the PRO-CTCAE |
| Year of Contract: | Year 1 |
| Task Leader: | Dr. Deb Bruner |
| Backup Task Leader: | Dr. Deb Schrag |

**5.a.1. Overview**

Pursuant to the RFP, our first task will be to produce a report, *Ideal Conditions, Potential Barriers, and Strategies for Implementing the PRO-CTCAE.* The report will consider the theoretical, methodologic, and logistical issues, challenges, and benefits of developing and implementing the PRO-CTCAE system. This report will be based on literature searches and content prepared for this Technical Proposal, as well as key informant interviews.

This Task and coordination of its deliverables will be led by Dr. Deb Bruner at the University of Pennsylvania, with project management coordination and data management support centralized at Memorial Sloan-Kettering Cancer Center. For this Task, we have budgeted 10% of Dr. Bruner's time in year 1 and 5% in year 2, as well as 20 hours of dedicated contractor time from Dr. Andy Trotti at Moffitt Cancer Center (an expert in the history and development of the CTCAE and a prominent radiation oncologist in RTOG; see Letter of Support). In addition, key individuals representing stakeholder groups have agreed to participate in interviews for this Task at no cost (see Letters of Support and Table A.5-1).

Limitations related to the design and administration methods of current CTCAE symptom items, as well as the potential for PROs to provide solutions, have previously been documented (Bruner, J Clin Oncol 2007; Trotti et al, J Clin Oncol 2007; Basch et al, Lancet Oncol 2006) and are also described above. Interest in and need for valid, reliable, and sensitive PRO assessments in clinical trials in general are demonstrated by the increasing number of pharmaceutical, cooperative group, and other clinical trials that include PROs (Wiklund, Fundam Clin Pharmacol 2004; Bruner et al, Qual Life Res 2004; Lipscomb et al, J Clin Oncol 2007), the numerous meetings and publications devoted to PROs (Sloan et al, Value Health 2007; Sloan et al, J Clin Oncol 2007), and the interest in the use of PROs in drug labeling claims (Willke et al, Control Clin Trials 2004; FDA PRO guidance). Numerous organizations have shown interest in the development of consistent validated PRO measures, among them the International Society for Quality of Life Research (ISOQOL), the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), the Pharmaceutical Manufacturer's Association Health Outcomes Committee (PhRMA-HOC), and the European Regulatory Issues on Quality of Life Assessment (ERIQA) (Acquadro et al, Value Health 2003). The groundswell of this effort has contributed to the current RFP.

**5.a.2. Methods, specific steps, and deliverables**

The report will be based on literature searches and content prepared for this Technical Proposal, as well as key informant interviews. An outline of the report will be prepared by Drs. Bruner and Trotti, and delivered to NCI staff during month 1 of the contract period. Discrete sections in this report will include but will not be limited to topic areas specified in the RFP Statement of Work (section 1, items a through l) and in the RFP section titled "Additional Technical Proposal Instructions" (RFP pp. 61-62), as detailed in Table A.5-1.

During months 1-4, all sections of the report will be developed under Dr. Bruner's supervision and in collaboration with NCI program staff. As noted in the table, key stakeholders will be interviewed by Drs. Bruner, Schrag, and Basch at cooperative group meetings, and by telephone by the MSKCC Project Manager. Drs. Basch and Schrag have already arranged for such focus group sessions to occur at CALGB meetings through arrangements with chairs of committees, which consist of clinical investigators, data managers, and patient advocates, with support and endorsement from Dr. Rich Schilsky (CALGB chair) (see attached Letter of Support).

**Table A.5-1. Key Informants and Estimated Number of Interviews**

| KEY STAKEHOLDER INFORMANTS | Number of Planned Interviewees |
|---|---|
| **Patients advocates** | 3-5 |
| **Disease-specific cooperative group leaders** | 5-10 |
| **Clinical investigators (academic)** | 10-15 |
| **Industry/biopharmaceutical sponsor representatives** | 2-5 |
| **Adult oncologists (CCOPs and tertiary centers)** | 5-15 |
| **Pediatric oncologists/researchers** | 2-5 |
| **Radiation oncologists** | 3-5 |
| **Surgical Oncologists** | 2-5 |
| **Research nurses** | 5-10 |
| **Data managers/clinical research assistants** | 2-3 |

| | |
|---|---|
| **PRO instrument developers/psychometricians** | 2-5 |
| **Health/cultural literacy experts** | 2-3 |
| **Cognitive interviewing experts** | 1-2 |
| **ePRO technology vendors** | 2-3 |
| **CDISC representative** | 1 |
| **NCI staff (Outcomes Research Branch)** | TBD |
| **NCI staff (DCP)** | TBD |
| **NCI staff (CTEP)** | TBD |
| **NCI staff (caBIG/NCICB)** | TBD |
| **FDA (SEALD)** | 2 |
| **FDA (Oncology Division)** | 2-3 |
| **Representatives of other federal agencies** | TBD |

*The CALGB Research Communications Subcommittee, of which Dr. Basch is a member, predominantly consists of patient advocate members; Deb Collyar, chair of this committee, has agreed to facilitate these interviews and group discussions, including patient advocate members of this committee (see attached Letters of Support).

For each section of the report, information will be gathered from appropriate sources and stakeholders during their interviews (Table A.5-2). Hence, interviews for each stakeholder type will include modules pertaining to the particular sections of interest, as delineated below.

**Table A.5-2. Selected Sections of Report and Applicable Interviewees**

| REPORT SECTION | APPLICABLE INTERVIEWEES |
|---|---|
| Review current CTCAE use in NCI-sponsored trials | Clinical investigators, data managers, research nurses, NCI staff |
| Review adverse event reporting systems used by other federal agencies | Representatives of other federal agencies |
| Identification of multiple approaches for integrating PRO-CTCAE into CTCAE (current and future versions) | Patient advocates, disease-specific cooperative group leaders, clinical investigators, industry |

| | |
|---|---|
| and ideal approach for multiple types of symptoms | representatives, oncologists, nurses, data managers, NCI staff, instrument developers |
| Create blueprint for how the open-source PRO-CTCAE platform will be interoperable with caAERS, AdEERS, CDUS, caBIG Study Calendar, and other applicable software platforms and standards | SemanticBits (software developer subcontractor), other ePRO vendors, caBIG/NCICB representatives, CDISC representative, CTEP representatives |
| Determine which symptoms of CTCAE may be enhanced by PRO | PRO instrument developers, clinical investigators, oncologists, nurses, patient advocates, NCI staff, FDA staff |
| Identify barriers for implementing the PRO-CTCAE in NCI-sponsored trials | Cooperative group leadership, clinical investigators, data managers, community oncologists and nurses (at CCOPs), patient advocates, NCI staff |
| Determine ideal measurement properties of PRO-CTCAE (validity, reliability, sensitivity, recall period, questionnaire burden, etc). | Instrument developers/psychometricians, NCI staff, FDA SEALD staff |
| Identify existing PRO measures that may serve as sources of items/terms for PRO-CTCAE | Instrument developers/psychometricians, clinical investigators, pharmaceutical representatives, ePRO vendors, NCI staff, FDA SEALD staff |
| Determine minimal criteria to collect patient self-reports (age, educational level, cognitive ability) | Health/cultural literacy experts, cognitive interviewing experts, instrument developers/psychometricians, pediatric oncology researchers, patient advocates, NCI staff, FDA staff |

Investigators on this proposal also have expertise specific to all of these areas.

An *initial draft* of this report will be delivered to NCI staff during month 6 of this contract. Revisions will be made based on NCI staff feedback and lessons learned during the other year 1 Tasks of this proposal, and input from the investigative team, and a *revised draft* will be delivered to NCI staff during month 12. We will continue to update this report throughout the contract, and a *final version* will be delivered to NCI staff during month 24.

The collected data will also be used during year 1 as background data to prepare our study plan titled ***Detailed Design of Validation/Feasibility Studies Plan*** (which will be carried out in Tasks 7 and 8). The study plan will be delivered to NCI staff in draft form for review during month 13 and in final form during month 14. The collected data will also be used to inform a document titled ***Plan for Best Practices for Application of PRO-CTCAE in Community and Comprehensive Cancer Care Settings,*** which will be delivered to NCI staff during month 12 and used to guide our work in Tasks 7, 8, and 9 of this proposal.

The schedule of deliverables to NCI staff is detailed in Table A.5-3, and timeline in Tables A.7-1 and A.7-2.

**5.a.3. Analytic approach**

The principal investigator (Dr. Basch) and the investigative team will review and critique each draft before the final report. Subject to NCI review and approval, the report will be submitted for peer-review publication and disseminated to the key stakeholders through formal presentation at cooperative group meetings and through electronic communication.

In general, information elicited from literature searches and interviews will be used to develop frameworks for the planned Tasks in three classification categories: theoretical, methodologic, and logistical. For example, theoretical considerations regarding item design will include how the PRO-CTCAE may advance measurement of the patient experience with treatment and relate to other measures of efficacy and safety. Conceptual frameworks will assist the investigators in defining variables of interest and positing causal relationships (Ferrans, *Outcomes Assessment in Cancer* 2005; Rothman et al, Value Health 2007). Although the investigative team is very familiar with methods for item development and evaluation, we will nonetheless turn to the scientific literature and stakeholder interviews to remain open to alternative approaches to those described in this proposal. The input of NCI staff will also be highly valued in this regard. Specific design elements of specific Tasks in this proposal will be informed by the gathered data, for example, refining the list of CTCAE items amenable to patient self-reporting in Task 2 or design specifications for the technology platform in Task 5.

Logistical considerations will primarily be informed by stakeholder interviews and focus groups. At this level, key stakeholders, including regulatory (FDA, EMEA), NCI (CTEP, DCP, NCICB, Outcomes Research Branch) and cooperative group leadership, clinical trialists, pharmaceutical industry representatives, community clinician representatives, and patient advocates will address the contexts in which PRO-CTCAE items might be most informative and consider what the anticipated barriers and strategies to implementation of the PRO-CTCAE system will be. Issues of administration, such as electronic vs. paper or telephone, timing, frequency, language translations, and feasibility among patients and staff, will specifically be addressed to determine optimal content of the planned training manuals in Task 9. Patient acceptance/adherence issues will be reviewed with patient advocates and cooperative group and community clinical trialists and data managers/clinical research associates. Discussions with participating cooperative groups, particularly CCOP leadership and DCP, will inform Task 8, the development of a protocol to test feasibility in a cooperative group trial.

### 5.a.4. Specific deliverables and timeline

The schedule of deliverables to NCI staff is detailed in Table A.5-3.

**Table A.5-3. Specific Deliverables and Timeline for Task 1**

| SPECIFIC DELIVERABLES | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| ***REPORT: Ideal Conditions, Potential Barriers, and Strategies to Implement the PRO-CTCAE*** | |
| ▪ Outline of Report | Month 1 |
| ▪ Initial Draft of Report | Month 6 |
| ▪ Revised Draft of Report | Month 12 |
| ▪ Final Version of Report | Month 24 |
| ***DOCUMENT: Detailed Design of Validation and Feasibility Studies Plans*** | |
| ▪ Draft Documents | Month 13 |
| ▪ Final Documents | Month 14 |
| ***DOCUMENT: Plan for Best Practices for Application of PRO-CTCAE in Community and Comprehensive Cancer Care Settings*** | |
| ▪ Final Document | Month 12 |

### *5.b. Substantive Approach to Task 2: Develop PRO-CTCAE Item*

| | |
|---|---|
| Task 2: | Identify CTCAE items that are amenable to patient self-reporting, determine which items require additional clinician judgment after patient self-report, determine what concepts are appropriate for measurement in PRO-CTCAE items, and create draft items based on analysis of existing instruments, systematic literature review, and key stakeholder feedback |
| Year of Contract: | Year 1 |
| Task Leader: | Dr. Charlie Cleeland |
| Deputy Task Leaders: | Dr. Jeff Sloan, Dr. Tito Mendoza |

**5.b.1. Overview**

A particular challenge we will face when creating the PRO-CTCAE items is that although we would ideally wish to start from scratch and design items with formats and measurement properties that are optimal for administration in patients with cancer, at the same time we may be limited by structural constraints of the clinician CTCAE, with which PRO-CTCAE items must ultimately remain compatible. Therefore, we believe it is an essential initial step in development of the PRO-CTCAE to analyze the structure of current CTCAE v3 "symptom items" (ie, items with subjective content that may be amenable to patient self-reporting) to identify elements that may present problems in the context of PRO. In particular, double-barreling of concepts in response options and the limited range of the CTCAE ordinal response scale may be limitations for patient self-reporting.

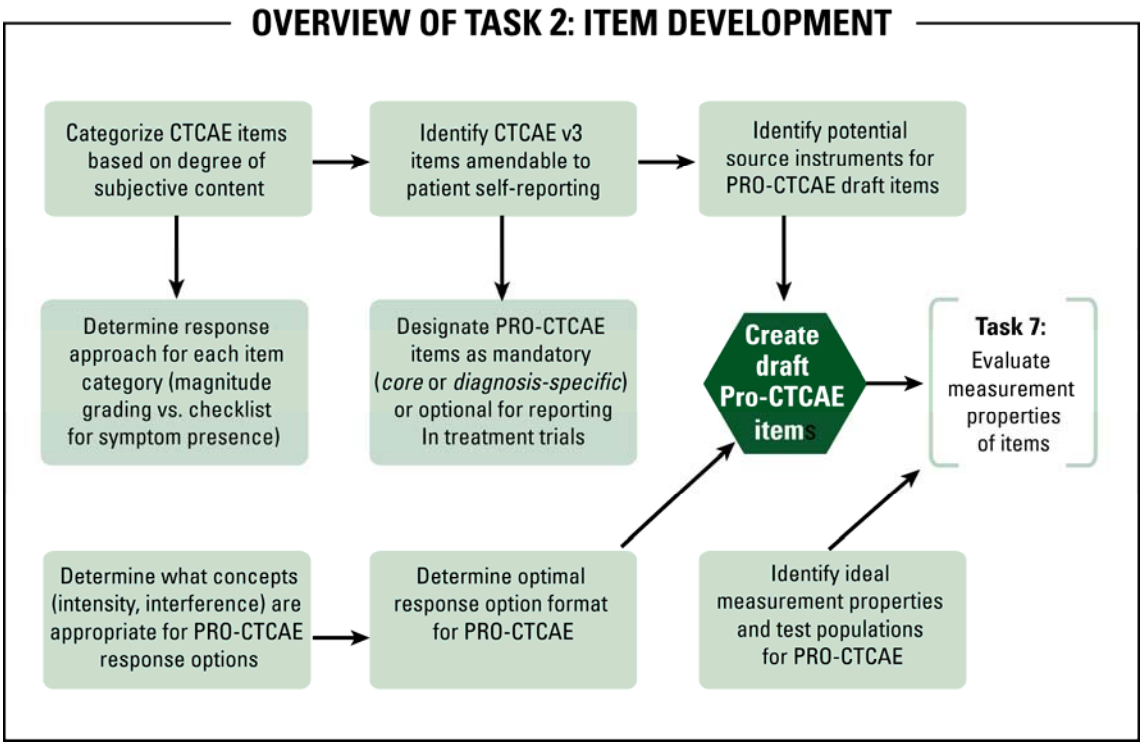



**Figure A.5-1. Overview of Specific Steps Toward Development of PRO-CTCAE Draft Items**

In this section of the proposal, we walk through a sequence of steps to identify these potential problem issues, review existing literature in each applicable area, and describe why we feel it is scientifically justified to diverge from the clinician CTCAE format for PRO-CTCAE items in some cases (Figure A.5-1). As described below, we propose that that fundamental questions regarding which concepts should be included in PRO-CTCAE items (intensity +/- interference) and response option formats (0-10 vs. 0-4) should be addressed and decided through a consensus process early in the award period, involving the investigators, NCI staff, and selected advisors, based on best available evidence and expert opinion. However, we also include optional research approaches to addressing these questions in Appendix 2 (not included in the current budget, but available for consideration of future study via alternate funding mechanisms). We feel that these "sub-studies" will not be necessary if these questions can be resolved through consensus; this process will be a topic of immediate discussion with NCI upon award of the contract.

In this section, we also describe our rationale for categorizing CTCAE items according to their relative degree of subjective content, explain our process for identifying CTCAE items amenable to PRO, and illustrate our designations of whether particular items should be considered as mandatory across all cancer treatment trials (core), mandatory in selected trials (diagnosis specific), or optional. Finally, we describe our proposed process for creating the PRO-CTCAE items themselves, once we have explored the above areas.

The investigators have several unique advantages in conducting this work. Drs. Basch and Schrag and the team at MSKCC have extensive experience developing and evaluating prior draft patient versions of CTCAE items (MSKCC STAR Items), which mirror the format and concepts of existing CTCAE v3 items. Therefore, they are aware of the potential difficulties associated with this approach, which have informed the methods described in this section. Drs. Cleeland and Mendoza and their group at MDACC have extensive experience with instrument development and therefore also have access to large data sets of patient questionnaires from this work, which we can analyze toward creation of PRO-CTCAE items. The investigators bring access to large and diverse patient populations in our associated institutions. The investigative team also brings multidisciplinary perspectives, including clinical, statistical, psychometric, informatics, and health services, and we have included key advisers from relevant stakeholder groups, particularly patient advocates.

**5.b.2. Identification of existing CTCAE items amenable to patient self-reporting**

Our first step in the development of PRO-CTCAE items will be to review the current CTCAE and identify which items are amenable to patient self-reporting (and thus should be included in the PRO-CTCAE).

Dr. Basch's team at MSKCC has previously developed a method for selecting and categorizing items in the CTCAE to determine if they are amenable to PRO (described below) and has conducted a preliminary analysis using this method to compile an initial list of candidate items. This list, which is shown in Table A.5-4, will be refined by the investigators and presented to NCI staff during <u>month 1</u> of the contract period. This item list will then be further refined through consultation with our key advisers and NCI staff to yield a final item list for version 1 of the PRO-

CTCAE, which will be delivered to NCI staff in month 2 of the contract period. The process used for item selection and categorization is shown in Figure A.5-2, and each aspect of this process is explained in the following subsections.

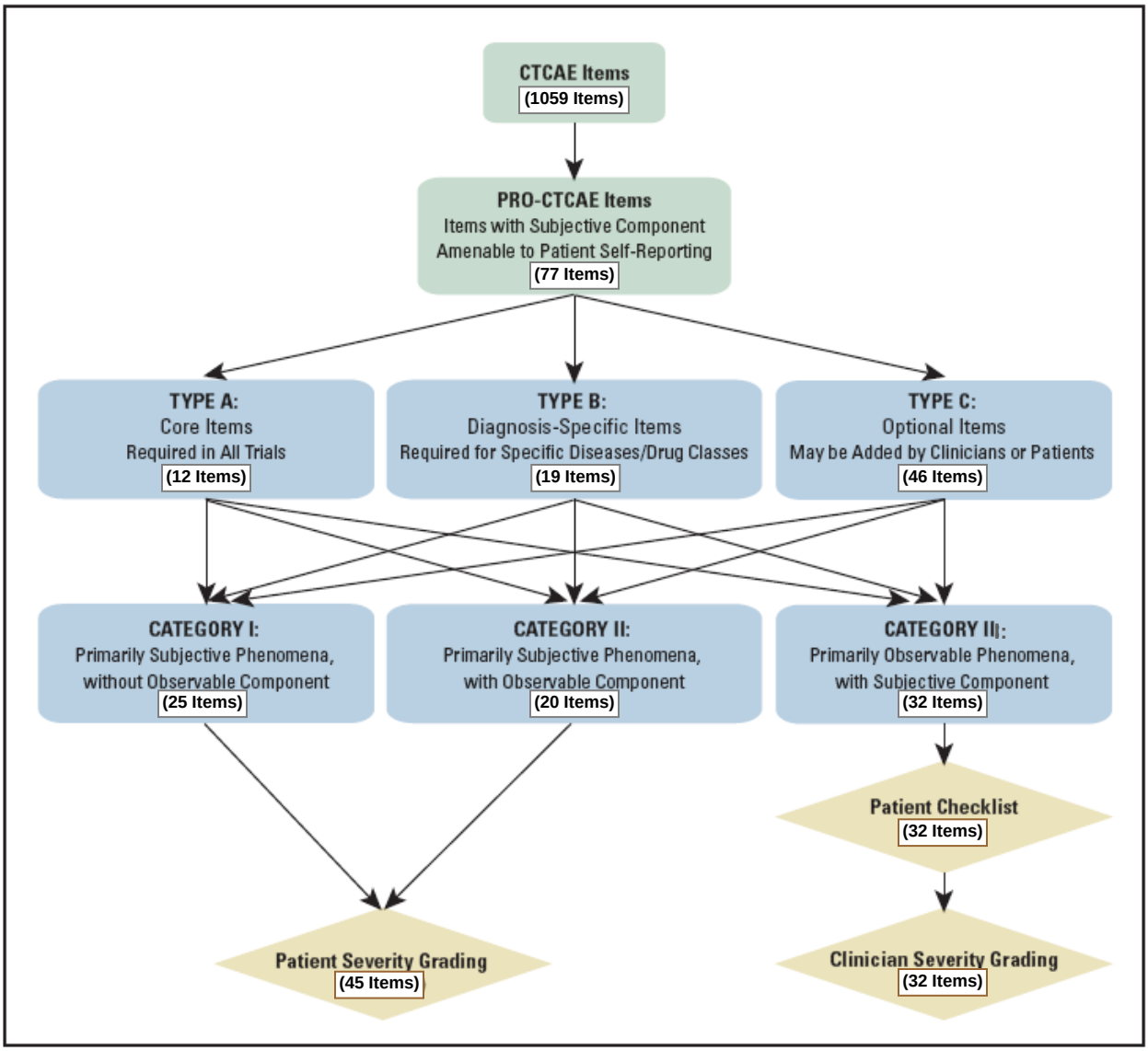



**Figure A.5-2. Approach to Selecting CTCAE Items for PRO-CTCAE**

### 5.b.3. Categorization of CTCAE items based on degree of subjective content

To generate a list of candidate items, the 1059 items in CTCAE v3 can be divided into four mutually exclusive categories based on degree of subjective (sensory) vs. objective (observable) content (Table A.5-4).

**Table A.5-4 Categories of CTCAE v3 Items Based on Degree of Subjective Component**

| CATEGORY | CHARACTERISTICS | EXAMPLES |
|---|---|---|
| **I** | Primarily subjective phenomena, without observable component | Nausea, fatigue |
| **II** | Primarily subjective phenomena, with observable component | Vomiting, diarrhea |
| **III** | Primarily observable phenomena, with subjective component | Nail changes, peripheral edema |
| **IV** | Laboratory/measurable | Hemoglobin, prolonged QTc |

For example, to prepare the preliminary list shown in Table A.5-4, this categorization process was conducted independently by two medical oncology clinical investigators, and results were compared and resolved through discussion and review of the final list by members of the investigative team. Results were also compared to a list generated during an informal meeting at NCI in March 2007, during which Dr. Basch and representatives of CTEP, DCP, and the Health Outcomes Branch reviewed each CTCAE item and designated each as potentially amendable to PRO based on consensus. Our categorization process yielded a list of 77 CTCAE items in categories I, II, or III (all believed to be amenable to patient self-reporting in the PRO-CTCAE).

However, the investigators recognize that the current list is *preliminary and may not exhaustive,* and the criteria used to delineate among items (as well as the categorization of controversial items) will likely require adjustment through dialog with NCI staff and key advisers during the early contract period.

A particular issue meriting discussion with NCI staff is that in the current CTCAE many items we classify as category IV actually require some evaluation of a patient's subjective experience to yield a grade. For example, evaluation of left ventricular systolic dysfunction in CTCAE v3 requires determination of whether a patient is symptomatic or asymptomatic. However, because detection of the phenomenon is via a measurement test, and the subjective component is a sequelae of the measured phenomenon rather than the measured phenomenon itself, we designate this item as category IV and therefore not ideal for PRO in the initial PRO-CTCAE. In the future, once the PRO-CTCAE has been established and there is an opportunity for bidirectional collaborative grading by clinicians and patients, such items could be included in a reporting model in which clinicians recognize the presence of a phenomenon via a measurement test, and subsequently patients are asked to report their subjective experience related to that phenomenon.

As shown in Figure A.5-2 and Table A.5-5, using these criteria our preliminary estimate is that there are 25 items in category I, 20 in category II, and 32 in category III. As noted above, the criteria and list will be refined by the investigators and presented to NCI staff during month 1 of the contract period. This item list will then be further refined through consultation with our key advisers and NCI staff to yield a final item list for version 1 of the PRO-CTCAE, which will be delivered to NCI staff in month 2 of the contract period.

**5.b.4. Determination of patient response approach based on degree of subjective content in each PRO-CTCAE item: Severity grading vs. checklist for presence/absence**

Based on development work with the identified items at MSKCC (including patient feedback/debriefing) and consensus among members of Dr. Basch's ongoing Working Group for the Development of a PRO Version of the CTCAE, the investigators have concluded that items in categories I, II, and III are amenable to patient reporting of the presence or absence of the symptoms/phenomena, but only categories I and II are amenable to patient *severity* grading. In contrast, items in category III, due to their substantial objective (observable) component, require clinical training and experience to grade along the full spectrum of severity. For example, during cognitive debriefing we have found that patients with drug rash or hand/foot syndrome are generally not aware of the existing full spectrum of severity, whereas clinicians are. Therefore, the investigators conclude that items that are primarily observable yet have a subjective component of the phenomenon itself (or can be initially observed by patients) (ie, category III) should be included in the PRO-CTCAE as binary items (yes/no) to detect the presence/absence of a phenomenon (ie, via a checklist). However, severity rating should be conducted by clinical staff. Therefore, the investigators plan to include these items in the PRO-CTCAE system toward collaborative reporting by patients and clinicians. In this model, patients will be able to screen themselves for the presence or absence of these symptoms via the PRO-CTCAE system, and then a clinician will be required to assign a grade for that item. The PRO-CTCAE technology platform is planned to accommodate this design element, as described in the approach to Task 5 (platform development). As shown in Figure A.5-2 and Table A.5-5, in our preliminary list of items there are 45 items intended for patient grading and 32 items for checklist format (which when positive will trigger mandatory clinician grading). This approach will be discussed with NCI and FDA staff immediately after award of the contract to gain consensus on logistics of the approach.

**5.b.5. Designation of PRO-CTCAE items as mandatory or optional for reporting in trials**

Parsimony and minimization of patient burden are key to developing a widely accepted PRO-CTCAE. Therefore, all items cannot be asked at every visit. The investigators assert that a core set of common and clinically important symptoms should be required for administration in the PRO-CTCAE at every visit in every trial regardless of diagnosis. Thus, a core item set will be defined as a part of this proposal.

A preliminary list of core items has been selected based on consultations with oncology disease experts at MSKCC and patient advocates, literature review, and consensus among the investigators (shown in Table A.5-5). This list is included for the purpose of demonstrating the approach, and the final list will be based on consensus between the investigative team, selected advisors, and NCI staff early in the contract period. This list includes 12 core items, 19 diagnosis-specific items, and 46 optional items. (Notably, all of the identified core and disease-specific symptoms and approximately one-third of optional symptoms have previously been included in evaluations of the MDASI, and therefore these MDACC data sets will be highly useful in the planned development work for the PRO-CTCAE.)

In this model, for any given trial, as described below in the "Substantive Approach to Task 5," when a PRO-CTCAE CRF is created, the core items would always be mandatory. Then,

additional diagnosis-specific items would be mandatory, depending on the disease of participating patients. The clinical investigator building the form could then chose additional optional items from the menu of optional PRO-CTCAE items, depending on the characteristics of the study regimen or patient population. Subsequently, during the conduct of the study, patients completing PRO-CTCAE CRFs electronically at scheduled visits would have an opportunity to select additional optional items for grading from a checklist or dropdown menu. These items would then be automatically added to that patient's CRFs at following visits to track the course of that symptom longitudinally.

Two optional evaluations, not included in the proposed work or budget, could be added into the proposal/budget if NCI staff feels they are warranted prior to designating PRO-CTCAE core items: 1) Evaluation of prevalence of symptoms, correlations/redundancy measured among symptom items, and relative severity levels in different cancer populations based on MDACC initial patient interviews and the validation set for the MDASI in which 2093 patients with diverse cancer diagnoses were administered 52 distinct symptom items (encompassing all of the above "core" items and most of the "diagnosis-specific" items), of which longitudinal assessments were collected from 638 patients receiving active cancer treatment (Cleeland et al, Cancer 2000; Cleeland et al, N Eng Jrl Med 1994b; Mendoza et al, Cancer 1999; Armstrong and Mendoza, J Neurooncol 2006; Rosenthal and Mendoza, Head and Neck 2007); and 2) evaluation of prevalence of symptoms according to CTCAE reporting based on all CDUS and AdEERS reports since 2004 (analysis could be done in collaboration with CTEP using anonymous data).

**Table A.5-5. Preliminary List of CTCAE Items for PRO Adaptation and Item Categories and Types, Based on Meetings from 2006 to the Present of Dr. Basch's "Working Group for the Development of a PRO Version of the CTCAE," Investigator Consensus, and Preliminary Literature Review**

| NO. | CURRENT CTCAE SUPERORDINATE TERM | ITEM CATEGORY<br>**I:** Primarily subjective phenomena, without observable component<br>**II:** Primarily subjective phenomena, with observable component<br>**III:** Primarily observable phenomena, with subjective component | ITEM TYPE<br>A: Core items<br>B: Diagnosis-specific items<br>C: Optional items | PLANNED STUDY POPULATIONS<br>a: Adjuvant breast<br>b: Metastatic prostate/bladder<br>c: H/N, gastro-esophageal (RT)<br>d: Metastatic lung<br>e: Metastatic colon<br>f: Lymphoma/myeloma |
|---|---|---|---|---|
| 1 | **Alopecia/hair loss (scalp or body)** | III | C | a,d |
| 2 | **Anorexia/appetite disturbance** | I | A | b,d,e |
| 3 | **Arthralgia/joint pain** | I | C | b,f |
| 4 | **Ataxia (incoordination)** | III | C | c,d |
| 5 | **Bronchospasm/wheezing** | III | B | d |
| 6 | **Bruising** | III | C | f |
| 7 | **Cheilitis** | III | C | c |
| 8 | **Constipation** | II | A | b,c,e |
| 9 | **Cough** | II | B | d |
| 10 | **Dehydration** | III | B | c,e |
| 11 | **Diarrhea** | II | A | a,b,c,d,e,f |
| 12 | **Distension/bloating, abdominal** | III | C | e,f |
| 13 | **Dizziness/lightheadedness** | I | B | c,d |
| 14 | **Dry mouth/salivary gland (xerostomia)** | I | B | c |
| 15 | **Dry skin** | II | C | c,d,f |
| 16 | **Dysphagia/difficulty swallowing** | I | B | c,e |
| 17 | **Dyspareunia** | I | C | a,e |
| 18 | **Dyspnea** | I | A | d |
| 19 | **Dysuria** | I | C | b |
| 20 | **Edema (peripheral)/leg swelling/fluid retention** | III | C | b,e,f |
| 21 | **Ejaculatory dysfunction** | II | C | b |
| 22 | **Epistaxis (nosebleed)** | III | C | c |
| 23 | **Erectile dysfunction** | II | C | b |
| 24 | **Fatigue*** | I | A | a,b,c,d,e,f |
| 25 | **Flatulence** | II | C | a,e |
| 26 | **Flushing** | II | B | a,b |
| 27 | **Gynecomastia** | III | C | b |
| 28 | **Heartburn/dyspepsia** | I | B | a,b,c,d,e,f |
| 29 | **Hot flashes** | I | B | a,b |
| 30 | **Hiccoughs** | III | C | c,d |
| 31 | **Hyperpigmentation** | III | C | c,f |
| 32 | **Incontinence, anal** | III | C | b,e |
| 33 | **Incontinence, urinary** | III | C | b |
| 34 | **Induration/fibrosis (skin and subcutaneous tissue)** | III | C | c,f |

| NO. | CURRENT CTCAE SUPERORDINATE TERM | ITEM CATEGORY<br>**I:** Primarily subjective phenomena, without observable component<br>**II:** Primarily subjective phenomena, with observable component<br>**III:** Primarily observable phenomena, with subjective component | ITEM TYPE<br>A: Core items<br>B: Diagnosis-specific items<br>C: Optional items | PLANNED STUDY POPULATIONS<br>a: Adjuvant breast<br>b: Metastatic prostate/bladder<br>c: H/N, gastro-esophageal (RT)<br>d: Metastatic lung<br>e: Metastatic colon<br>f: Lymphoma/myeloma |
|---|---|---|---|---|
| 35 | **Injection site reaction/extravasation changes** | III | C | a,b,c,d,e,f |
| 36 | **Insomnia/sleep disturbance** | I | A | a,b,c,d,e,f |
| 37 | **Irregular menses (change from baseline)** | II | C | a,d,e,f |
| 38 | **Libido** | I | B | a,b |
| 39 | **Mood: anxiety** | I | A | a,b,c,d,e,f |
| 40 | **Mood: depression** | I | A | a,b,c,d,e,f |
| 41 | **Mucositis/stomatitis (functional/symptomatic)** | III | B | c,d,e |
| 42 | **Myalgia (muscle pain)** | I | C | d,f |
| 43 | **Nail changes** | III | C | b,d |
| 44 | **Nausea** | I | A | a,b,c,d,e,f |
| 45 | **Neuropathy (sensory)** | I | A | b,d,e |
| 46 | **Odor (patient odor)** | III | C | c,e |
| 47 | **Orgasmic dysfunction** | II | C | B |
| 48 | **Pain** | I | A | b,c,d,e,f |
| 49 | **Pain, abdominal** | I | B | E |
| 50 | **Palpitations** | II | C | D |
| 51 | **Photosensitivity** | II | C | E |
| 52 | **Prolapse of stoma, GI** | III | C | E |
| 53 | **Pruritus/itching** | I | B | c,f |
| 54 | **Rash/desquamation** | III | C | c,f |
| 55 | **Rash: acne/acneiform** | III | B | A |
| 56 | **Rash: Dermatitis associated with radiation** | III | B | a,c |
| 57 | **Rash: hand-foot skin reaction** | III | B | c,e |
| 58 | **Rigors/chills** | III | C | d,f |
| 59 | **Skin breakdown/decubitus ulcer** | III | C | b,e |
| 60 | **Speech impairment (eg, aphasia)** | II | C | c,d |
| 61 | **Striae** | III | C | E |
| 62 | **Sweating (diaphoresis)** | II | B | b,d,e,f |
| 63 | **Taste disturbance** | I | C | c,d |
| 64 | **Telangiectasia** | III | C | C |
| 65 | **Tinnitus** | I | C | c,d |
| 66 | **Tremor** | III | C | c,d |
| 67 | **Urinary frequency/urgency** | II | B | B |
| 68 | **Urine color change** | III | C | b,e |

| NO. | CURRENT CTCAE SUPERORDINATE TERM | ITEM CATEGORY<br>**I:** Primarily subjective phenomena, without observable component<br>**II:** Primarily subjective phenomena, with observable component<br>**III:** Primarily observable phenomena, with subjective component | ITEM TYPE<br>A: Core items<br>B: Diagnosis-specific items<br>C: Optional items | PLANNED STUDY POPULATIONS<br>a: Adjuvant breast<br>b: Metastatic prostate/bladder<br>c: H/N, gastro-esophageal (RT)<br>d: Metastatic lung<br>e: Metastatic colon<br>f: Lymphoma/myeloma |
|---|---|---|---|---|
| 69 | **Urticaria (hives, welts, wheals)** | III | C | c |
| 70 | **Vaginal discharge (noninfectious)** | II | C | A |
| 71 | **Vaginal dryness** | I | B | a,e |
| 72 | **Vision/ocular: blurred vision** | II | C | C |
| 73 | **Vision/ocular: flashing lights/floaters** | II | C | C |
| 74 | **Voice changes/dysarthria (eg, hoarseness, loss or alteration in voice, laryngitis)** | III | C | c,d |
| 75 | **Vomiting** | II | A | c,e |
| 76 | **Watery eye (epiphora, tearing)** | II | C | C |
| 77 | **Weakness** | I | C | b,c,d,e,f |

**PRELIMINARY ESTIMATES:**
**TOTAL CTCAE ITEMS AMENABLE TO PRO: 77**
**TOTAL AMENABLE TO PATIENT GRADING AS PURE PRO (CATEGORIES I or II): 45**
**TOTAL AMENABLE TO PATIENT CHECKLIST Y/N → CLINICIAN REVIEW (CATEGORY III): 32**
**TOTAL "CORE" ITEMS (TYPE A): 12**
**TOTAL "DIAGNOSIS-SPECIFIC" ITEMS (TYPE B): 19**
**TOTAL "OPTIONAL" ITEMS (TYPE C): 46**

**Note: Asthenia May Be Split from Fatigue in CTCAEv4.*

### 5.b.6. Determination of what concepts are most appropriate for measurement in PRO-CTCAE response options (eg, intensity, interference, frequency)

*Overview.* Current CTCAE v3 symptom items include double-barreling of concepts, asking respondents to consider either both intensity and interference (for most category I items) or both intensity and frequency (for most category II items). However, when these items are adapted for the purpose of patient self-reporting, multiple methodologic disadvantages emerge that are related to this approach, as Dr. Basch and his team have learned during their experience developing and evaluating STAR Items (particularly, patients may not be able to fully consider multiple concepts when responding to these items). Therefore, a fundamental question that must be addressed early in the proposed contract period is whether double-barreling of concepts in response options should be included in the PRO-CTCAE or if this approach, which is prevalent in CTCAE v3, should be abandoned in favor of a different response option format. We propose that this question be decided through a consensus meeting early in the award period, including investigators, NCI staff, and selected advisors, based on best available evidence and expert opinion. However, we also include an optional research approach to addressing this question in Appendix 2, which is not included in the current budget. This research could be considered in the future via an alternate funding mechanism.

The investigators concur that "severity" should be assessed for all category I and II items. Based on existing evidence, we believe that there is only limited additional information provided by interference or frequency items, such that administration of such items in addition to intensity does not merit the additional administrative burden in the setting of adverse event screening. We propose that this question be decided through a consensus meeting early in the award period, including investigators, NCI staff, and selected advisors, based on best available evidence and expert opinion. However, we also include an optional prospective study and an optional retrospective data mining study to addressing this question in Appendix 2, which is not included in the current budget. This research could be considered in the future via an alternate funding mechanism.

(A note on *"Severity" vs. "Intensity:"* Notably, an emerging issue in the symptom research community is whether there is a distinction between "intensity" and "severity." Traditionally, these terms have been considered to represent the same concept (ie, subjective/sensory proprioception). More recently, it has been considered whether these are different – ie, "intensity" may represent subjective/sensory proprioception, while "severity" represents subjective perception of magnitude. For example, a patient could experience "severe interference" or "severe frequency" (or "severe intensity"). If we take the latter perspective, then PRO-CTCAE questions which ask patients "how severe" a particular symptom is may actually capture both "intensity" and "interference" experiences (as well as other concepts, such as "bother," "frequency," and "duration"). Likely, the determination of whether we are measuring "intensity" or "severity" when we ask a patient "how severe is your…" is an academic issue when it comes to creating PRO-CTCAE items, due to the necessary brevity of these items: it would be challenging if not impossible to create items which distinguish "intensity" from "severity" with limited verbiage, and doing so would likely yield complex items not amenable to translation, use in children, and free of cultural bias. In general, in this application, when referring to the results of symptom research, we will use the term "intensity." However, in our discussion of creating PRO-CTCAE items which ask patients to globally rate the magnitude of an experience without specific reference to a concept (such as "interference" or level of bother), we will use the term "severity." Unfortunately, due to ambiguity regarding the distinction between these terms at this time, in cases when PRO-CTCAE development is being discussed in the context of prior research or publications, we will use both terms or include one of the terms in parenthesis following the other. We intend to clarify this issue of nomenclature in discussions with NCI staff and key advisors at the time of contract commencement, towards clarity during the item development process. The issue we address in this section is different, however. We are interested to determine whether it is necessary to specifically ask patients about their "interference" experience in PRO-CTCAE items, or whether we can gather enough information from a single question which does not specifically ask about interference (regardless of whether we believe the alternative is asking about "intensity" or "severity" and whether or not these two terms represent the same thing).

*Patient understanding of double-barreled response options.* As noted above in section 4 (Rationale for Approach), the response options in all current CTCAE symptom items include double-barreling of concepts, either asking respondents to consider both intensity and

interference (for most category I items) or asking respondents to consider both intensity and frequency (for most category II items). In previous development work, Dr. Basch's team at MSKCC created and evaluated initial pilot patient versions of CTCAE symptom items (STAR Items), which mirrored the double-barreling of concepts in the parent CTCAE items. This method was adopted at the time because it was believed that this approach was most likely to be accepted by clinical investigators and regulators (because the PRO approach to adverse event reporting in oncology was then novel) (Basch et al, J Clin Oncol, 2005; Basch et al, J Am Med Inform Assoc 2007; Basch et al, Lancet Oncol, 2006). The generic structure of the response options for these items is shown in Table A.5-6, with category I items such as neuropathy generally combining intensity with interference, and category II items such as diarrhea combining intensity with frequency (eg, counting the number of stools per day).

**Table A.5-6. Generic Structure of the MSKCC STAR PRO-CTCAE Symptom Item Response Options***

| GRADE | CATEGORY I RESPONSE CONCEPTS | CATEGORY II RESPONSE CONCEPTS |
|---|---|---|
| **0** | Not present | Not present |
| **1** | Mild intensity AND no interference with applicable function/ADL/enjoyment of life | Occasional (optional to quantify) AND no interference with applicable function/ADL/enjoyment of life |
| **2** | Moderate intensity AND/OR some (a little) interference with applicable function/ADL/enjoyment of life | Frequent/persistent (optional to quantify) AND/OR some (a little) interference with applicable function/ADL/enjoyment of life |
| **3** | Severe intensity AND/OR much (a lot) interference with applicable function/ADL/enjoyment of life | Frequent/persistent (optional to quantify; may be >grade 2 magnitude) AND/OR much (a lot) interference with applicable function/ADL/enjoyment of life |
| **4** | Excruciating intensity AND/OR attributable disability | NA (life-threatening in CTCAE) |

*Concepts included in specific individual items in the parent CTCAE items in each category can vary from this structure, which was accounted for in the MSKCC STAR CTCAE patient adaptations.

Although mirroring the double-barreling of parent CTCAE items was an appropriate approach for initial proof-of-concept development work, for an operationalized PRO-CTCAE, the investigators are inclined against the use of double-barreled concepts.

As described in Table A.5-7, our rationale is based on the following: (1) health literacy and cognitive interviewing experts and patient advocates and FDA SEALD representatives in Dr. Basch's Working Group for the Development of a PRO Version of the CTCAE have expressed that inclusion of multiple concepts would likely not be accepted or clearly comprehended by many patients; (2) patients in pilot work at MSKCC have reported "dissatisfaction" with such items because they felt "forced" into response criteria they did not feel they fit; (3) double-barreled items require abundant verbiage to convey the complexity of conceptual choices and

thus take longer to administer, increasing questionnaire burden (eg, at MDACC and MSKCC, average administration time for 20 MDASI items using 0-10 NRS is 7 minutes, whereas average administration time for 20 STAR Items is 13 minutes); and (4) complex verbiage in double-barreled responses presents challenges when translating items into different languages or adapting items for administration in children or patients with cognitive disabilities.

**Table A.5-7 Disadvantages of Double-Barreling of Concepts in CTCAE Symptom Item Response Options**

| |
|---|
| May not be comprehended or interpreted similarly by all patients |
| May "force" patients to select responses they feel they don't "fit" |
| Requires abundant verbiage in responses, which increases questionnaire burden |
| Requires complex verbiage, which is challenging for language translation |
| May not be feasible for administration in children or cognitively disabled patients |

Based on prior work at MSKCC with STAR items (in the pilot patient adaptions of CTCAE symptom items) and at MDACC with the MDASI, BPI, and BFI, the investigators believe that asking about intensity alone is sufficient for the purpose of adverse event monitoring, and we hypothesize that the incremental additional information provided by also asking about interference or frequency is not sufficient in this setting to merit the additional associated questionnaire burden in the setting of adverse event screening. For example, in retrospective comparisons of patient self-reports of intensity vs. interference for the same symptoms, we found that Pearson correlation coefficients were generally greater than 0.74, suggesting that interference may be somewhat redundant of intensity (or severity). Although we agree that in the setting of in-depth symptom characterization (for example, for assessing a clinical trial efficacy end point) eliciting the full spectrum of concepts related to a symptom of interest is essential (including intensity, interference, frequency, duration, bother, attribution, and location), because adverse symptom event screening must be able to encompass a large number of symptoms efficiently, we believe that the additional questionnaire burden may not merit including additional concepts beyond intensity (or severity) in this setting. We feel that if a single item is to be used in the setting of adverse symptom event monitoring, it should be an intensity (severity) question.

### 5.b.7. Determination of what format is most appropriate for PRO-CTCAE response options (eg, 0-10 NRS, 0-4 NRS, or verbal descriptor scale)

*Background.* The first page of the instructions for CTCAE v3 state that:

| |
|---|
| Grade refers to the severity of the AE. The CTCAE v3.0 displays grades 1 through 5 with unique clinical descriptions of severity for each AE based on this general guideline: |
| Grade 1: Mild AE |
| Grade 2: Moderate AE |
| Grade 3: Severe AE |

| |
|---|
| Grade 4: Life-threatening or disabling AE |
| Grade 5: Death related to AE |

As such, each item is graded on an ordinal scale with numeric grades given primacy, but verbal descriptors are also included and mapped to each numeric grade. For any patient adaptation of CTCAE items, as described previously in his proposal, grade 5 ("death related to AE") would obviously be excluded. Therefore, if a patient adaptation of the CTCAE was intended to be consistent with the current CTCAE grading approach, such patient items would be limited to five response options (ie, grades 0-4).

This RFP provides a unique opportunity to momentarily set aside the notion that PRO-CTCAE response options *must* adhere to the format of the current clinician CTCAE. It allows us to take a step back and ask, if we were creating the PRO-CTCAE from the ground up, what response format would we select. Assuming that it is determined early in the contract period to dispense with concept double-barreling in items and that we dispense with the abundant verbiage in response options of CTCAE items for patient self-reporting, then we believe that there are three reasonable grading scale options for response options in the PRO-CTCAE, each of which has relative advantages (Table A.5-8).

**Table A.5-8. Alternative Choices for Response Format Options in PRO-CTCAE Items**

| SCALE | ADVANTAGES | DISADVANTAGES |
|---|---|---|
| **0-4 NRS** | Consistent with current CTCAE scale<br>May provide sufficient information | May provide less granularity than 0-10<br>Range may be unfamiliar to patients |
| **0-10 NRS** | Consistent with items in other instruments<br>Common scale, well established<br>May provide greater granularity | Not consistent with current CTCAE range |
| **Verbal Descriptor Scale** | Consistent with current CTCAE scale<br>May be easily understood in labels | May not be interpreted the same by all patients<br>Less amenable to translation than NRS |

*NRS, Numerical rating scale*

- *Numeric Rating Scales (NRS).* NRS ask patients to rate their symptoms on simple numerical scales. The scales typically range from 0–5, 0–10, or 0–100. Numeric scales usually are presented horizontally, with a verbal anchor given for each end of the scale. For example, the "0" end of a symptom scale assessing pain may be defined as "no pain," and the other end may be defined as "pain as bad as you can imagine" or "very severe pain." The NRS are generally considered easier for patients to understand than visual analog scales (VASs) (Cleeland, in Osoba 1991). The use of numbers instead of words may remove some sources of cultural and linguistic variation that may occur with verbal descriptor scales (Cleeland, in Foley et al 1990; Cleeland et al, in Jensen et al 1997).
- *Verbal Descriptor Scales.* These category scales (also known as verbal response scales) are used to ask a patient to choose a verbal descriptor that best describes a symptom. A simple verbal descriptor scale consists of "none," "mild," "moderate," and "severe." Although these scales can be useful in symptom assessment, these scales assume that patients have approximately the same meaning in mind when choosing the descriptor that best describes their

symptom. This assumption is questionable given that patients come from diverse educational, cultural, and linguistic backgrounds (Cleeland, in Osoba 1990; Cleeland et al, in Jensen et al 1997). Therefore, the investigators are predisposed against the use of this approach in the PRO-CTCAE because such scales may be difficult for some patients to understand, may not be interpreted in the same way by all patients, and may not be as amenable to language translation as numeric rating scales.

- *Visual Analog Scale (VAS)*. The investigators have not included this as an option to consider for the PRO-CTCAE, both because of logistics involved with electronic administration and because of our interpretation of the literature suggesting this is not an appropriate scale in the setting of adverse event screening. A VAS usually consists of a 10-cm horizontal straight line. One end of the line represents "no symptom," and the other end represents some concept such as "symptom as bad as you can imagine." A patient is asked to select a point along the line indicating how much of the scale is equivalent to the severity (or frequency, bother, etc.) of the symptom. Visual analog scales have been used extensively in pharmacologic and other types of symptom research. Considerable evidence of their reliability and validity has been found. However, the VAS requires that patients understand the analog concept. These scales often necessitate considerable instruction and supervision from clinical or research staff. It has been suggested that the VAS is not appropriate for patients with limited educational backgrounds (Cleeland, in Osoba 1991).

NRS, verbal descriptor scales, and VAS approaches have been used extensively in pain research. In pain assessment in clinical settings, these three scales of severity have been found to provide nearly equivalent data (Jensen et al., Pain 1986). Given this equivalency of information, other criteria become more important in selecting a scale, such as clarity, ease of administration, and simplicity of scoring. In general, all three scales are highly intercorrelated, although the NRS and VAS are most highly correlated with one another. In clinical research, the NRS has been found to be more reliable than the VAS, especially for patients with limited education (Ferraz et al., J Rheumatol 1990; Hollen et al, Qual Life Res 2005). For very ill patients, oral versions of the NRS can be administered. Numerical scales have been endorsed for use in cancer clinical trial instruments because they are easier to understand and score than VAS or VRS (Moinpour et al., 1989). Eleven-point rating scales (0–10) maximize a trade-off between subject ease of responding and increasing reliability with longer numerical scales (Nunnally, Psychom Ther 1978).

Regarding the number of points on a response scale (ie, 0-4 vs. 0-10), it has previously been evaluated whether the number of points makes a difference (and the extent to which the inclusion of adjectival descriptors modifies responses) (Sloan et al, Mayo Clin Proc 2002; Sloan et al, J Clin Oncol 2007; Bogen, in Proc Res Meth 1996). Some have suggested that minor changes in format result in measurement error or bias (Benoit, Pt Rpt News 2006; Cheung et al, Qual Life Res 2004; Cheung et al, J Clin Epidemiol 2006; Cleopas et al, J Clin Epidemiol 2006), whereas others have proposed that the size of measurement error or bias introduced by such response format modifications is within the realm of chance variability and therefore not worthy of concern (Brundage et al, J Clin Oncol 2005; Grassi et al, Euro Resp Hlt QOL Grp 2008).

The choice of response format has been discussed by the investigators before the issuance of this RFP during meetings of Dr. Basch's ongoing Working Group for the Development of a PRO Version of the CTCAE, during which there has been overall consensus that if we were starting from scratch in the absence of an existing CTCAE which already uses a 5-point NRS, then a 0-10 scale would be preferable (and consistent with standard symptom evaluation practices in other settings). However, because there would be a resulting incongruence if the PRO-CTCAE used a 0-10 scale and the CTCAE continued to use a 5-point scale, there would likely be logistical and labeling challenges caused by such an approach. We propose that this question be decided through a consensus meeting early in the award period, including investigators, NCI staff, and selected advisors, based on best available evidence and expert opinion. However, we also include an optional research approach to addressing this question in Appendix 2, which is not included in the current budget. This research could be considered in the future via an alternate funding mechanism.

### 5.b.8. Identification of existing instruments for use as potential sources of terms or items for the PRO-CTCAE

*Background.* Once the above issues are determined through consensus with NCI staff to finalize the format for draft PRO-CTCAE items, and the list of symptoms for adaptation is finalized, we will proceed to create the draft PRO-CTCAE items. The first step toward creating these items will be to evaluate existing instruments that may be used as sources of terms or complete items through a literature review. The investigators and their respective teams have extensive experience in the areas of item development and evaluation and therefore are already aware of many of the existing instruments, relevant publications, and appropriate criteria for evaluating candidate items and terms. A recent systematic review of cancer symptom assessment instruments has been published by Kirkova et al., which will be used as a reference point for our literature review (Kirkova et al, J Clin Oncol 2006).

Our overall aim in creating questions for the PRO-CTCAE items will be to include terms that map to the specific symptom of interest, with more than one term used if necessary to capture responses from patients who may characterize the experience of that symptom differently. For example, an item to evaluate the CTCAE v3 symptom *fatigue* may include the terms *fatigue* and *lack of energy* and *tiredness*. An item to evaluate the CTCAE v3 symptom *neuropathy: sensory* may include the term *numbness* and *tingling*. Avoidance of redundant terms will be a priority, with literature review for each source item and data mining if prudent, to optimize the information gathered from each item while minimizing the number of symptom terms as much as possible (in the interest of parsimony and facilitating more straightforward subsequent language translation). Publications that include data from cognitive interviewing will be particularly helpful, as will studies that demonstrate that specific terms map to symptoms of interest. Data from PROMIS data banks will also be harnessed in this regard, as demonstrated in the "Example Approach to Fatigue" (described later in this proposal), in which PROMIS items are included in the analysis of what terms to consider for the draft PRO-CTCAE fatigue item. We anticipate that this approach will save time and effort during the contract period.

*Literature Review.*

- *Overview.* A literature review will be conducted to identify single-item and multi-item

symptom questionnaires that have been developed and evaluated in cancer populations. Non–cancer-specific questionnaires will also be considered for symptoms with a paucity of prior high-quality development work. As noted above, in 2006, Kirkova et al. identified 21 symptom instruments as appropriate for clinical use in oncology and described in detail 14 of them that assessed more than five symptoms, rated by such criteria as comprehensiveness, psychometric properties, time to complete, and utility for decision making (Kirkova et al, J Clin Oncol 2006). Thirteen of the instruments assessed the patient's perception of the impact of the symptoms, using scales that measure distress, bother, or interference with normal activities and life. Dr. Basch's team at MSKCC has already created binders containing literature review and published instruments for 25 symptoms in the preliminary candidate CTCAE symptom list. Dr. Cleeland's group at MDACC similarly has amassed a large database of supporting literature and instruments used during the development of the MDASI.

- *Approach to Search.* The literature review search strategy and search terms will be designated by Drs. Basch and Cleeland, with input from the investigative team. The search itself will be conducted by Dr. Tito Mendoza at MDACC and Dr. Thomas Atkinson at MSKCC, with symptom-specific files, including search criteria, results, and identified publications maintained by the data coordinator at MSKCC. The investigators have extensive prior experience conducting similar searches.

- *Data Abstraction.* Abstraction of identified publications will be coordinated by Drs. Mendoza and Basch, with input from Drs. Cleeland and Sloan. Table A.5-9 shows the information of interest that we plan to abstract from identified publications for each symptom of interest.

**Table A.5-9. Information to Be Abstracted From Identified Publications for Each Symptom**

| |
|---|
| Available intensity (ie, severity) items for each selected symptom of interest. |
| Available interference items for each symptom (only if it is determined in the above studies that interference items should be included in the PRO-CTCAE, which the investigators suspect it will not). |
| Source instrument of each candidate item. |
| Exact verbiage of question and response options for each candidate item. |
| Patient cognitive interviewing results from development of these questions, with particular attention to the rationale for selection (or rejection) of specific terms, based on patient responses. |
| Measurement properties available for each item: validity, reliability, sensitivity, clinical relevance of score changes, recall period. |
| Correlations available for each item with other measures of the symptom or measures of overall health state or disease status. |
| Specifications of the population(s) in which these items were developed and evaluated. |

- *Analysis.* We anticipate that identified items that will be of the greatest interest to us for the PRO-CTCAE will have been developed with patient input; will have supporting literature demonstrating a clear scientific rationale for inclusion of specific term(s); will have been developed and evaluated in patients with cancer (and preferably evaluated in diverse cancer populations); and will have compelling measurement properties and correlations with other measures of the symptom of interest. These criteria will be used by the investigative team to priority rank the available items, after which the rank order list will be reviewed with NCI staff.

- *Timeline and Deliverables to NCI*. The search strategy and search terms will be developed during month 2 of the contract. The search will be conducted during months 2-3, with abstraction of publications and a rank order list produced during month 4. This list will hence be available for NCI review toward creation of draft PRO-CTCAE items in months 4-5.

### 5.b.9. Issues of intellectual property

The item formats and/or verbiage chosen for the PRO-CTCAE may be similar or identical to items in copyrighted scales. If this similarity of the final items exists, the copyright holders of these existing instruments or their designated representatives will be contacted by Dr. Basch or a designated member of the investigative team, and permission will be requested to use the item/format. Dr. Cleeland's team at MDACC has extensive experience in this area, because the MDASI was developed using items sourced from existing instruments. This issue has been specifically discussed by Dr. Basch with the MSKCC legal department during preparation of this proposal, and they have agreed to provide legal support in this regard at no additional cost to the contract (ie, included in MSKCC indirect costs). In no case will the PRO-CTCAE be represented as other existing instruments or as an equivalent of existing instruments.

### 5.b.10. Creation of Draft PRO-CTCAE Items

*Overview.* Once the questions of item structure, concepts, and response option format have been determined, as well as the finalized list of CTCAE v3 symptoms for adaptation and the ranked list of candidate items/terms from other instruments, we will be ready to proceed with creation of the draft PRO-CTCAE items.

*Generic Structure and Concepts in Planned Questions.* The structure of questions in PRO-CTCAE items will include three elements: periodicity, the specific concept being measured (ie, intensity, interference), and the specific symptom term(s). Magnitude is also included in this format, but by default in the CTCAE, magnitude is the "worst" a particular phenomenon has been during the period of measurement.

Currently, in most cancer treatment trials, periodicity and magnitude for CTCAE v3 items are required to be reported as the worst magnitude since the prior treatment. Hence, although all magnitudes at varying intervals between treatments may be considered, only the worst magnitude is reported. Although such a model may be appropriate for assessment of laboratory evaluation items, for symptoms this approach is problematic because there is growing evidence that patients cannot accurately recall the magnitude of their symptoms beyond several days (Broderick et al, Pain 2008). This issue is specifically addressed in a subsequent section of this proposal, "Substantive Approach to Task 7," in which we propose a preliminary evaluation of recall period for PRO-CTCAE items. Nevertheless, because the recall period for CTCAE assessment is generally study specific, the initial PRO-CTCAE items will be created with a default to inquire about magnitude since the last treatment but with the capacity for clinical investigators to alter this variable. We anticipate that in the future, based on future evaluations of recall period, between-visit evaluations (likely minimally 1-week apart) may become mandatory for PRO-CTCAE reporting. However, such a policy shift is beyond the scope of this proposal and is an issue we plan to discuss with NCI staff should we be awarded this contract.

Items and terms will be considered from candidate instruments as described above. A structure and feel will be selected across all PRO-CTCAE symptoms in a generic format that could be applied to any added future items.

*Example Question Structures.* If we dispense with double-barreled items and only assess severity in PRO-CTCAE screening items, then for category I and II items that are graded, questions may be similar to the following, depending on the source items selected and discussions with NCI staff:

| |
|---|
| **How [concept] has your [symptom term(s)] been since [period of recall]:** |
| **What is the worst [concept] your [symptom term(s)] has/have been since [period of recall]?** |
| **Since [your last treatment], how [concept] has/have your [symptom term(s)] been?** |
| **Since [your last treatment], what was your worst severity of your [symptom term(s)]?** |

Alternatively, a checklist format may be adopted that is similar to the MDASI or Memorial Symptom Assessment Scale (MSAS), in which an overall severity contextual question is asked initially, followed by a list of symptom terms with severity grading scales (with one or multiple items potentially on a page, an issue to be evaluated during usability testing in Task 6).

The feel of such a questionnaire might be as follows:

**Since [your last treatment] please rate the worst [severity] you have had for these symptoms:**

Symptom #1 term(s):

[SCALE----------]

Symptom #2 term(s):

[SCALE----------]

Symptom #3 term(s):

[SCALE----------]

Specific examples questions for fatigue and neuropathy using these formats are provided in the "Example Approaches to Fatigue and Neuropathy" section of this proposal.

For category III symptoms, items will be included that allow for patients to select yes or no for the presence of that symptom (ie, "Since [your last treatment], have you had [symptom term(s)]?").

We recognize that the items envisioned in this proposal will not inherently have the ability to discriminate between symptoms that occur because of drug effect (toxicity) vs. disease vs. a combination of both. However, the investigators concur that attribution is highly complex in the setting of cancer care and is beyond the scope of CTCAE items (and rather can be left to expedited reporting through a linkage of the PRO-CTCAE system to caAERS and AdEERS).

Items will be presented to NCI staff with supporting data in month 5 of the contract, although we envision continuous NCI and advisory input throughout item creation prior to month 5.

### 5.b.11. Ideal properties of the PRO-CTCAE (including validity, reliability, sensitivity, and minimal administrative burden)

Any acceptable symptom measure, whether intended to evaluate an efficacy end point or an adverse event, must demonstrate adequate validity, reliability, and sensitivity to the symptom of interest, with limited variability of interpretation or responses across diverse patient populations. The Medical Outcomes Trust has previously developed a list of recommended review criteria for evaluating the performance of patient-reported measures, including reliability, validity, responsiveness, interpretability, burden of administration, alternative forms/modes of administration, and cultural and language adaptations (Scientific Advisory Committee of the Medical Outcomes Trust; 2002). As such, we concur that an ideal measure will be brief to place minimal burden on ill patients, particularly if the symptom scale is used frequently. A symptom measure should be easy to understand such that patients with limited educational backgrounds can complete it with minimal assistance, demonstrated via cognitive debriefing.

Aside from the inherent measurement properties of the measure, logistical considerations are essential. Feasibility of administration across populations is vital to demonstrate. Multiple methods of administration should be available with demonstrated rough equivalence, including electronic (Web, handheld device, and IVR), paper, and via interviewer. Availability in multiple languages is important, particularly in settings where patients come from diverse racial and ethnic groups. In addition, a measure ideally should have applications in both clinical and research settings to ensure widespread acceptance. The recall period for items should be well understood, as should the clinical relevance of specific score changes in populations and contexts of interest. In multisymptom instruments, symptoms should be included that are clinically relevant to the patient population of interest and to investigators.

We propose to develop and evaluate PRO-CTCAE items adherent to these principles. Although not all of these properties are feasible to evaluate during the 2-year contract period, we are cognizant of all of these issues with plans ultimately to evaluate all of them, including beyond the 2-year contract. Notably, the investigators have been involved in this area of research in various forms for many years and will continue involvements in these areas (now including the ongoing development of the PRO-CTCAE).

In the subsequent section, “Approach to Task 7: Measurement Properties,” we describe a study in 300 patients to evaluate the validity, reliability, sensitivity, clinical relevance of score change (ie, “minimally important difference”), and (preliminary) recall period for the developed PRO-CTCAE items.

### 5.b.12. Identification of ideal patient populations for optimal evaluation of draft PRO-CTCAE items

We propose to conduct the studies described in this application in the routine care setting at the five participating institutions: Dana-Farber, Duke, Mayo, MDACC, and MSKCC (in both tertiary hospital and community satellite clinic sites). We have opted to develop the PRO-CTCAE in these settings because we are confident that we can obtain IRB review and accrue patients quickly at our home institutions. Furthermore, this network will provide access to diverse patient populations.

For ultimate feasibility assessment (described in Task 8), we will amend Dr. Basch's current CALGB protocol 70501 of patient self-reporting of CTCAE items to evaluate the new draft PRO-CTCAE items in the cooperative group setting.

For the proposed studies, we have selected the following disease states for inclusion: breast cancer receiving adjuvant chemotherapy, metastatic prostate or bladder cancer receiving chemotherapy, head and neck/gastroesophageal carcinoma receiving daily chemo-radiotherapy, metastatic lung cancer receiving chemotherapy, metastatic colon cancer receiving chemotherapy, and lymphoma/myeloma receiving chemotherapy. The intention of including these specific diagnoses is to ensure that the pool of interviewees will include patients who both have and have not experienced specific PRO-CTCAE symptoms under evaluation. Our sampling strategy includes patients receiving primary therapy with curative intent (eg, myeloma/lymphoma, adjuvant breast cancer therapy) and with palliative intent (eg, metastatic lung or prostate cancer). As shown in Table A.5-5, all symptoms in the preliminary list of candidate symptoms are prevalent in one or more of these diagnostic categories.

### 5.b.13. Initial language for PRO-CTCAE and subsequent translation plan

The initial language of the PRO-CTCAE will be English. As described in the "Substantive Approach to Task 3," during item development, Dr. Vish Viswanath and his staff in the Health Communication Core of the Dana-Farber/Harvard Cancer Center will advise the investigative team regarding considerations for future translation of items to other languages, initially Spanish and French. The Core will evaluate the tone and terminology of items to ensure that they are accessible, professional, and free from slang or colloquialisms. In addition, the use of cultural references or regional variations will be avoided. Aside from the terms themselves, we will prepare to adapt the organization and presentation of items to fit the differing lengths and structures of various anticipated languages. After the contract period, the investigators are interested in continuing collaborative work with Dr. Viswanath's Core toward initial translation of the PRO-CTCAE items into Spanish and French and potentially other languages. In addition, Larry Shiman's team in the Data Technologies Core in the Population Sciences Division at the Dana-Farber Cancer Institute has the capability to translate the planned education/training materials for the PRO-CTCAE into multiple languages.

### 5.b.14. Determination of cut-point scores for each item

As described previously in section 4, "Rationale for Approach," and discussed in the later "Substantive Approach to Task 7: Measurement Properties" section, an important function of the PRO-CTCE system will be to trigger automated alerts to clinicians for concerning levels of symptom severity (ie, "moderate" levels) and also to prompt expedited reporting at prespecified threshold levels (ie, "severe" levels). Therefore, we will use established methods to determine threshold cut-point scores for each PRO-CTCAE item toward this system functionality (Palos et al, J Pain 2006; Serlin et al, Pain 1995; Anderson, Pain 2005; Jensen et al, Pain 2001; Paul et al, Pain 2005; Zelman et al, Pain 2003; Mendoza et al, Clin J Pain 2004; McConahay et al, Acad Emerg Med 2006; Chow et al, J Palliat Med 2006; Dihle et al, Clin J Pain 2006; Zelman, 2003; Jensen et al, Pain2001). Determinations of cut points will be based on data obtained in Task 7, through mining of data sets at MDACC and MSKCC, and through literature review. Clustering of severity scores within the selected scale with respect to interference anchor measures will be

used to establish cut-point boundaries for alerts to clinicians for each PRO-CTCAE item (at "moderate" levels, generally ≥5 on 0-10 NRS) and for expedited reporting (at "severe" levels, generally ≥7 on 0-10 NRS), although we recognize that cut points may differ for different symptoms (for example, we have observed in work at the MDACC and MSKCC that cut points for shortness of breath tend to be relatively slightly lower than other symptoms by one point on average). Both empirically derived and clinician recommended thresholds will be used in determining the final cut points. If a 0-4 NRS (or verbal descriptor scale) is used, then ≥3 will likely be used as the trigger threshold for alerts.

### 5.b.15. Specific deliverables and timeline for Task 2

The schedule of specific deliverables to NCI staff is detailed in Table A.5-10.

**Table A.5-10. Specific Deliverables and Timeline for Task 2**

| SPECIFIC DELIVERABLES | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| List of candidate CTCAE v3 items for PRO-CTCAE | Month 2 |
| List of candidate "source" instruments for PRO-CTCAE | Month 4-5 |
| Creation of draft PRO-CTCAE items | Month 5 |

### *5.c. Substantive Approach to Task 3: Assure Cultural Literacy & Lack of Bias*

| | |
|---|---|
| Task 3: | Health Literacy, Avoidance of Bias, Minimal Criteria for Collecting Patient Reports, and Language Translation Plan |
| Years of Contract: | Year 1 and 2 |
| Task Leader: | Dr. Vish Viswanath |
| Deputy Task Leader: | Dr. Jennifer Hay |

**5.c.1. Overview**

We will use techniques to ensure that PRO-CTCAE items are developed at an appropriate education level, are designed and evaluated to avoid bias (eg, age, sex, ethnicity, education, geography), and are structured to be optimally amenable to language translation.

It is of paramount importance to the investigators that the developed PRO-CTCAE items and methods of administration be widely generalizable without bias across cancer patient populations regardless of age, sex, educational level, ethnicity, geography, or other patient characteristics. In addition, we aspire to create items that are amenable to future language translation and adaptation for administration to children and patients with cognitive disabilities. Toward these goals, Dr. Vish Viswanath, who leads the Health Communication Core of the Dana-Farber/Harvard Cancer Center, will actively participate in Tasks 1, 2, 4, and 6 of this proposal, as described below.

**5.c.2. Assurance of health literacy and minimization of bias**

During Task 1 of this proposal (preparation of the report titled, *Ideal Conditions, Potential Barriers, and Strategies to Implement the PRO-CTCAE*), Dr. Viswanath and his staff will consult with the Task leader, Dr. Deborah Bruner, to ensure consideration of pertinent issues of health communications, literacy, and cultural bias. During Task 2 of this proposal (item development), Dr. Viswanath and his staff will actively consult with Dr. Basch and other key members of this Task, including Drs. Cleeland, Mendoza, and Sloan, to ensure readability and avoidance of cultural bias in items. The Health Communication Core will provide readability guidelines in advance of item creation and will review and provide feedback about drafts of items as development progresses. Particular attention will be given to vocabulary and sentence structure, organization of ideas, layout, and presentation. To evaluate readability and ensure that the items are presented in clear, plain language, the Health Communication Core's medical writers will review item drafts. If necessary, the Health Communication Core will define or explain technical terms to minimize any barriers to understanding items.

During Task 4 of this proposal, Dr. Viswanath's team will consult with Dr. Basch and the Task leaders, Dr. Jennifer Hay and Larry Shiman, to include verbal probes in cognitive

interviews that address issues of cultural literacy. Responses will be analyzed to refine items.

Patients from diverse sociocultural, ethnic, racial, educational, and geographic backgrounds will be asked to participate in usability and feasibility testing (Tasks 6 and 8). This may include eligibility criteria that specify enrichment for recruitment at particular clinical sites. Feasibility and usability setting will be distributed across the four key collaborating institutions (MSKCC, DFCI, Duke, and M.D. Anderson and their vast networks of affiliated cancer clinics) specifically to ensure that diverse representation is included. Dr. Viswanath's team and Dr. Hay will review the results of usability and feasibility testing to assess for systematic differences in response based on nonclinical patient characteristics.

Input from the Health Communication Core during item development and evaluation is intended to ensure (1) the use of plain language and natural tone; (2) avoidance of unnecessarily academic or clinical language; (3) avoidance of content or examples that may evoke unnecessarily emotional responses; (4) balanced representations of sex, age, ethnicity, culture, and socioeconomic status; (5) that the context of items is clear and makes no assumptions; (6) that the organization and language used are consistent from item to item; and (7) evaluation of usability and feasibility data for systematic differences in responses based on nonclinical patient attributes.

**5.c.3. Determination of minimal patient criteria to collect PROs**

During month 1 of the contract, and in collaboration with Dr. Viswanath's Health Communication Core and NCI staff, the investigators will finalize a list of minimal criteria for collection of PROs using the initial version 1 of the PRO-CTCAE system that we develop during the contract period. Categories to be considered will include age, educational level, and cognitive ability. At the time of preparation of this proposal, the perspective of the investigators is as follows:

*Consideration of Educational Level.* The investigators propose that PRO-CTCAE items be designed for use by patients of all educational levels through use of a standard plain language approach and consultations with health literacy experts (described above) and through item refinement via cognitive interviewing (described in Task 4).

*Consideration of Age.* The investigators propose to design items with the intention of initial administration and evaluation in adults and future planned adaptation for use in children and their proxies. Toward this, Dr. Viswanath's Health Communication Core will evaluate items to ensure adaptability for pediatric use. In addition, two experts in symptom and adverse event evaluation in children are included as advisers on this proposal (Dr. Joanne Wolfe and Dr. Steve Joffe; see Letters of Support). Drs. Wolfe and Joffe will actively participate during the item development process with an eye toward

future administration to children. Children are not included in the initial planned evaluation studies of the PRO-CTCAE, because the investigators believe the additional required steps to adequately evaluate items in children are not feasible during the contract period, given the other complex Tasks of this application. In addition, we believe that the overall model and logistics of administration should be worked out in adults before subjecting children to questionnaire burden. Evaluation and administration of PRO items in young children pose specific challenges that the investigators recognize are essential to address, and thus such work is planned once the model has been proven in adult patients (Varni et al, J Pediatr Psychol 2007; Langer et al, J Clin Epidemiol 2008; Varni et al, Health Qual Life Outcomes 2007).

*Consideration of Cognitive Ability.* Ultimately, it is the goal of the investigators that the PRO-CTCAE items and technology platform be adapted for and evaluated in patients with varying levels of cognitive disability, for example, in patients who have undergone resection of brain tumors. It is our intention to create an open-source system that could be adapted by investigators in the future for use in such populations, and the investigators are also themselves interested in such development work in the years after completion of this contract. However, development work explicitly in this population is beyond the scope of this proposal. It is understood that patients with cancer receiving chemotherapy may have varying levels of cognitive ability, and it is the intention of the proposed work to create items that are appropriate for administration across this spectrum of ability in general, as represented in the population of patients enrolled in NCI-sponsored clinical trials. It is toward this aspiration that Dr. Viswanath's team is involved in the item development process.

**5.c.4. Planning for Future Language Translations**

Dr. Viswanath and the staff of his Health Communication Core will advise the Task 2 leaders about considerations for future translation of the items to other languages, initially Spanish and French. The Core will evaluate the tone and terminology of items to ensure that they are accessible, professional, and free of slang or colloquialisms. In addition, the use of cultural references or regional variations will be avoided. Aside from the language itself, we will be prepared to adapt the organization and presentation of the items to fit the differing lengths and structures of different languages. After the contract period, the investigators are interested in continuing collaborative work with Dr. Viswanath's Health Communication Core toward initial translation of the PRO-CTCAE items into Spanish and French.

### 5.c.5. Specific deliverables and timeline

The schedule of deliverables to NCI staff is detailed in Table A.5-11.

**Table A.5-11 Specific Deliverable and Timeline for Task 3**

| SPECIFIC DELIVERABLE | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| Finalized list of minimal criteria (age, educational level, cognitive ability) for collection of patient reports | Month 1 |

### *5.d. Substantive Approach to Task 4: Conduct Cognitive Interviews*

| | |
|---|---|
| Task 4: | Design and Conduct Cognitive Interviews in Diverse Patient Populations; Clinical Staff Members to Ensure Comprehension of Verbiage in Draft PRO-CTCAE Items and to Ensure that Terms Map to Intended Concepts |
| Years of Contract: | Year 1 |
| Task Leader: | Dr. Jennifer Hay |
| Deputy Task Leader: | Mr. Larry Shiman |

**5.d.1. Overview**

Since the 1980s, cognitive interviewing has become widely accepted for ensuring patient comprehension of questionnaire items and assessing whether items correspond to concepts of interest (Willis, *Cognitive Interviewing* 2004; Murtagh et al, Palliat Med 2007; Rosal et al, Diabetes Educ 2003; Drennan, J Adv Nurs 2003). Cognitive interviewing techniques will be used at two different points during item development in Task 2.

Design of the cognitive interviewing script and protocol will be under the leadership of Dr. Jennifer Hay, who will devote 15% effort to this work during year 1 only. Dr. Hay and her staff at the MSKCC Behavioral Sciences Core have extensive experience in cognitive interviewing. Drs. Basch and Hay are ongoing collaborators evaluating the initial MSKCC "STAR" (Symptom Tracking and Reporting) PRO adaptations of CTCAE items.

The cognitive interviews will be conducted by trained staff under the supervision of Larry Shiman, Director of the Data Technologies Core in the Population Sciences Division at the Dana-Farber Cancer Institute. Experienced interviewers from his Core will travel to each participating site to conduct interviews, and his Core will compile the results. Dr. Hay and her staff will then be primarily responsible, with Dr. Basch, for analyzing and reporting interview results toward item refinement in consultation with the investigative team and NCI staff.

**5.d.2. Optional study design to evaluate patient understanding of double-barreled concepts in response options and preferences for various response formats**

An optional study is presented in Appendix 2, but is not included in the proposed budget.

**5.d.3. Study design to evaluate draft PRO-CTCAE items**

*Rationale for Study*. Once we have developed initial draft PRO-CTCAE items based on the processes described in the "Substantive Approach to Task 2" section, the investigators will consider it essential to evaluate whether patients understand the verbiage used in these draft items and that terms map to concepts of interest. Cognitive interviewing with semistructured questionnaires and verbal probing is an ideal approach to address these issues and to generate practical data for refining items.

*Study Design and Patient Population.* We will conduct two rounds of cognitive interviews in a total of 60 patients with cancer receiving chemotherapy at four of the participating centers, including the main hospital and satellite sites for each site (DFCI, Duke, MSKCC, M.D. Anderson).

Cancer diagnoses for the target populations will be selected based on prevalence of specific symptoms of interest in such populations, as specified in Table A.5-5 in the "Substantive Approach to Task 2" section. As described in Table A.5-5, eligible disease groups will include the following: breast cancer receiving adjuvant chemotherapy; metastatic prostate or bladder cancer; head and neck/gastroesophageal carcinoma receiving chemo-radiotherapy; metastatic lung cancer; metastatic colon cancer; and lymphoma/myeloma. The intention of including these multiple specific diagnoses is to ensure that the pool of interviewees includes patients who both have and have not experienced the symptoms of interest. Patients will be stratified based on age, ethnicity, sex, and educational level. An approach will be used, which the investigators have used in the past, to ensure that data are available for each item from patients with varying representative demographic characteristics.

The sampling strategy includes patients receiving primary therapy with curative intent (eg, myeloma/lymphoma, adjuvant breast cancer therapy) and palliative intent (eg, metastatic lung or prostate cancer). Patients receiving adjuvant treatment are unlikely to have symptoms related to their underlying malignancy since they have been rendered free of measurable disease by surgery. Therefore, symptoms encountered in patients receiving adjuvant treatment are most likely to be attributable to the therapeutic agents themselves. Evaluation of PROs in the adjuvant setting is optimal for determining treatment toxicity. However, when therapy is administered with curative intent, both patients and clinicians have a greater willingness to tolerate adverse symptoms. The rationale is that therapy is of finite duration and that greater toxicity translates into higher potential for long-term cure. This creates a paradox. The adjuvant setting is ideal for systematic evaluation of symptoms attributable to treatment (as opposed to the underlying malignancy), but yet the degree of symptom burden is simultaneously less salient when cure is the goal of therapy. Most initial drug development in cancer takes place in the setting of advanced incurable disease. In this circumstance, it is often difficult to disentangle those symptoms that are attributable to disease from those attributable to the therapeutic intervention. In the context of randomized trials, comparison of two or more study arms eliminates this obstacle, but in the phase II setting it is not easy to discern whether symptom burden is disease or treatment related. Recognizing these issues and the rationale for including PRO instruments in the setting of phase II and phase III trials in the metastatic setting and phase III trials in the adjuvant setting, patient populations that span the spectrum have been selected for inclusion in system development.

Unlike the cognitive interviewing study described above, this study will be conducted under an IRB protocol because we intend to screen patients for eligibility based on stratification variables

and collect PHI. The protocol will be developed by Drs. Hay and Basch with input from Larry Shiman and the investigative team, according to the timeline described below.

*Questionnaire.* A semistructured interview script will be developed by Dr. Hay and will be reviewed by Larry Shiman, the investigative team, and NCI staff. Experienced interviewers from Dr. Shiman's Core will travel to each study site to administer questionnaires, including draft PRO-CTCAE items of interest, and then "probe" respondents using the script to assess comprehension (level of understanding of instructions, questions, and response options), clarity (level of straightforward meaning of instructions, questions, and response options), knowledge/memory (ease of recall of information needed to respond), and judgment (ease of fitting personal responses to response options).

*Enrollment and Study Procedures.* Staff members at each site have been previously contacted and will identify eligible patients based on the stratification variables and will also provide private space for each interview. Local staff will acquire consent from the patients, and we have budgeted a per-patient financial incentive for each site to identify patients and provide interviewing space. Stratification and patient registration will be coordinated by the MSKCC Coordinating Center Project Manager who will be in constant contact with local clinical research associates/data managers, as well as with interviewers from Dr. Shiman's staff.

The first round of interviews will include up to 60 patients. Each patient will complete a questionnaire, including selected PRO-CTCAE draft items, and then will undergo retrospective semistructured "probing" by the interviewer. All interviews will be audio recorded. Based on our prior experience in this area, we will design the questionnaire and interview to be conducted in a total of 60 minutes. After analysis of interview results and item refinement (described below), a second round of interviews will be conducted in up to 40 patients via similar methods. (Interviews in 54 patients have been budgeted for by the DFCI Data Technologies Core, and other interviews will be conducted by MSKCC staff).

*Analysis.* Qualitative analysis of the results of the first round of cognitive interviews will follow methods outlined by Willis (Willis, *Cognitive Interviewing* 2004). A codebook will be maintained by Larry Shiman's Core to synthesize responses to questions and verbal probes. Two members of the Core will code and organize responses for each item and prepare a summary for each item and patient stratum with particular attention to the following dimensions: comprehension, clarity, knowledge/ memory, and judgment. Members of Dr. Hay's team will then review summary notes and coding discrepancies, with final decisions made by consensus. Coded comments will be summarized in a table. Items eliciting confusion or misinterpretation will be considered for revision. The second round of cognitive interviews will then be conducted for the revised draft items, and analysis will follow the same overall approach. Items will again be refined after the second round of interviews and subsequently will be uploaded to the PRO-CTCAE electronic platform for usability testing (Task 6).

*Timeline and Deliverables to NCI Staff.* The semistructured interview script will be created by Drs. Hay and Basch with input from Larry Shiman and the investigative team and advisers during month 1 of the contract period and will be available to NCI staff in month 2. An IRB protocol will be developed, and we have allowed a 4-month period for IRB review and approval at the participating sites (between month 2 and month 5). The first round of interviews will be conducted during months 5-6. Analysis and item refinement will occur during month 7 of the contract. The second round of interviews will occur during months 7-8, and items will undergo refinement toward finalized form for upload to the PRO-CTCAE platform and usability testing in month 9. Final results of the 2 rounds of cognitive interviewing will be available for NCI staff review in month 10 of the contract.

**5.d.4. Specific deliverables and timeline for Task 4**

Detailed timelines for processes in each cognitive interviewing study are described above. The schedule of specific deliverables to NCI staff is detailed in Table A.5-12.

**Table A.5-12. Specific Deliverables and Timeline for Task 4**

| SPECIFIC DELIVERABLES | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| ▪ Cognitive interview script for NCI review | Month 2 |
| ▪ Results of interviews to NCI | Month 10 |

### *5.e. Substantive Approach to Task 5: Develop PRO-CTCAE Software System*

| | |
|---|---|
| Task 5: | Develop an open-source software system for patient self-reporting of adverse symptom events that capitalizes on prior evaluations of similar platforms, adheres to standards and federal regulations for data security and privacy, is mapped to standardized lexicons, and is interoperable with current and emerging platforms for adverse event and data management in cancer trials (CDUS, AdEERS, MedDRA, caAERS/CTCAE v4, caBIG/NCIB, CDISC). |
| Year of Contract: | Year 1 |
| Task Leader: | Mr. Ram Chilukuri and Mr. Vinay Kumar, SemanticBits, LLC |
| Backup Programmer: | Mr. Marwan Shouery, MSKCC |

**5.e.1. Overview of PRO-CTCAE platform**

We envision a PRO-CTCAE assessment software system (ie, “PRO-CTCAE platform”) that allows clinical investigators to create electronic case report forms (CRFs) populated with PRO-CTCAE items via an online Publishing Tool/Form Builder; includes linkage to the caBIG study calendar to prompt and track CRF administration; provides a Web-based Patient Interface for completion of items at visits; generates paper CRFs for backup questionnaire administration; incorporates a rules-engine that can trigger electronic pre-populated expedited reporting forms for caAERS/AdEERS, email alerts to clinical staff, or prompt administration of additional items or questionnaires to patients in real-time; includes an automated reminder system for missed CRFs; provides a query engine for reviewing data at the level of patient, site, or study; interoperates with caAERS, CDUS, AdEERS and other key systems; adheres to caBIG/NCICB and CDISC standards for interoperability, security, and privacy; is open-source; and is designed to incorporate interactive voice technology (IVR) and handheld device-based administration in the future. These platform functionalities and interrelationships are shown in Figure A.5-3 on the following page.

To build the PRO-CTCAE platform, we propose to subcontract to SemanticBits, LLC, a software development company with extensive experience with caBIG/NCICB projects that is highly relevant to this proposal. Dr. Basch and developers at MSKCC have met numerous times with SemanticBits and thoroughly reviewed their development processes and past projects, and concluded that SemanticBits is uniquely positioned and qualified to build this system. The basis for this decision and a description of SemanticBits’ qualifications are described in the following sections, as are technical details of our planned approach to platform functionality, interoperability, and security.

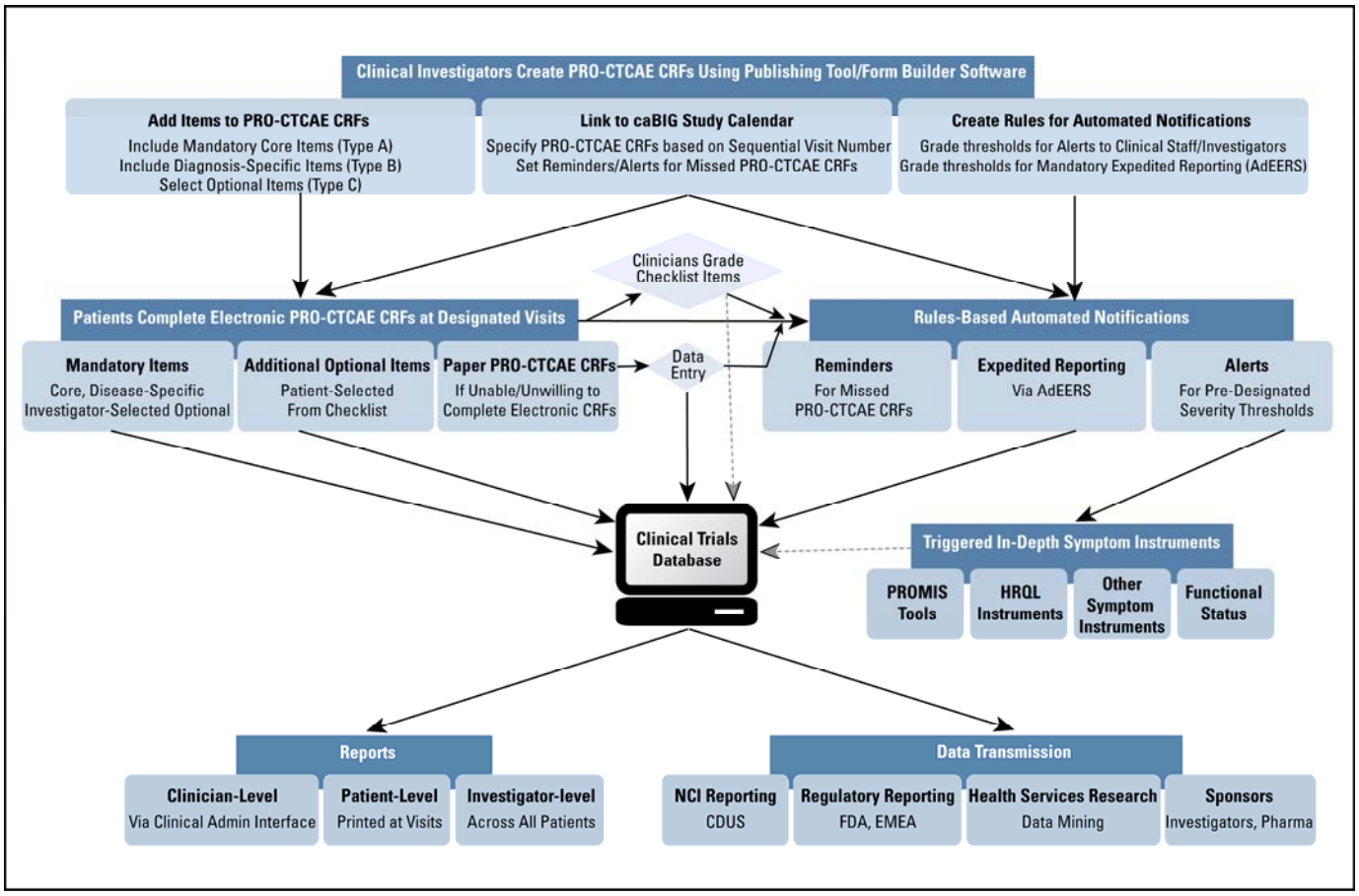



**Figure A.5-3 . Proposed Functionality and Systems Interoperability for PRO-CTCAE Platform**

As shown in Figure A.5-3, prior to initiating a trial, the clinical investigative team will login to the PRO-CTCAE Publishing Tool software to specify which symptoms they wish patients to self-report during the trial, at which specific visits. Based on this information (and optional linkage to the electronic study calendar), the Publishing Tool software will automatically generate the electronic CRF that patients will be asked to complete at applicable visits. At the time of trial enrollment, clinical research staff will train patients to self-enter data into the PRO-CTCAE questionnaire (CRF) via a clinic-based computer. This might be via a desktop computer using a mouse, via a touchscreen computer kiosk (eg, similar to used by airlines for check-in), or via a wireless touchscreen tablet computer (the approach currently used in CALGB 70501). Patients will self-enter data at that time as a practice/baseline measurement. Subsequently, at each followup visit, patients will be reminded by staff to self-enter their data into the PRO-CTCAE questionnaire (CRF). The CRF will include the items specified by the investigator via the Publishing Tool, but patients will also be able to add additional “optional” items from a checklist of other symptoms. Clinical staff will be trained to assist patients with technical questions, and a telephone support hotline will be available. After patients self-report, their information will be available in a longitudinal Report in real-time. This Report will be available to clinicians

for review and sign-off either electronically or by paper, depending on the configuration of the particular clinic or trial (electronic is preferable and is envisioned as the standard approach ultimately). For particular adverse symptom grades, clinical staff will be required to complete CRFs/expedited reporting forms (based on rules established by the investigator up front using the Publishing Tool software). Emails to clinical staff with links to these forms will be automatically triggered when patients hit the pre-specified grade thresholds. The PRO-CTCAE system will track compliance during the study, and will recognize if patients do not login at pre-specified times, and will track if clinicians complete their required forms. This information will be available to data managers, and automated email reminders will be triggered to staff to complete required forms. An audit trail will be maintained of all data entry.

The PRO-CTCAE system fits into the process of care at the interface between patients, caregivers, and investigators. Patients are asked to login when they arrive for appointments, and this information is available for review by clinicians in real-time. Therefore, PRO data can be used as the basis for discussions during visits, and as the basis for interventions.

Patient-entered data will be available to clinicians in real-time. Therefore, clinicians will be able to review this information at the time of clinic visits to consider interventions for adverse symptoms of concern. Clinicians are required to sign off on selected PRO-CTCAE symptoms in the proposed model. Specifically, these are a subset of items for which patients can report presence/absence, but not level of severity (eg, ataxia, bruising, edema) (Section 5.b.4; Table A.5-5). For such items, we envision a model in which patients self-report presence of the phenomenon, and then clinicians receive this information and must respond to it by grading or further characterizing the phenomenon. This functionality is planned for the PRO-CTCAE platform (as described in Table A.5-14 and Figure A.5-3), and will be included in the cooperative group feasibility trial.

**5.e.2. Selection of SemanticBits, LLC to create the PRO-CTCAE platform**

*SemanticBits Experience with NCI projects.* Because of their strong background in software development for NCI and caBIG, SemanticBits is uniquely and strategically positioned to deliver a product that is adherent to the requirements of CTEP, NCI/CBIIT, CDISC, and industry stakeholders. They are the primary developers for the NCI Cancer Adverse Events Reporting System (caAERS), which is an essential system for compatibility with PRO-CTCAE. caAERS, as described below in detail, is an open-source, standards-compliant application designed to collect, assess, and manage adverse events in cancer clinical trials. As such, while building this system, SemanticBits has gained an in-depth understanding of the adverse event reporting mechanism in cancer trials, AdEERS, CDUS, CTCAE, and relevant NCI requirements and personnel. SemanticBits is also the lead designer for the Cancer Central Clinical Participant Registry (C3PR) and developers on the caGrid and Patient Study Calendar teams. From this

experience they have gained a deep understanding of the software development philosophy, tools, and technology stack of NCI/CBIIT and CTEP. All of their application development is open-source with a focus on open access and federation. Table A.5-13 shows examples of NCI projects which they either lead or in which they actively participate.

**Table A.5-13. SemanticBits Leadership and Participation in Key NCI Informatics Projects**

| |
|---|
| ***caAERS:*** The Cancer Adverse Event Reporting System (caAERS) is an open source, standards-compliant application designed to collect, assess, and manage adverse events in cancer clinical trials. On this project, SemanticBits serves as the lead. The funded caBIG Adopters are the Mayo Clinic and Wake Forest University. The first phase of this contract ran from December 2006 to February 2008. In March 2008, they were awarded a follow on 12 month contract for work on phase 2. |
| ***caBIG Patient Study Calendar (PSC):*** The Patient Study Calendar (PSC) is an open source, standards-compliant, Web-based application that assists with the management of the activities of subjects on clinical trials. SemanticBits serves as a key member of the development team. |
| ***caGrid:*** caGrid provides the underlying service-oriented infrastructure for caBIG. As part of an initial 12-month contract, they developed several key components of caGrid, including the caGrid Installer, caGrid Authz, caGrid Browser and caGrid Portal. They also developed caBIO and caArray data grid services and deployed them on NCI grid node. In April 2007, they were awarded a 9-month contract extension, during which they developed and released caGrid Portal 2.0.Cancer Common Ontologic Research Environment (caCORE). |
| ***C3PR:*** The Cancer Central Participant Registry (C3PR) Release 2 is a Web-based application used for end-to-end registration of patients to clinical trials. SemanticBits is the development lead on the project. Duke University is the project lead. The funded caBIG Adopters are the Duke University and Wake Forest University. The initial contract ran from November 2006 to March 2008. They were recently awarded a follow on contract for an additional 12 months. |
| ***caTRIP:*** The Cancer Translational Research Informatics Platform (caTRIP) was developed as part of the Cancer Biomedical Informatics Grid (caBIG) project. It allows users to post a query across a number of caBIG data services and view the results in a user-friendly interface. Its initial focus is outcomes analysis, which allows clinicians to query across data from existing patients with similar characteristics to identify patients that were treated successfully. This project was a joint collaboration between SemanticBits and Duke Comprehensive Cancer Center. |
| ***QBT:*** The Quantitative Biological Tool (QBT) is being developed under a project funded by National Heart, Lung, and Blood Institute (NHLBI). QBT is designed to define, visualize, annotate, and simulate biological systems of interest. The intended user is a scientist or clinician who needs to conceptualize their system of interest, organize their experiments, review relevant literature, and simulate the system's behavior. |

SemanticBits has undergone the caBIG silver compatibility review process for three of their applications for which they generated well defined, strongly typed application programming interfaces (APIs) backed by robust semantics. Furthermore, they have leveraged the caGrid infrastructure in all projects, providing for a way to seamlessly interoperate between systems. For example, they exposed services in the caBIG Clinical Trials Suite (CCTS) for exchanging study definitions and subject registration details.

SemanticBits leverages caCORE and caGrid technologies wherever appropriate for generating open APIs. The caCORE SDK facilitates model driven architecture/ development (MDA/MDD) by providing mechanisms for generating data objects, XML Schemas, and services directly from UML. Furthermore, they leverage the robust security infrastructure of caCORE and caGrid. The Common Security Module (CSM) provides a way to perform local authentication and make local authorization decisions based on a user's role and the resource he or she is trying to access. The Grid Authentication and Authorization with Reliably Distributed Services (GAARDS) in the caGrid toolkit provides services and tools for the administration and enforcement of security policy in an enterprise Grid. They leverage caGrid security for identity management, trust, authentication, authorization, single sign-on (SSO), and delegation. All of their applications are derived from rich information models defined in the unified modeling language (UML), which ultimately are registered as common data elements (CDEs) in the caDSR with standards-based vocabularies registered in the EVS. When modeling a domain, they leverage community standards wherever possible. For all caBIG projects, they have undergone several rounds of intensive harmonization with the comprehensive domain analysis model that is being developed by the Biomedical Research Informatics Domain Group (BRIDG), providing a robust information foundation for many of the critical aspects of their application from data modeling all the way through to the user interface. SemanticBits will adopt a similar approach in the development of PRO-CTCAE system. The areas of programming and domain expertise that SemanticBits will bring to this project, based on their ongoing work with caBIG/NCICB are shown in Figure A.5-4.

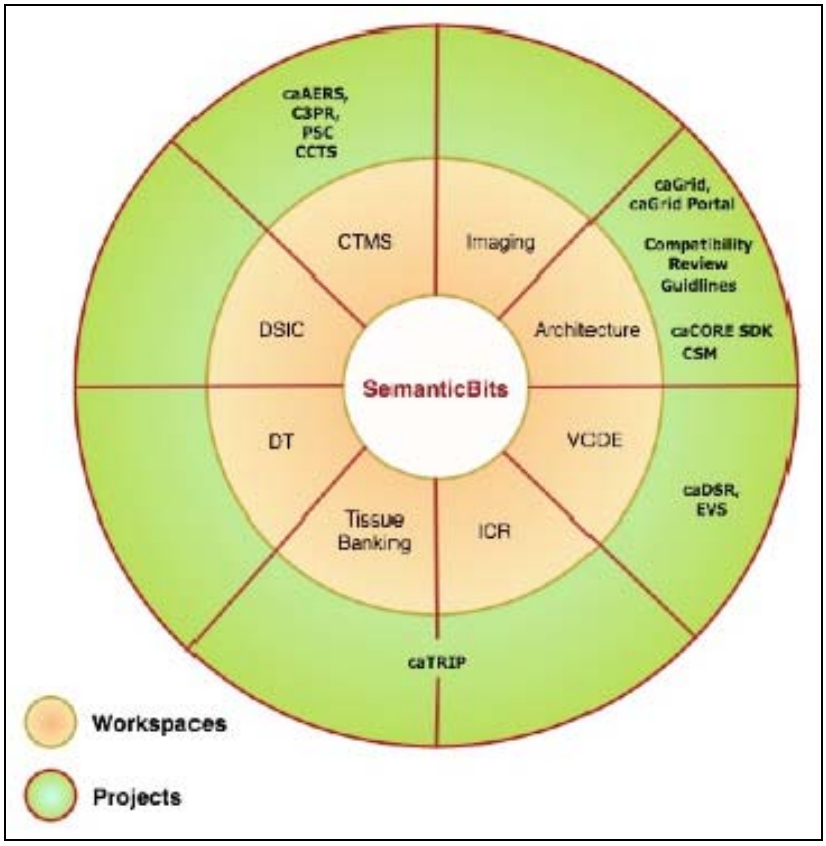



**Figure A.5-4. SemanticBits caBIG/NCI expertise**

**5.e.3. Rationale for selecting SemanticBits to create the PRO-CTCAE platform**

Although both Dr. Basch and Dr. Abernethy have well-developed programming teams with extensive experience creating PRO platforms, we have decided to subcontract the technology development for the PRO-CTCAE to SemanticBits, LLC. This is because:

1) SemanticBits works closely with NCI and has in-depth inside knowledge of systems which are essential for integration with PRO-CTCAE (including caAERS, AdEERS, CDUS, caBIG study calendar);
2) SemanticBits creates open-source platforms in keeping with caBIG/NCICB requirements so there will be no issues related to source code or intellectual property;
3) SemanticBits has been highly reliable in contracts with caBIG/NCICB and has an excellent working relationship with John Speakman, Ann Setser, and other key individuals to this contract;
4) Many of the design elements necessary for PRO-CTCAE have already been built by SemanticBits for caAERS, such as automated notifications, integration with expedited reporting, and integration with study calendar. This will expedite development for PRO-CTCAE.

We believe that these factors will assure that a high-quality product is produced according to timeline, with full system functionality, security and privacy provisions, and compatibilities we feel are essential to the PRO-CTCAE platform. In other words, working with SemanticBits will assure that we get it right the first time. SemanticBits has unique access and experience that no other developer in this space has, and thus working with them is an enormous asset on this project.

Notably, we have also budgeted effort of Marwan Shouery, a lead programmer in Dr. Basch's group at MSKCC, who will work closely with SemanticBits throughout the project to assure quality and comprehensiveness. Mr. Shouery has programmed numerous PRO platforms in the past and is himself capable of creating this system. He and other programmers in Dr. Basch's group will therefore serve as a backup to SemanticBits throughout the contract period. We have budgeted site visits for Mr. Shouery and Dr. Basch to SemanticBits during the contract.

**5.e.4. SemanticBits development of the NCI caAERS system**

As noted above, a key system of relevance to the PRO-CTCAE is the Cancer Adverse Event Reporting System (caAERS). We envision that PRO-CTCAE will be a stand-alone plug-in to caAERS. This system has been developed by SemanticBits.

In December 2006, SemanticBits began development of caAERS under the sponsorship of caBIG. In March 2008, they released the first official version of the software (caAERS 1.0). Their partners included the Mayo Clinic Comprehensive Cancer Center (MCCC), Wake Forest Comprehensive Cancer Center (WFCCC), Northwestern University's Robert E Lurie Cancer Center, and Akaza Research. SemanticBits also worked hand-in-hand with representatives from several pivotal groups in the adverse event reporting field, including the Cancer Therapy Evaluation Program (CTEP) and the Division of Cancer Prevention (DCP). In addition, their team received valuable input from subject matter experts at City of Hope, Duke University, Harvard University, and a number of Cooperative Groups, including the Cancer and Leukemia Group B (CALGB).

caAERS v1.0 is currently under adoption at MCCC, WFCCC, and CALGB. It is an open source, Web-based application for documenting, managing, reporting, and analyzing adverse events. The system operates both as a repository for capturing and tracking routine and serious AEs and as a tool for preparing and submitting expedited AE reports to regulatory agencies. caAERS works with cancer prevention and therapeutic trials and can accommodate a range of interventions and study modalities, including investigational and commercial agents, radiation, surgery, and medical devices. Adverse events can be coded in caAERS using either CTCAE or MedDRA.

The caAERS 1.0 system features a powerful, state-of-the-art rules engine, which can capture a range of sponsor, institution, and protocol-level reporting requirements. Using these rules, caAERS can automatically determine if an adverse event requires expedited reporting and when and to whom the report must be submitted -- for any of an organization's trials. The business rules used by caAERS can be authored within the application itself or imported from a library of approved rule sets. caAERS also features an advanced email-based alert system that can be customized along a number of dimensions (message content, recipients, delivery times) to ensure that notifications and reminders are sent out as needed. Also included with caAERS is an easy-to-use report template generator, which allows users to build and customize reports.

As part of the 1.0 release, caAERS has been integrated with CTEP's AdEERS reporting system, allowing for the electronic submission of expedited reports from caAERS directly to AdEERS. The development of this feature required the encoding of over 80 rules to ensure the accuracy, completeness, and integrity of the data being submitted. This feature enables caAERS to serve as a front end, and eventually it is intended to replace the AdEERS system.

The SemanticBits team is currently engaged in the second phase of the caAERS project. Versions 1.5 and 2.0 are scheduled for release in Fall 2008 and early 2009. These new versions will feature additional functionality, to include enhanced usability, advanced routing and review of reports, and expanded integration with the AdEERS system. The system architecture for caAERS is shown in Figure A.5-5.

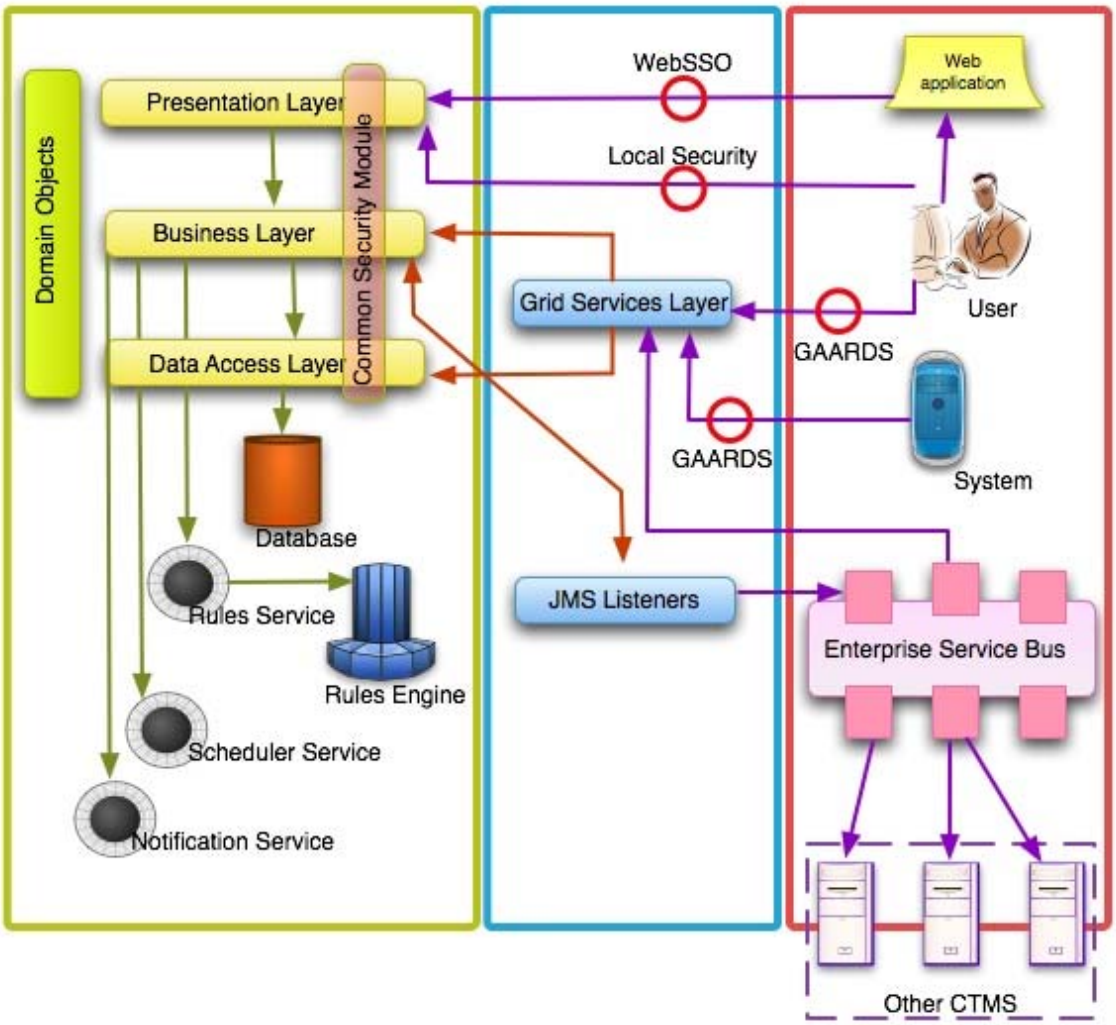



**Figure A.5-5 High-Level Architecture of caAERS**

**5.e.5. Plan for interoperability of PRO-CTCAE with caAERS, AdEERS, CDUS**

SemanticBits will develop the PRO-CTCAE system based on a well-defined domain model derived from the BRIDG model, which is a comprehensive domain analysis model for the biomedical/clinical research space. The BRIDG Model is a collaborative effort of stakeholders from the Clinical Data Interchange Standards Consortium (CDISC), the HL7 Regulated Clinical Research Information Management Technical Committee (RCRIM TC), NCI, and FDA to produce a shared view of the dynamic and static semantics that collectively define a shared domain-of-interest, ie, the domain of clinical and pre-clinical protocol-driven research and its associated regulatory artifacts.

SemanticBits has extensive experience developing systems using this approach. caAERS is one such application in which they successfully applied this approach. Using this approach in the development of the PRO-CTCAE system will greatly facilitate integration

with caAERS at the data level. caAERS supports multiple terminologies such as CTCAE (versions 2 and 3) and MedDRA (versions 9 and 10) for coding adverse events. Since the PRO-CTCAE is going to be an extension of the CTCAE, and caAERS supports CTCAE, it should be fairly straightforward to integrate the PRO-CTCAE system with caAERS.

One of the salient and popular features of caAERS is its seamless integration with AdEERS. caAERS can be used to submit CTEP 5 and 10 calendar day reports to AdEERS. To implement this functionality in caAERS, it was necessary to harmonize study and adverse event sections of the caAERS domain model with CTEP-ESYS. SemanticBits also implemented over 50 business rules that pertain to CTEP serious adverse event (SAE) reporting in caAERS. During this process they closely worked with the AdEERS development team, and will also be working with them to implement sending amended SAE reports and the CTEP 24-hour notifications to AdEERS. This clearly indicates a deep understanding of AdEERS and in general CTEP-EYS, which should greatly aid in integration efforts of the PRO-CTCAE with AdEERS and CDUS.

**5.e.6. Timeline for completion of PRO-CTCAE assessment system components**

The PRO-CTCAE system to be created by SemanticBits will include all of the system components ("deliverables") noted in Table A.5-14 below. The scope of the subcontract will include completion of all items noted in the description of each deliverable in keeping with the specified timeline in the month column (pertaining to the specific month the corresponding deliverable is due, following initiation of the contract). This schedule will allow for completion of version 1 of the PRO-CTCAE Assessment System in month 6 of the contract, and version 2 in month 12 of the contract, as required by the RFP. Please see the accompanying business proposal, containing the signed agreement with SemanticBits which recapitulates the below table.

**Table A.5-14. Schedule for Completion of PRO-CTCAE Assessment System Components**

| Deliverable | Description | Version | Month |
|---|---|---|---|
| **Publishing Tool/Form Builder** | Web-based publishing tool/form builder for creating multi-item electronic and paper PRO-CTCAE CRFs, for any given specific study. This interface will be similar in general design to the questionnaire builder used in caAERS. It will be user-friendly and not require specific technical background to use it, but rather will be designed for form creation by clinical investigators or clinical trial data managers. The Form Builder will construct study-specific questionnaires (CRFs) which include PRO-CTCAE items selected by clinical investigators. The form builder will include groupings of symptoms into pre-defined Item Sets, including "core" (mandatory) PRO-CTCAE items for inclusion in all studies, and "disease-specific" Item Sets. Clinical Investigator Users will therefore be able to build multi-item PRO-CTCAE CRFs based either on selection of Item Set(s) and/or selection of individual items from a checklist of all available items. The system will be designed for future capacity to include additional items/questionnaires (ie, non-PRO-CTCAE items) that can be loaded into the system by clinical investigators, and for full integration of PROMIS tools. The Form Builder will include customizable Rules for triggering automated alerts to clinical investigators for pre-specified grade thresholds entered by patients into the generated electronic CRFs, and rules for triggering required expedited reports through caAERS/AdEERS for pre-specified grade | Initial | 2 |
| | | Revision | 3 |
| | | Final | 4 |

| | | | |
|---|---|---|---|
| | thresholds. The Publishing Tool will include the ability for different Electronic PRO-CTCAE CRFs to be required at different visits during a specific study based on linkage to a Study Calendar and for question skip-pattern rules to be included. | | |
| **Electronic PRO-CTCAE CRFs** *(Electronic Patient Data Entry Interface)* | Generic online Data Entry Pages/Forms will be designed for electronic administration of PRO-CTCAE question items. These will be designed for completion of one item per online page. Multi-item electronic CRFs generated by the Publishing Tool will include one of these Data Entry Pages for each PRO-CTCAE item selected by clinical investigators for that CRF, as well as a drop-down list or checkbox list of the additional available PRO-CTCAE items not selected by the clinicians which patients can select in real-time ("Patient-Selected" items). This Deliverable includes design and functionality of Web pages in keeping with specifications provided by the Investigative team, and will be fully integrated with other PRO-CTCAE system elements. A design option for a popup touchscreen login virtual keyboard will be included in the design, for use with touchscreen computers. A link will be included to view/print longitudinal Patient-Level Reports of previously entered responses. | Initial | 3 |
| | | Revision | 4 |
| | | Final | 5 |
| **Paper PRO-CTCAE CRFs** *(Paper Patient Data Entry Interface)* | Design of a paper version of the PRO-CTCAE CRF. These forms will be generated as PDFs by the Clinical Administrative Interface as an optional backup data collection method for situations in which patients are unwilling/unable to use online data entry. The Paper CRFs will include the same items in the same order as the analogous Electronic CRFs for the same patient on the same visit day. The paper CRFs will be designed based on specifications provided by the Investigative Team. Multiple items will be included per page. | Initial | 4 |
| | | Revision | 5 |
| | | Final | 6 |
| **Clinical Administrative Interface** | A Web-based system for clinical investigators to monitor completion of forms, the schedule for completion of forms, and to generate Paper CRFs. User-level access will allow for permissions based on which patients a particular User should have access to in a trial, such as clinician-specific, clinic-specific, institution-specific, or study-specific. | Initial | 3 |
| | | Revision | 4 |
| | | Final | 5 |
| **Alert System** *(Rules-Based Notifications)* | An email-based alert system will be integrated into the system, based on Publishing Tool/Form Builder Rules (or modifiable by clinical investigators building CRFs with the Publishing Tool), to trigger messages to clinical investigators for pre-specified grade thresholds self-entered by patients. There will be an option in the Rules for these either to be warnings without required expedited reporting, or with expedited reporting linked to caAERS/AdEERS. The model for tracking these alerts and responses will be similar to caAERS. | Initial | 6 |
| | | Revision | 7 |
| | | Final | 8 |
| **Patient-Level Report** | A standard individual-patient-based longitudinal report generating function will be included that will be usable both for clinical trial data review, and for real-time clinical use. The design of these "Patient-Level Reports of previously entered responses" will be based on the specifications of the Investigative Team. It will be designed both for on-screen review, and for printing into reports. Reports will be printable by patients themselves via a link in the Electronic PRO-CTCAE CRFs (Electronic Patient Data Entry Interface), or by clinicians via the Clinical Administrative Interface. | Initial | 7 |
| | | Revision | 8 |
| | | Final | 9 |
| **Study-Level Queries** | Capacity for clinician administrators to query summary information across patients in a specific trial, with multiple options for data display including date; visit number; level of grading severity; specific symptom item(s). It will be | Initial | 7 |
| | | Revision | 8 |

| | | | |
|---|---|---|---|
| | designed both for on-screen review, and for printing reports. Access to specific patients will be by User level permissions. | Final | 9 |
| **Study Calendar Integration** | Linkage to the caBIG Study Calendar system will be included to determine the date that any given patient in a specific trial is scheduled to complete a PRO-CTCAE CRF. This will include the capability for variations in the items in CRFs for any given visit (ie, items 1-10 on Visit 1, and items 11-20 on visit 2, etc). An email alert system will be integrated from Study Calendar to inform clinical investigators when a patient has missed a scheduled CRF self-report due date. | Initial | 13 |
| | | Revision | 14 |
| | | Final | 15 |
| **Item Skipping Rules Capability** | Optional Rule in the Publishing Tool which allows for logic to be built into electronic CRFs to skip items based on responses to prior items (for example, “if responses to item #1 and item #2 are “none” then skip item #3). | Initial | 9 |
| | | Revision | 10 |
| | | Final | 11 |
| **Optional Item Recurrence** | Optional Rule in the Publishing Tool which specifies that if a patient opts to respond to a “Patient-Selected” item on the CRF from the drop-down list or checkbox menu, then that item will automatically recur the next time the patient responds to the CRF. | Initial | 10 |
| | | Revision | 11 |
| | | Final | 12 |
| **Refinement of PRO-CTCAE System** | Adjustments of the design and functionality of all elements of the PRO-CTCAE system will be continuous during the contract period, based on communications between Semantic Bits and the Investigative team and NCI staff, as well as based on the specific finding of usability testing and cognitive interviewing. | NA | 1-24 |
| **Interoperabilty, Compatibility, Data Standards, Security, Privacy, Open-Source, Intellectual Property** | All elements of the platform will be designed for full integration with each other and with caAERS/AdEERS; will include the capacity for full integration with PROMIS tools; will be open-source; will adhere to standards for data security and privacy and regulatory considerations in keeping with caBIG/NCICB and CDISC; will function on PC or Macintosh; will function on all commonly available browsers. Source code will be available at all times for review and will be delivered at completion of the project to the NCI and to the Investigative Team. The system will be designed for use on Web-enabled computers, but will be designed towards future integration of IVR and PDA data entry interfaces, and for capacity for downloadable software. No intellectual property rights will be retained by the company, designers, programmers, Investigators or any other involved individuals. | NA | 1-24 |
| **System and Database Maintenance** | During the contract period, for the purpose of studies being conducted to evaluate the PRO-CTCAE system by the Investigative Team, SemanticBits will maintain and keep the PRO-CTCAE Platform fully operational, and a study database repository will be created and maintained for storage and analysis of PRO-CTCAE data entered by patients via Electronic CRFs. Storage of this patient information will adhere to regulatory standards for security and privacy of PHI. | NA | 1-24 |
| **Install Manual** | SemanticBits will create the user manual for IT programmers to download/install and maintain local versions of the PRO-CTCAE software. | NA | 24 |

**5.e.7. Technical Approach**

The technical approach and proposed technical solution will be driven by various functional and non-functional requirements, which includes the need to interoperate with other systems such as the caBIG Patient Study Calendar. The SemanticBits team has analyzed the vision of the project in detail with Dr. Basch, Mr. Shouery, and the investigative team, and has identified a number of specific architectural considerations for his project, as listed below.

*Architectural Considerations*. Architectural considerations specific to this project are listed below:

- *User-friendly GUI:* A user-friendly, highly interactive and intuitive Web interface is required for the PRO-CTCAE system. It should not require specific technical background to use it, but rather will be designed for study subjects, clinical investigators and clinical trial data managers.
- *Extensible:* The PRO-CTCAE application should be developed in such a way that it is extensible to meet the future requirements. It should include well-designed APIs, which should be open for extension and closed for a modification, which is commonly referred to as the "Open-Close" principle.
- *Secure:* The PRO-CTCAE application will handle sensitive patient data, therefore it is imperative to secure the data exposed by the PRO-CTCAE system with flexible access control mechanisms.
- *Layered Architecture:* The PRO-CTCAE application should be developed using a layered architecture consisting of presentation, business logic, data access, and technical services layers. Each layer must expose a well-designed interface through which the layer immediately above it interacts with it.
- *Interface based design:* The PRO-CTCAE application should be developed using the core object-oriented analysis and design (OOAD) principle of Interface-based design. The application should expose a well-defined set of interfaces. These interfaces will not only facilitate communication between PRO-CTCAE components, but also greatly streamline integration with external systems. This will lead to a loose coupling between the components of PRO-CTCAE, which greatly enhance reusability and maintainability.
- *Regulatory Compliance and Auditing:* The PRO-CTCAE application should meet regulatory requirements as determined by use cases. To achieve this, a robust auditing and logging mechanism should be incorporated in the PRO-CTCAE system.
- *Configurable and Customizable:* The PRO-CTCAE application should be configurable and customizable. This will allow adopters to customize the PRO-CTCAE to meet their UI requirements and use their organization specific look and feel.

*BRIDG Harmonization.* The domain model for the PRO-CTCAE system should be harmonized with the BRIDG model. The domain objects in the PRO-CTCAE system should be derived from the analysis classes defined in the BRIDG model to the extent possible:

- *Message based integration with external systems:* The PRO-CTCAE system is required to communicate with other external systems. Some information (e.g. patient visit schedule from local study calendar system) may be required in the PRO-CTCAE system. To achieve this goal, the PRO-CTCAE system should be able to communicate with external systems using messages, rather than point-to-point API based integration.
- *Rules-Engine:* There are various use cases where the PRO-CTCAE system needs to make a decision based on data entered into the system by the user. The subsequent actions based on this decision may vary also vary based on different criteria. Including this volatile logic in the application may not be a good decision from a software design standpoint. This should be handled by leveraging a Rules-Engine just like it is being implemented in the caAERS application.
- *Notification System (Use Case Model):* Numbers of use cases warrant a sophisticated rules-based notification system. The use case view models the functional requirements of the system by showing the various users (actors) and the actions they perform while using the system. It is important to start the inception phase with a solid idea of the use case model so that it can be circulated to the stakeholders and serve as a basis for further requirements gathering. Below is the use case model that reflects our current conceptualization of the system requirements (Figure A.5-6).

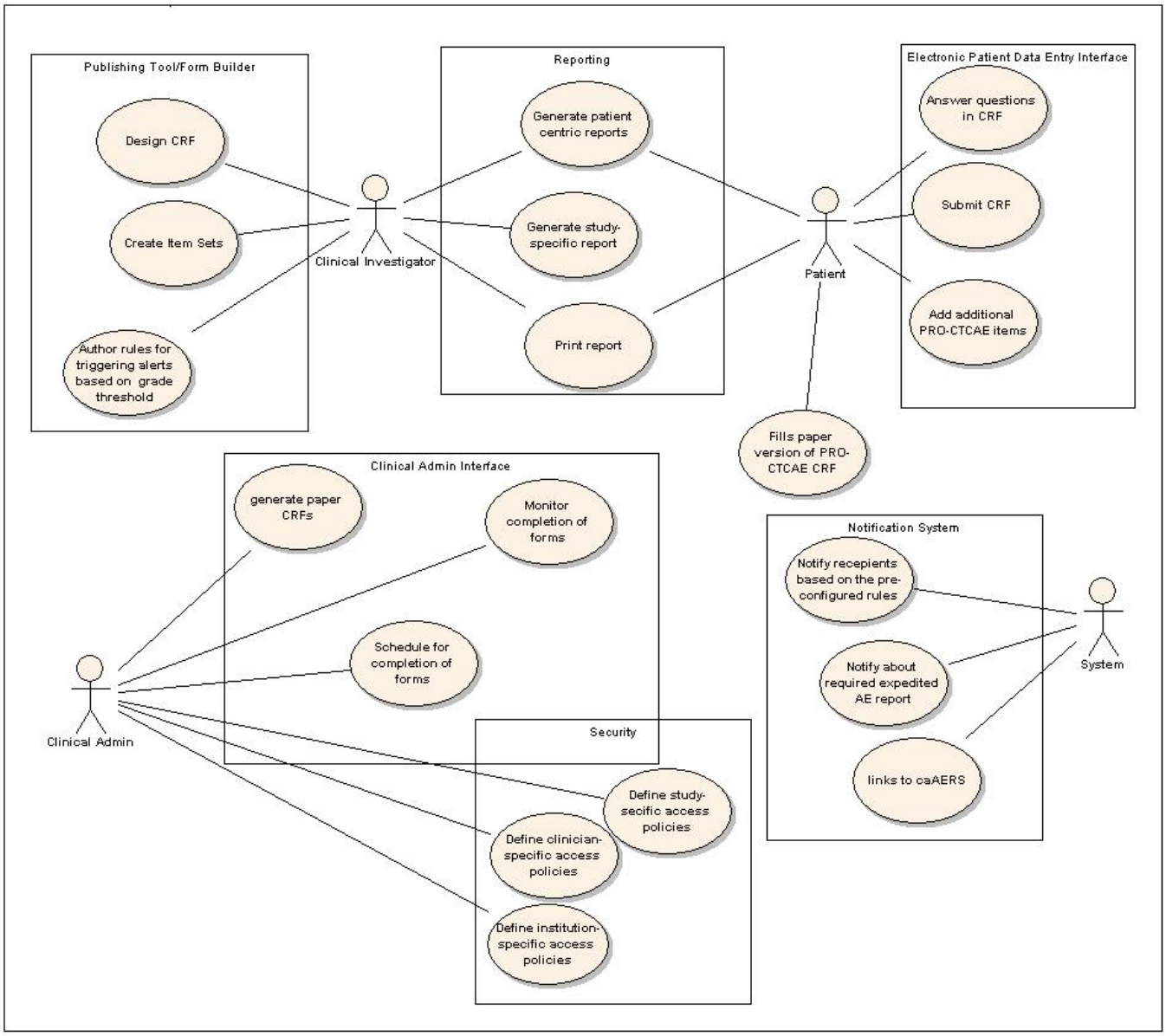



**Figure A.5-6. PRO-CTCAE Use Case Model**

*PRO-CTCAE Modules*. As shown above in the table of deliverables and timeline, the main modules in the PRO-CTCAE system are:

- *Publishing tool/Form builder:* used by clinical investigators to design PRO-CTCAE CRFs.
- *Simple clinical admin interface:* used by clinical investigators to monitor completion of forms, the schedule for completion of forms, and to generate paper CRFs.

- *Simple patient interface (electronic self-report, interviewer, or paper):* used by patients to answer PRO-CTCAE CRFs.
- *Query/reporting interface:* Used by clinical investigators and patients to generate appropriate reports.
- *Rules-based notification system:* Used for automated email alerts, triggereing expedited reporting forms, or triggering additional items/questionnaires; listing of non-responders on scheduled days.

*Web Application Architecture*. The SemanticBits team proposes to develop the PRO-CTCAE system using a layered architecture. SemanticBits has successfully implemented several software systems (including caAERS, C3PR, and PSC) using this approach. Figure A.5-7, below, illustrates the various logical layers in the proposed architecture:

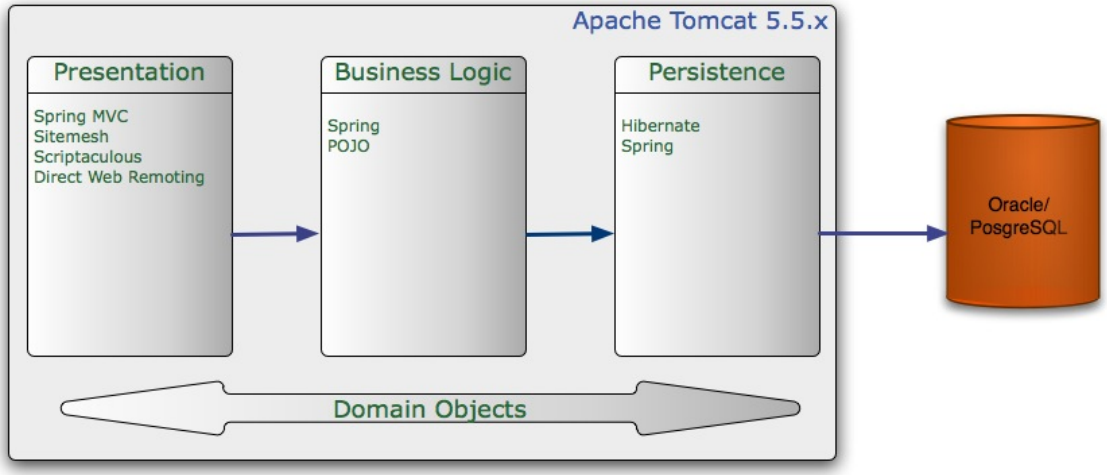



**Figure A.5-7. Multi-layered Web application architecture**

*Presentation/UI Layer*. The presentation layer contains Web components such as Java Server Pages (JSP) and Java servlets that can generate dynamic HTML content. The components in this layer are responsible for providing site navigation, workflow, user authentication, authorization, and access to the business logic tier. The SemanticBits team will use the Spring MVC framework for developing the presentation tier of the PRO-CTCAE system. Wherever necessary and appropriate, to improve the user experience, they will use Asynchronous JavaScript and XML (AJAX) technology. They will extensively work with the end users (adopters) during the user interface development, as described in the Usability Testing (Task 6) section of this proposal.

*Domain/Business Layer*. The business layer consists of domain object classes and a set of business services, each of which contains a number of coarse-grained methods that implement a group of related use cases. Some of the key advantages of developing business services are:

- It provides a simpler interface to the clients by hiding all the complex interactions between domain objects. This leads to better manageability, centralization of interactions (responsibility), greater flexibility, and greater ability to cope with changes.
- It avoids exposing the underlying business objects directly to the client to minimize the tight coupling between layers.
- It provides a uniform coarse-grained service layer to separate business object implementation from business service abstraction.
- It improves the performance by reducing the number of over the network calls since the service methods are coarse grained.

The absence of a well-designed business service layer may lead to the following issues:

- Tight coupling that can lead to direct dependence between clients and domain objects.
- Too many method invocations between the client and server leading to network performance issues

Data Access Layer

The Data Access layer contains classes that encapsulate logic for accessing the underlying database. Data Access Object (DAO), which is one of the most important and commonly used core J2EE patterns, will be used for implementing data access logic. The main intent of the DAO pattern is to separate business logic from persistence code. The DAO pattern introduces an interface for defining persistence operations that the business/domain layer can access, hiding the actual implementation. The SemanticBits team will use Hibernate for implementing O/R mapping, which involves mapping objects in the model to database tables.

**5.e.8. Development approach**

*Agile Development*. SemanticBits will adhere to agile development practices. In this approach, development will be organized into a series of short, fixed-length mini-projects called iterations. The outcome of each iteration is a tested, integrated, and executable partial system. Each iteration includes its own requirements analysis, design, implementation, adoption, and testing activities. Some of the other core practices include:

- Tackle high-risk and high-value issues in early iterations;
- Continuously engage stakeholders for evaluation, feedback, and requirements to develop a system that closely meets their needs;

- Build a cohesive, core architecture in early iterations, which in this case is the elaboration phase - in an adoption-oriented project this also includes deployment of the core architecture of the different tools being developed;
- Continuously verify quality; test early, often and realistically;
- Manage requirements carefully;
- Practice change request and configuration management.

*Disciplines*. Multidisciplinary teamwork is one of the success criteria in both user-centered approaches and agile methods. It is considered to be essential for the Agile Unified Process (AUP) in order to meet project objectives and to ensure stakeholder buy-in to both process and results (Figure A.5-8). This can be divided across the sub-disciplines of the AUP:

- *Requirements*: the overarching goal of this project is to ensure that the applications are generally usable by the stakeholder community, so it is important to make sure that the stakeholder use cases and requirements generalize to meet the needs of the larger community;
- *Analysis and design*: System requirements will be analyzed to create the domain and design models for the system;
- *Implementation*: Implementation of the system will rely heavily upon existing, well established software development methodologies;
- *Testing*: In line with AUP, testing will be performed early and often, and will inform all aspects of each iteration by both the developers and stakeholders;
- *Change and Configuration Management*: The overarching goal of this project is adoption of existing tools, so changes to them will be reviewed carefully before a final decision is made;
- *Project Management*: SemanticBits will leverage a light-weight management approach focused on communication between all teams and time-boxed tasking;
- *Environment*: Development teams will maintain a development, test, and production environment. The user acceptance testers will maintain test and production environments.

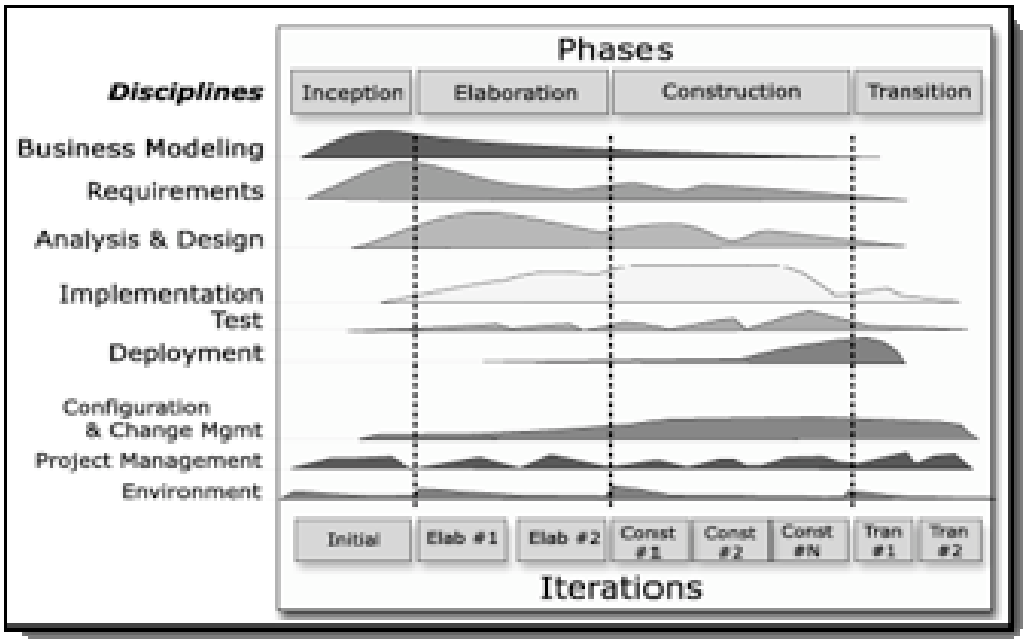



**Figure A.5-8. Agile Unified Process (AUP)**

*Phases*. The project will be developed in two major phases with the Agile/UPF sub-phases divided amongst them:

- Phase 1: document approach and develop the software system
  - Inception
  - Elaboration
  - Construction
  - Transition

- Phase 2: evaluation studies and creation of training materials (in collaboration with investigators, described separately in Task 9 of this proposal)
  - Inception/Elaboration
  - Construction
  - Transition

Some of the main activities in the Inception phase will include:

- Establish the project scope and boundary conditions
- Outline the use cases and key requirements that will drive the design tradeoffs
- Outline one or more candidate architectures
- Identify risks

The second project phase has a combined Inception and Elaboration Phase because its nature is less oriented towards development and the scope should be well established by that time. Some of the main activities in the Elaboration phase will include:

- The core software architecture with the most risk will be developed and tested based on a subset of use cases
- The majority (close to 95% percent) of the requirements will be discovered but not completely fleshed out
- Close to 30% of the architecturally significant, high business value and high risk use cases will be described completely and implemented
- Major risks will be mitigated or retired

Some of the best practices that the team will adopt during this sub-phase are:

- Performing short time-boxed and risk-driven iterations
- Beginning programming early in the phase
- Adaptively designing, implementing, and testing the core and risky parts of the architecture
- Testing early, often, and realistically
- Adapting based on feedback from tests, users, and developers
- Writing a subset of use cases and other non-functional requirements in detail through a series of workshops, once per elaboration iteration

During the Construction sub-phase, all remaining components and application features will be developed and integrated into the product, and all features will be thoroughly tested. This includes the following tasks:

- Continue to develop and refine the use case specification and use case realization document.
- Continue to develop and refine the requirements specification
- Continue to develop and refine the domain model and data model
- Continue to develop and refine the design model
- Continue to develop and refine the codebase (presentation, business logic, and data access layers). All systems developed in the elaboration phase will iteratively be enhanced and refined.
- Continue to refine and update all developer testing documentation
- Develop release notes that contain a listing of new features and functionalities, known bugs and their status
- Develop an installation guide that contains detailed technical instructions about the software system installation. It will also contain various configurations that are supported.
- Develop a system administration guide which contains detailed instructions on how to update and maintain the software system.

The purpose of the Transition sub-phase is to transition the software product to the sites and ultimately the user community. This will begin when a baseline is mature enough to be deployed in the end-user domain. This typically requires that the system has been completed to an acceptable level of quality and that user documentation is available so that the transition to the user will provide positive results for all parties. This sub-phase will include:

- Rollout of the systems to the sites
- "Beta testing" to validate the new system against user expectations
- User acceptance testing (described separately, in the Task 6 section of this proposal)

*Stakeholder Involvement*. A large number of stakeholders from different organizations will be involved in the project. They fall into the following categories:

- Clinical and research organizations (e.g. cancer centers)
- Industry organizations (e.g. pharmaceutical companies)
- NCI Program Staff
- NCI sponsored clinical trials networks
- NCI Symptom Management and Quality of Life (SxQOL) Steering Committee
- CTCAE developers
- Adverse event reporting systems developers (as needed beyond SemanticBits)

Managing stakeholder involvement is critical throughout the course of the project, particularly so in the processes of requirements gathering and analysis. The two most critical factors of success in the area of Stakeholder Involvement are 1) Making sure communication is open, engaging, and efficient; and 2) balancing the need for comprehensiveness and consensus with a bias for action that will enable for on-time delivery. The plan for communications with stakeholders and NCI staff is described separately in this proposal.

SemanticBits has experience engaging a wide variety of stakeholders and subject matter experts. In the caAERS project, they have engaged registrars, site coordinators, and study coordinators from Mayo Clinical, Wake Forest Cancer Center, and Duke Cancer Center. In this project, the Project Manager planned and oversaw interactions between the stakeholders and the development team, the business analyst translated business requirements into concrete use cases, and the architect turned use cases into implementable infrastructure models. In addition, SemanticBits is the primary developer of the Cancer Translational Research Informatics Platform (caTRIP), whereby we engaged clinicians and pathologists from Duke Cancer Center and Duke Health Systems. In this project, they took abstract, high-level requirements for outcomes analysis and built a distributed system for integrating disparate datasets. They were able to effectively apply complex grid technologies to solve a real clinical problem while effectively keeping the technology away from users and “under the hood.”

The SemanticBits project team will follow a similar approach to combine broad requests for information/feedback with involvement of specific individual stakeholders who have shown the capability and willingness to contribute meaningfully to solving challenges in the PRO-CTCAE problem domain. In all cases, the team will strive to produce correspondence that includes specific action items and a timeline within which those items should be fulfilled. The stakeholder communications approach will allow measurable progress while still enabling open, stimulating discussion of the problem domain and innovative solutions with the stakeholders.

*System Refinement and User Acceptance Testing*. A key reason that SemanticBits will adopt an Agile methodology is that it will allow them to put a prototype version of the system in the hands of stakeholders very early in the development lifecycle. SemanticBits will therefore deliver a prototype system to the PI and investigative team to verify that the developers and stakeholders agree on the solution. This will happen at the end of each iteration of the Elaboration phase. During the Construction phase, this prototype system will be maintained and updated as needed, such that the stakeholders and developers can review progress and changes, thereby mitigating risk and fostering involvement and awareness.

User acceptance/usability testing will be key facets in the development process. This process, under the leadership of Dr. Abernethy at Duke, is described separately in Task 6. Key stakeholders from participating organizations will be identified to perform user acceptance testing. The ideal candidates will be a cross section of users from the target population of end users. Therefore, we will engage patients and medical staff. The process will be to place the application in the hands of users and ask them to perform specific tasks and workflows within the system. During this time, a SemanticBits business analyst will observe how the users function within these realistic scenarios and document issues that the users have with the application. This feedback will be fed directly into the software development process on an iterative basis.

*Leverage Existing Work*. One of the basic tenants of the SemanticBits software development approach is reuse. In this vein, they will leverage work already done in this area at MSKCC and Duke, both of which have existing systems for performing patient-reported outcomes. They will spend time in the Elaboration sub-phase analyzing the existing systems in light of the scope and requirements defined from the Inception sub-phase. This should give the development team a solid foundation to begin development of the system, which will likely consist of the following modules:

- Publishing tool: generates electronic or paper [PDF] PRO-CTCAE CRFs – consider scannable forms
- Simple staff/clinician interface
- Simple patient interface (electronic self-report, interviewer, or paper)

- Query/reporting function; automated email alerts; listing of non-responders on scheduled days

**5.e.9. Plan to assure security, privacy, regulatory considerations**

Systems dealing with sensitive patient health data are required to comply with CFR 21 Part 11, which is a specification from the FDA. It contains comprehensive rules that govern collection and management of electronic records. Compliance to such regulations has introduced several challenges such as rigorous security, privacy, data integrity and audit standards including the requirement to maintain electronic records for years or even decades. Security and privacy are closely related technologies, although there are important differences that are important to appreciate in order to design new systems that address both. *Privacy* is about informational self-determination—the ability to decide what information about one goes where. *Security* offers the ability to be confident that those decisions are respected.

Since the PRO-CTCAE will handle sensitive patient information, security must figure prominently in the systems architecture so that this information remains safe. The planned architecture will provide components to protect against the following common security threats.

- *Masquerading:* Only legitimate "actors" (users or systems) should be able to authenticate to the system.
- *Improper Access:* Once authenticated, actors should be able to access and manipulate data only according to the configured security policy.
- *Repudiation:* The actions of each actor must be audited in such a way that they cannot be repudiated.
- *Eavesdropping:* Communication among components of the system and external systems should not be able to be interpreted if they are intercepted.
- *Message Integrity:* Intercepted messages should not be able to be manipulated (or substituted) before being forwarded to the intended recipient.

An effective authentication approach is necessary to ensure that only legitimate users gain access to the system. A fine-grained authorization mechanism is also necessary to ensure that authenticated users can access and manipulate data according to the security policy. An auditing and logging component is needed to provide non-repudiation. Secure communication protocols are needed to prevent eavesdropping and ensure message integrity. SemanticBits has demonstrated and implemented this type of security addressing all of the above issues. For authentication and authorization, the SemanticBits team will be using Common Security Module (CSM) in conjunction with Acegi security. The SemanticBits team has vast experience in using PKI infrastructure and has used the same extensively in caGrid, and will use the same approach for the PRO-CTCAE. To support complete auditing and logging features, we propose the use of the NCICB's

Common Logging Module (CLM). CLM provides an object-state logging feature, which can be used to capture data snap-shots prior to modification.

### 5.e.10. Plan for Open-Source Approach

All software developed in this effort will be published as open-source under a non-viral license (eg, LGPL). SemanticBits has extensive experience developing open-source software for the caBIG program administered by CBIIT/NCI. All software we have developed thus far has followed the open source principles of:

- The software is to be freely available to biomedical researchers, educators, and institutions in the non-profit sector, such as institutions of education, research institutions, and government laboratories.
- The terms of software availability are to permit the commercialization of enhanced or customized versions of the software, or incorporation of the software or components of it into other software packages.
- The terms of software availability are to include the ability of research institutions outside the offeror and affiliated organizations to modify the source code and to share modifications with other colleagues.

Furthermore, as part of the software package, we plan to document and publish openly all data standards and programming interface specifications that are leveraged and exposed by the system. This will greatly facilitate adoption of the software into existing clinical and research settings that rely on existing systems for storage and transmittal of data.

### 5.e.11. Specific Deliverables to the NCI, and Timeline for Task 5

In an above section, we have provided a detailed timeline for completion of individual components of the PRO-CTCAE platform (assessment system). That timeline is largely intended for project management during the contract period, and contractual purposes with SemanticBits. We anticipate that NCI staff will be closely involved with the software development process, and hence will be able to review the system at various timepoints throughout the contract. Below, we list the specific deliverables to NCI as required by the NCI, which are broader than the level of detail provided above (Table A.5-15).

**Table A.5-15. Specific Deliverables and Timeline for Task 5**

| SPECIFIC DELIVERABLES | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| Completion of online security awareness training by all relevant project participants | Month 1 |
| PRO-CTCAE Assessment System, version 1 | Month 6 |
| PRO-CTCAE Assessment System, version 2 | Month 12 |
| Source Code/Object Code/System Delivery | Month 24 |

### *5.f. Substantive Approach to Task 6: Conduct Usability Testing*

| | |
|---|---|
| Task 6: | Conduct Usability Testing of the PRO-CTCAE Platform (Electronic and Paper) in Diverse Patient Populations and with Clinical Staff Members to Ensure Intuitive Navigation/Usability and Evaluate for Respondent Fatigue toward Platform Refinement. |
| Years of Contract: | Year 1 and 2 |
| Task Leader: | Dr. Amy Abernethy |
| Deputy Task Leader: | Dr. Deb Schrag and Mr. Marwan Shouery |

**5.f.1. Overview**

Usability of the PRO-CTCAE system is broadly defined as the capacity of the system to allow users to carry out their tasks safely, effectively, efficiently and enjoyably (http://www.usability.gov/basics/whatusa.html, accessed 6/1/08). We plan to create explicit usability testing protocols to ensure that the target users – patients from diverse backgrounds with varying educational levels, computer experience, and physical illness and clinical research staff (eg, nurses, research assistants, physicians) – can conduct the expected tasks. Measures will focus on: (1) *Ease of learning* - How fast can a user who has never seen the user interface before learn it sufficiently well to accomplish basic tasks?; (2) *Efficiency of use* - Once an experienced user has learned to use the system, how fast can he or she accomplish tasks?; (3) *Memorability* - If a user has used the system before, can he or she remember enough to use it effectively the next time or does the user have to start over again learning everything?; (4) *Error frequency and severity* - How often do users make errors while using the system, how serious are these errors, and how do users recover from these errors?; and, (5) *Subjective satisfaction* - How much does the user like using the system?

**5.f.2. Usability testing study design**

During Task 5 of this proposal (Platform Development), Mr. Chilukuri and the vendor, SemanticBits, will design the blueprint for PRO-CTCAE platform, develop user interfaces, create a publishing tool, and adapt the system platform based on stakeholder feedback and usability testing. Expected deliverables include the blueprint for the PRO-CTCAE electronic platform data flow, the system design, and then a functional beta open-source PRO-CTCAE platform.

Usability testing is key to this process. In conjunction with Dr. Basch, the usability task leader, Dr. Abernethy, and Deputy Task Leaders, Dr. Schrag and Mr. Shouery will consult with Mr. Chilukuri and SemanticBits frequently during the system design phases to understand and inform the platform and system design characteristics, recommend modifications as needed, and plan usability testing of the beta model.

Once the beta model is prepared, full-scale usability testing will proceed with 2 main iterations planned. All usability tests will be conducted in the environment where the

system is expected to be used most frequently, i.e., outpatient oncology clinics. Inpatient tests are also planned, to ensure that the system accommodates sick individuals in the hospital. The goal is to carefully create realistic scenarios that replicate future uses of the system in order to observe *in situ* interactions and identify iterative opportunities for improvement. Participants will be recruited from MSKCC, M.D. Anderson, Duke, Mayo, and Dana-Farber and their respective satellite clinics, in order to obtain a broad representation of academic and community-based users as well as diverse oncology-based disciplines including medical oncology and radiation oncology.

Two groups of users will be included: clinical research staff and patients. We have planned for 54 clinical research staff evaluations and 108 patient evaluations, divided into the two rounds of usability tests. Among clinical research staff, each round will include approximately 10 nurses, 10 clinical research assistants, and 10 physician investigators. Among patients, each round will minimally include 60 patients with at least 10 individuals who are over the age of 75 years, at least 10 individuals who have not completed a high school education, at least 10 individuals who are non-white, at least 10 who are inpatients or severely ill (ECOG $\geq$2), and at least 30 individuals who have metastatic disease. The same individual may represent multiple characteristics. After 40 patients have been involved in a round of usability testing we will enrich the group with more patients who have a specific characteristic (eg, low education) based upon prior experience in that usability testing round.

Usability testing will comprise three main components: (1) observation of users interacting with the system coupled with think-aloud protocols and extensive field notes; (2) user feedback using semi-structured interviews and short surveys; and, (3) Web analytics including tracking for different Web pages, with clickstream analysis. It is unlikely that eye tracking protocols and video-based data collection of screen movements will be necessary for this project, although both of these options are available if needed. Field observations, interviews, and survey questions will capture information from all five critical usability domains: ease of learning, efficiency of use, memorability, error frequency and severity, and subjective satisfaction. Web analytics will provide supplemental information on efficiency, memorability, and errors.

Drs. Abernethy and Schrag, together with Mr. Shouery (at MSKCC) and an experienced staff trained in usability testing protocols from Duke University Medical Center, will work together with Mr. Chilukuri and programmers from SemanticBits to develop the think-aloud protocol and to ensure that the Web tracking software is incorporated into the site. A formal research protocol will be developed for Task 6, which will be submitted to the relevant IRBs for approval. Research staff at each site will be trained by Dr. Abernethy and the usability testing team from Duke University Medical Center in the protocol, expectations, and data collection processes. The importance of field notes and extensive written data collection will be reviewed. Field notes can be handwritten or

dictated; all field notes will be transferred to typewritten text. Participants who meet eligibility criteria will be recruited at each site in accordance with best clinical research practice as outlined in the Human Subjects section of this proposal (Section D); informed consent will be obtained. Site research personnel will conduct the usability according to the defined protocol. Data will be collected on case report forms specifically designed for the usability testing. If site research staff does not feel comfortable conducting these tests, research staff from Duke University Medical Center can guide the process through video conferencing or travel for onsite training and participant accrual.

**5.f.3. Usability testing analysis**

Analysis corresponds to the testing round (1 or 2), type of usability testing, and the participant characteristics.

Observational data obtained through field testing is the most critical information. Typed field notes generated through think-aloud protocols and observation will be carefully reviewed and coded according to themes using a codebook analysis approach (MacQueen, *Cultural Anthro Meth* 1998). The codebook structure will include the code, a basic definition, a full definition, guidelines for using the code, and examples. Codes will be categorized into themes and sub-themes, which will minimally include the five usability domains: ease of learning, efficiency of use, memorability, error frequency and severity, and subjective satisfaction. Text will be reviewed by research staff at Duke University Medical Center trained in codebook analysis and codes will be applied. A relational database will be used to facilitate developing the code and applying it to text. The codebook analysis will provide a qualitative summary of the types of problems encountered and a quantitative summary of the aggregate number of times a particular problem was encountered.

User feedback will be solicited through semi-structured interviews and short surveys. Interview data will be added to the field notes and incorporated in the codebook analyses. Survey data will be compiled using descriptive statistics, divided into the five usability domains (ease of learning, efficiency of use, memorability, error frequency and severity, and subjective satisfaction).

Web analytics and clickstream analyses will supplement the field notes and user response data, providing supplemental information on efficiency, memorability, and errors. Benchmarking information that will be tracked will include but is not limited to: (1) total time required to use the PRO-CTCAE system for new and repeat users; (2) aspects of the system (eg, Web pages) that create unnecessary delays; (3) order of pages viewed and number of times the user needs to “go back”; and, (4) number of missed or skipped items.

Results will be reported in aggregate and by user type, ensuring that the different categories of users are equally capable of interacting with the system as intended.

Summaries of the information captured in the first round of usability testing will be reviewed with Mr. Chilukuri and the programmers from SemanticBits to inform iterative improvements in the system. The modified PRO-CTCAE system will be subjected to a second round of usability testing to ensure that first round items were addressed, and that further development is not needed.

### 5.f.4. Specific deliverables and timeline

The schedule of deliverables to NCI staff is detailed in Table A.5-16.

**Table A.5-16. Specific Deliverable and Timeline for Task 6**

| SPECIFIC DELIVERABLE | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| Summary reports of usability testing | Month 12 |
| Documentation of resulting system modifications | Month 12 |
| Refined PRO-CTCAE platform based on results | Month 12 |

### *5.g. Substantive Approach to Task 7: Evaluate Measurement Properties*

| | |
|---|---|
| Task 7: | Systematically Evaluate in Diverse Patient Populations the Validity, Reliability, and Sensitivity of the PRO-CTCAE Items, as well as the Clinical Significance of Score Changes for Each Item, and Appropriate Recall Period |
| Years of Contract: | Year 2 |
| Task Leaders: | Dr. Charles Cleeland, Dr. Jeff Sloan |
| Deputy Task Leader: | Dr. Tito Mendoza |

**5.g.1. Overview**

As described previously in the "Substantive Approach to Task 2: Item Development" section, an ideal symptom measure should demonstrate adequate validity, reliability, and sensitivity, with limited variability of interpretation of responses across diverse patient populations (Scientific Advisory Committee of the Medical Outcomes Trust, 2002). The recall period for items should be well understood, as should the clinical relevance of specific score changes in populations and contexts of interest.

The investigators bring extensive experience in instrument development and evaluation of these properties, with a deep background in the applicable analytic methods. We have previously evaluated the measurement properties of instruments and item banks, which include many of the same symptoms in which we are interested for the PRO-CTCAE (eg, MDASI, BPI, BFI, STAR), and therefore the study proposed in this section capitalizes on prior research observations and collected data. In addition, the investigators bring access to a large and diverse patient population across our institutions, which will be available for timely accrual. The investigators also bring access to highly developed clinical research infrastructures, which will ensure efficiency and comprehensiveness of data collection.

We propose to conduct this study during year 2 of the contract period (ie, phase II of the RFP), using PRO-CTCAE items developed and refined in Tasks 2, 3, and 4, administered via the technology platform created and refined in Tasks 5 and 6.

**5.g.2. Access to patients and target patient populations**

*Selection of Nontrial Settings for Initial Development Work.* We propose to conduct the studies described in this section in the routine care setting at four participating institutions: Dana-Farber, Duke, MDACC, and MSKCC (in both tertiary hospital and community satellite clinic sites). We have opted to develop the PRO-CTCAE in these settings because we are confident that we can obtain IRB review and accrue patients quickly at our home institutions. Furthermore, this network will provide access to diverse patient populations. Although the ultimate goal of this program is to generate a system for use in the NCI cooperative group setting, we believe that it is not realistic to do the initial development work within the cooperative groups. This is primarily because of the aggressive timeline of the proposed work and the pace of protocol review and implementation in the cooperative groups. Notably, Dr. Basch currently has an open

CALGB trial, 70501, which is exactly in the area of this proposal (see Appendix). In CALGB 70501, patients enrolled in selected NCI-sponsored treatment trials self-report their own adverse symptoms via computer at scheduled visits using MSKCC preliminary adaptations of selected CTCAE items (ie, STAR Items). This protocol took almost 3 years to move through the cooperative group mechanism before opening, and therefore we aim to wait until the PRO-CTCAE is already well developed before assessing feasibility in the cooperative groups. That said, because CALGB 70501 is already open, a fairly straightforward amendment will allow us to exchange the new PRO-CTCAE items for the current STAR items, and thus we will be able to accelerate assessment in the cooperative groups, as described later in the section "Substantive Approach to Task 8: Feasibility Design." Furthermore, we plan to open this study in other cooperative groups as well (see attached letters from representatives of CALGB, RTOG, NCCTG, and ECOG).

*Access to Patients.* An advantage of our investigative team is that we bring access to large numbers of diverse patient populations across our respective institutions in both tertiary hospital settings and community satellite clinics. There are more than 2 million cancer outpatient visits each year at our combined institutions, with more than 20,000 patients enrolled annually in cancer treatment trials at our sites. As a result, we will have access to more than adequate patient samples to rapidly conduct the proposed studies and will also have access to patients who vary in characteristics by age, sex, ethnicity, educational level, and geography (see attached targeted enrollment tables for each participating institution).

*Selection of Disease-Specific Patient Groups.* For the evaluations of measurement properties, we have identified patient disease groups that span the spectrum of symptom burden to ensure adequate representation of individuals who have experienced (or are currently experiencing) the symptoms of interest (shown in Table A.5-17). In the prior section of this application, "Approach to Task 2: Item Development," we mapped each of the preliminary candidate CTCAE symptom items to one or more of these specific diagnoses based on prevalence data. These populations have also been selected to include some patients whose symptoms are most likely due to treatment effect rather than to burden of disease (eg, breast cancer receiving adjuvant chemotherapy), as well as those for whom the etiology of symptoms is likely to be less clearly attributable to treatment versus the underlying tumor. Patients in some selected groups will be receiving primary therapy with curative intent (eg, myeloma/lymphoma, adjuvant breast cancer) and palliative intent (eg, metastatic lung or prostate cancer). For evaluations of test-retest reliability, it will be essential to include patients who return daily for treatment (head and neck/gastroesophageal carcinoma receiving chemo-radiotherapy).

**Table A.5-17. Planned Patient Populations for Assessments of Measurement Properties**

| |
|---|
| Breast cancer receiving adjuvant chemotherapy |
| Metastatic prostate or bladder cancer receiving chemotherapy |
| Head and neck or gastroesophageal carcinoma receiving daily chemo-radio therapy |
| Metastatic lung cancer receiving chemotherapy |
| Metastatic colorectal cancer receiving chemotherapy |
| Lymphoma/myeloma receiving chemotherapy |

This combined population overall is representative of many of the patients accrued into adult cancer treatment trials in the cooperative group setting. Most importantly, the treatment regimens used to treat the cancers above have been carefully selected to encompass most of the important clinical symptoms encountered in clinical oncology. The conditions noted above are treated with a diverse array of chemotherapy drugs and several involve radiation treatments as well. We recognize that some disease diagnoses are not included in which patients may experience different symptoms from the above included groups. After the PRO-CTCAE is fully developed, evaluation in additional groups of patients will of course be warranted and can be conducted by us or others. Out team has the capacity to easily expand evaluation to additional groups of patients given supplemental resources. We prefer to develop the PRO-CTCAE system on a well specified group of patients so as to ensure that we can optimally compare responses to particular items based on the specific disease/treatment scenario. By including adjuvant breast cancer, we are certain to include patients receiving anthracycline-based therapy. In this way, we know we will have the opportunity to characterize PROs for nausea. For example, we know we will have the opportunity to characterize PROs for neuropathy because taxanes are used for adjuvant breast cancer and oxaliplatin is used for metastatic colon cancer and each frequently causes neuropathy.

*Lack of Inclusion of Children.* We have not included children in the initial studies of the PRO-CTCAE. As discussed previously in the “Substantive Approach to Task 3: Cultural Literacy” and in the Human Subjects section of this application, we propose to design items with the intention of initial administration and evaluation in adults, with future planned adaptation for use in children and their proxies. Toward this, Dr. Vish Viswanath’s Core at the Dana-Farber will evaluate items to ensure adaptability for pediatric use, and in addition, two experts in symptom and adverse event evaluation in children are included as advisers on this proposal (Dr. Joanne Wolfe and Dr. Steve Joffe; see Letters of Support). Evaluation and administration of PRO items in young children pose specific challenges, which the investigators recognize are essential to address, and thus such work is planned once the model has been proven in adult patients (Langer et al, J Clin Epidemiol 2008; Varni et al, Health Qual Life Outcomes 2007).

**5.g.3. Proposed study to evaluate the validity, reliability, sensitivity, clinical significance of score changes, and recall period of PRO-CTCAE items**

*Overview.* We propose a study to assess the newly developed PRO-CTCAE items in a diverse patient population that represents the spectrum of disease burden and symptoms in cancer care. We expect that this study will provide supportive data for the psychometric integrity of the PRO-CTCAE items and will also provide insights toward possible modifications of individual items. 270 patients total (approximately 50 patients from each disease group described above in Table A.5-17) will be enrolled across the 5 institutions (we will target balance in accrual across the five institutions within each disease group). All patients will be asked to complete all Category I and II PRO-CTCAE items at multiple time points as well as other validated questionnaires (time points and questionnaires are described in the following sections). Clinical data including ECOG performance status will also be obtained for each patient over the course of the study via case report forms (CRFs). For each symptom item, we will correlate its responses to various anchor measures to assess validity (criterion and concurrent), examine the degree of its test-retest reliability, and examine its sensitivity and ability to detect change over time. We will also conduct a preliminary analysis of the recall period of the items. Our overall objective is that at the conclusion of the proposed study, the PRO-CTCAE will be ready for feasibility testing in the cooperative group setting (described in Task 8), in anticipation of broad implementation.

*Patients and Enrollment.* An IRB protocol will be developed and approved at each of 4 participating institutions for this study (Dana-Farber, Duke, M.D. Anderson, MSKCC). A sample of 270 patients will be enrolled at main hospital sites and satellite clinics of these institutions. This will involve approximately 50 patients from each of the six disease groups described above. Patients will be identified and approached by research study associates/data managers at each site and invited to enroll. Eligibility criteria will include adults in the above noted diagnosis/treatment groups who have the ability to understand English and to read questions on a computer screen (or listen to questions via headphones and then select responses on a screen). Each eligible patient will then undergo informed consent. Enrollment and distribution of accrual across disease groups and institutions will be coordinated by the Project Manager at MSKCC, who will train and regularly communicate with local research staff. This model has worked effectively for similar ongoing multicenter studies based at MSKCC involving patient computerized self-reporting.

*Questionnaire Items.* Based on our preliminary work, described in the “Approach to Task 2: Item Development” section, we anticipate that there will be between 40 and 50 symptoms identified in CTCAE v3 that are amenable to severity grading by patients in the PRO-CTCAE (ie, in Categories I or II). All of these items will be included in the questionnaire for this study. Therefore, all patients will respond to all items, regardless of whether these items have been designated as core, symptom-specific, or optional via

the criteria previously described in Task 2. In our prior work at Duke, M.D. Anderson, and MSKCC, we have found it feasible to administer more than this number of symptom items to patients in clinic on repeat occasions via wireless touch-screen computers in the setting of conducting symptom research. This number of items on average takes approximately 20 minutes to administer, with very few refusals or incomplete questionnaires (<5%), even in patients with marginal performance status (given that they are well enough to come to clinic visits). Notably, any CTCAE v3 items identified as Category III will not be included in this study (because our approach to the PRO-CTCAE specifies that patients respond to these items via a yes/no checklist to evaluate for presence/absence of symptoms, without severity grading).

*Anchor Measures: Clinician-Reported, Patient-Reported, and Clinical Indicators.* For the assessments of validity, sensitivity, and clinical significance of score changes, it will be necessary to collect external anchor data. These data will include clinician-reported ECOG performance status and a patient-reported global health state item taken from the EuroQoL EQ-5D at each visit when patients report PRO-CTCAE items. In addition, a CRF will be developed for clinical research associates/data managers at each site to collect a wide array of clinical indicators pertinent to the symptoms under evaluation. ECOG performance status is a validated functional measure that is predictive of death and other clinical events (Oken et al, Am J Clin Oncol 1982). It is also a common metric used to determine patient eligibility for clinical trials. The EQ-5D single global item has similarly been used and evaluated extensively in the clinical trial setting as measure of patient current health state (Rabin and de Charro, Ann Med 2001; Brooks et al, *Measurement Valuation Health Status* 2003). Clinical indices for collection will be determined by the investigators before the study based on literature review and expert consultation. Because numerous symptoms are being collected, myriad indices will be of interest. A table will be created that lists the particular indices for each PRO-CTCAE symptom. The investigators have previously used this method in validation work. During conduct of the study, the CRF will therefore be used to collect information on use of selected supportive medications, chemotherapy and radiation treatments administered, cancer treatment dose modifications or schedule changes, blood transfusions, palliative radiation treatments, hospitalizations, emergency department visits, and referrals to hospice. In addition, three general interference items from the MDACC instruments (BPI, BFI, MDASI) will be included (specifically, interference with sleep, general activity, and enjoyment of life).

*Questionnaire Administration.* Each enrolled patient will be asked to complete questionnaires on three separate occasions: on the day of enrollment and at two subsequent visits. For most patients, this will be at consecutive medical oncology office visits (generally every 3-4 weeks). An exception to this is the 50 patients undergoing daily chemo-radiation (i.e., head and neck/gastroesophageal carcinoma patients), who will be asked to complete questionnaires on six occasions to evaluate test-retest properties: specifically, on days 1 & 2, days 15 & 16, and days 35 & 36. All participating patients will be approached by the site clinical research associate on scheduled

reporting days during the study. Electronic questionnaires will be completed by patients using wireless touch-screen tablet computers, which will be equipped with Verizon network cards. Therefore, these computers will transmit data directly into the secure patient database in real time over the cell-data network, without necessity to link into local institutional networks. This approach has been highly effective in prior multicenter work by Dr. Basch's group. We plan to use the PRO-CTCAE electronic platform created by SemanticBits, LLC, for this assessment, as previously described in the "Substantive Approach to Task 5: Platform Development" section. As a backup, the MSKCC STAR platform will be available, which will be modified by MSKCC programmer Marwan Shouery to have the exact same look, feel, and functionality as the PRO-CTCAE platform. This backup approach will ensure that no unanticipated delays occur in progress of this study for technological reasons.

*Analysis of Validity*. In addition to content validity (as outlined in Task 2), we will use known-group validity and criterion-related validity to demonstrate the validity of the PRO-CTCAE items. For this analysis, we will analyze each individual symptom across the population as it relates to clinician-reported ECOG performance status score (Oken et al, Am J Clin Oncol 1982), to a single patient-reported global health state item taken from the EuroQoL EQ-5D (Rabin and de Charro, Ann Med 2001; Rabin and de Charro, Ann Med 2001; Brooks et al, *Measurement Valuation Health Status* 2003) to the three generic interference items (interference with sleep, general activity, enjoyment of life), and to clinical indicators selected a priori that are expected to relate to that symptom. We will use ECOG performance status to show evidence of known-group validity (ie, patients with poor performance status are hypothesized to report higher relative levels of symptom severity and vice versa). Criterion-related validity will be used as another measure of validity. Clinical indicators specific to each symptom will be collected and documented by clinical research associates via study forms during the study. We hypothesize that clinical indicators will be correlated with specific symptom items. For example, clinical indicators pertaining to fatigue will include intervention for anemia (blood transfusion or administration of erythropoiesis-stimulating agent); referral to a fatigue clinic (relevant at two participating sites); initiation of a symptom-specific supportive medication (modafinil, methylphenidate, or a corticosteroid); change in narcotic analgesics; or hospitalization, emergency department visit, or referral to hospice. Clinical indicators for sensory neuropathy will include initiation or increase of pain medications, dose reduction of chemotherapy associated with neuropathy (eg, taxanes, platinums, vinca alkaloids), initiation/increase in supportive medications commonly given for neuropathy (amitriptyline, gabapentin, pregabalin), or sleep disturbance. We will examine the relationships between PRO-CTCAE items with interference items using Pearson/Spearman correlations. Further examination of the items will include comparison of PRO-CTCAE items across gender, ethnicity, age, etc. This will investigate whether responses are equivalent and are not biased depending on certain group memberships. Although we will collect clinician-reported CTCAE symptoms and will have this information available for comparison to PRO-CTCAE reporting, we have previously found patient and clinician CTCAE symptom reporting to

be discrepant (Basch et al, Lancet Oncol 2006), and therefore such correlation will not be considered a criterion for establishing validity.

*Rasch Model (one parameter IRT model)*. For the assessment of validity of items, in addition to using classical test theory approaches, we plan to employ methods of modern test theory, in particular the use of a family of Rasch models (eg, partial credit model, rating scale model) (Rasch, *Prob Models* 1980; Hambleton, Medical Care 2000; Andrich, Psychometrika 1978; Reeve and Masse, in *Method Test Surv Quest* 2004). The majority of studies evaluating psychometric properties of newly developed (or modifications of existing assessment tools) have been based on classical test theory, but the Rasch model provides a complementary approach to validation. Modern test theory allows for differential item functioning analysis (DIF). If the item measures the same construct/trait across groups of patients, we expect the same response curve regardless of group membership (Angoff, in *Diff Item Func* 1993). We plan to conduct such an analysis on the collected symptom questionnaire data.

*Analysis of Reliability.* The primary reliability testing will be performed via a test-retest evaluation in the subset of 50 patients receiving chemo-radiation treatment (i.e., head and neck/gastroesophageal carcinoma patients). These patients will be asked to provide data on consecutive treatment days. We will use the Spearman rank correlation and the intraclass correlation to estimate test-retest reliability. As few as 30 patients are adequate to achieve >90% power to test for significant correlation (i.e., correlation different from zero) when the true correlation is 0.60 using a two-sided test with 0.05 level of significance. Because each item in the PRO-CTCAE is intended to stand alone as an individual indicator and there will not be composite scoring, internal consistency reliability is not relevant toward operationalization. Nonetheless, we will calculate Cronbach's alpha coefficient among subsets of CTCAE items that we hypothesize are measuring related entities based on prior development work at the MDACC and MSKCC, pathophysiology, and clinical acumen (eg, nausea and vomiting; pain and neuropathy).

*Analysis of Sensitivity.* Contingency table analyses of sensitivity and specificity will be used to assess the level of agreement of the newly developed PRO-CTCAE items with clinician-rated ECOG performance status and the documented symptom-specific clinical variables. A reliable change index score will be calculated using longitudinal data from each patient's multiple visits to assess the amount of change that is seen in the PRO-CTCAE items and the measured clinical variables. The investigators have extensive experience with this approach.

**5.g.4. *Clinical significance of score changes (minimally important difference /responder definition)***

Both anchor-based and distribution-based methods will be used in the determination of a clinically meaningful score change for each item. Dr. Sloan's prior work has focused specifically on the assessment of clinically significant changes both between groups and

over time for individuals (Sloan et al, Mayo Proc 2002). We will anchor the PRO-CTCAE items against ECOG performance status and a patient self-reported global health state item and to the selected symptom-specific clinical variables. Distribution-based approaches will involve Dr. Sloan's generally accepted half standard deviation approach to determine a nonignorably large difference and then use one-fourth and one-third standard deviation estimates to compare with the anchor-based estimates.

*Preliminary Exploratory Study of Recall Period.* Recent evidence suggests that patients may have difficulty recalling the magnitude of their symptoms beyond several days, with likely variation depending on the measure (Broderick et al, Pain 2008). We will conduct a preliminary exploratory study of recall for PRO-CTCAE items in the cohort of 50 patients receiving daily chemoradiotherapy. Notably, in this cohort, patients will be reporting PRO-CTCAE symptoms on six occasions: days 1 & 2, days 15 & 16, and days 35 & 36. Since the retrospective inclusion time for symptom assessments by these patients will encompass the entire prior treatment duration, we will be able to assess several recall timeframes. We hypothesize that patient symptom ratings based on 2-week recall will not be highly correlated with ratings based on a 24-hour recall. To detect a correlation of at least 0.70 between the two recall periods, assuming a two-tailed test at a 0.05 level of significance, with 42 patients we will have 90% power. We recognize that this is a very preliminary and limited design for assessing recall. However, the proposed analysis is one which is feasible in the larger context of this planned study during the contract period. Notably, as an alternative design in order to more thoroughly evaluate recall period, we could opt for daily self-reporting over a 30-day period in these 50 patients, and this potential design element, which would increase the cost of the study, will be discussed with NCI staff prior to protocol development. An additional limitation of this analysis is that patients with head and neck/gastroesophageal carcinoma receiving chemo-radiation are not expected to experience the full spectrum of symptoms in the PRO-CTCAE. Therefore, there may be consistent scores of zero for some symptoms. In summary, we recognize that this is a limited initial evaluation of recall period which is essentially exploratory, and that further research of this question will be necessary. In the future, we are interested in conducting a more extensive evaluation of the recall period in a broader population to define the appropriate frequency of PRO-CTCAE reporting in trials. Such a study would likely require the use of a method for regular self-reporting between visits, such as via a handheld device or IVR. Such an evaluation is beyond the scope of this proposal but is recognized by the investigators as an important area to further evaluate.

*Power Calculation.* In evaluating the psychometric properties of the items, although we can calculate sample size based on reliability, we consider validity to be more important. Hence, our power analysis is based on the ability of the PRO-CTCAE to distinguish between patients with limited performance status and those with good ECOG performance status as a measure of known/group validity. To show that the PRO-CTCAE is sensitive enough to differentiate between patients based on their performance status, a single PRO-CTCAE score should be significantly different between patients

with reduced functionality (ECOG PS 2, 3, or 4) or fully active (ECOG PS 0 or 1). A one-point difference between patients with good and limited performance status at a given time point is deemed to be clinically important (Cohen, *Stat Power Behav Sci* 1988). This one-point difference corresponds to an effect size of 0.50 using Cohen's standards based on a standard deviation of 2.0 from previous research validating symptom items that uses a 0-10 NRS scale. Because we anticipate including approximately 45 distinct PRO-CTCAE items, we will control for our Type 1 error rate using Hochberg's step-up method (given the large number of tests, Bonferroni's method may be prohibitively conservative). For power calculations, however, we have used the more conservative Bonferroni's method. Thus, using a significance level of 0.05/45=0.0011, two-tailed, and power of 80%, a total of 276 patients (138 patients with ECOG PS 2, 3, or 4 and 138 patients with ECOG PS 0 or 1) will be needed to detect this one-point difference (an effect size of 0.5 standard deviations using a two-sample t-test). With a total of 300 patients, we would have 84% power to detect this (this analysis will be adjusted for the final tally of PRO-CTCAE Category I and II items included in the study, which will be based on the item development work previously described in Task 2). Adequate power will still be available even if a smaller proportion of patients with ECOG PS 2, 3, or 4 are accrued (with 120 patients with ECOG PS 2, 3, or 4 [180 patients with ECOG PS 0 or 1] the power is 82% and with 90 patients with ECOG PS 2, 3, or 4 [210 patients with ECOG PS 0 or 1] the power is 74%). The distribution of ECOG PS will be monitored during patient accrual and the protocol will be amended to allow for greater total accrual if less than 30% of patients are of limited ECOG PS. Note that confirmatory analysis of the symptom measures will include comparison between patients with limited and good ECOG PS while accounting for disease group.

*Determine Response Scale Cut Points for Triggering Automated Rules-Based Notifications.* Determinations of threshold cut points for PRO-CTCAE severity item alerts will be determined using established methods (Palos et al, J Pain 2006; Serlin et al, Pain 1995; Anderson, Pain 2005; Jensen et al, Pain 2001; Paul et al, Pain 2005; Zelman et al, Pain 2003; Mendoza et al, Clin J Pain 2004; McConahay et al, Acad Emerg Med 2006; Chow et al, J Palliat Med2006; Dihle et al, Clin J Pain 2006) and will be based on data obtained in the Task 2 and 7 studies (in which both severity and interference data are collected), through mining of data sets at the M.D. Anderson, MSKCC, and Michigan State University, and through literature review. Clustering of severity scores within the selected scale vis-à-vis interference anchor measures will be used to establish cut-point boundaries for alerts to clinicians for each PRO-CTCAE item (at "moderate" levels, generally $\geq$5 on 0-10 NRS) and for expedited reporting (at "severe" levels, generally $\geq$7 on 0-10 NRS), although we recognize that cut points may differ for different symptoms (for example, we have observed in work at the M.D. Anderson and MSKCC that cut points for shortness of breath tend to be relatively slightly lower than other symptoms by one point on average). Both empirically derived and clinician recommended thresholds will be used in determining the final cut points.

*Electronic vs. Paper Administration of the PRO-CTCAE*. Recently, a comprehensive meta-analysis summarized the scientific literature comparing electronic vs. paper reporting, including 46 clinical trials and 275 PRO measures (Gwaltney, *Value in Health* 2007). The average mean difference between the paper and electronic versions was very small (0.2% of the scale range or 0.02 point on a 10-point scale), and the average correlation between the paper and ePRO measures was very high (0.90). Overall, evidence indicated that the difference between paper-and-pencil administration and electronic assessment of patient-reported outcomes produces virtually identical results, with any differences observed being strictly due to random variation. Notably, fewer items tend to be skipped during electronic administration. For the PRO-CTCAE platform, we will include the ability to print questionnaires as paper CRF, and paper will therefore be an alternative backup administration strategy. However, it is the intention of this system that most patients complete forms electronically. It is beyond the scope of this proposal to compare paper vs. electronic administration of the PRO-CTCAE items, although this is an area of interest to the investigators for future evaluation. However, if the NCI staff believes that this is an important area for initial development work, we are open to integrating such a design into the above described study. The investigators bring extensive experience conducting such analyses, and three of the investigators are members of the ISPOR task force, which has recently completed a White Paper on this exact topic.

Specific Deliverables and Timeline for Task 7.
As per the RFP, the timeline for deliverables and timeline for this task are detailed below in Table A.5-18.

**Table A.5-18. Specific Deliverables to NCI and Timeline for Task 7**

| SPECIFIC DELIVERABLES | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| Draft protocol/detailed design for validation study | Month 13 |
| Final protocol/detailed design for validation study | Month 14 |
| Status report on implementation of validation study | Month 20 |
| Final report on findings from validation study | Month 22 |
| Recommendations for future studies | Month 22 |

### *5.h. Substantive Approach to Task 8: Design Feasibility Study*

| | |
|---|---|
| Task 8: | Develop a protocol in the NCI cooperative group setting to evaluate the feasibility of PRO-CTCAE items nested within treatment trials in patient populations with specific expected adverse symptom events and administered at cancer centers and Community Clinical Oncology Program (CCOP) sites. |
| Year of Contract: | Year 2 |
| Task Leader: | Dr. Deb Schrag |
| Deputy Task Leader: | Dr. Ethan Basch |

**5.h.1 CALGB 70501: Currently Open Feasibility Study of Patient CTCAE Reporting**

*Overview*

Dr. Basch is the National Study Chair (PI) for CALGB 70501, "Collection of Patient-Reported Symptoms and Performance Status Via the Internet" (included in full in the Appendix). This trial is currently open in the CALGB. In this study, patients enrolled in a cooperative group treatment trial are asked to self-report their own adverse symptom events at scheduled followup visits via wireless touchscreen computers. The electronic platform used in CALGB 70501 for patient self-reporting is the web-based "STAR" ("Symptom Tracking and Reporting") system, which was previoulsy developed by Dr. Basch's programming team at MSKCC. The items uploaded into the STAR platform in CALGB 70501 are pilot patient adaptations of selected CTCAE v3 symptom items which were developed and evaluated at MSKCC (Basch, JCO 2005; Basch, Lancet Oncol 2006; Basch, JAMIA 2007; Basch, JCO 2007).

*Study Design*

CALGB 70501 is a limited-access companion study. This means that it can be linked to multiple treatment trials, and can be opened successively at study sites to incrementally assess feasibility without having to open across the entire CALGB network at once. In addition, patients consent to 70501 independently from the "parent" treatment trial, and 70501 thus carries cancer control credits which are awarded for patient registration, which is an incentive to enroll at CCOP sites.

The sample size for CALGB 70501 is 175. A schematic for the study is shown below in Figure A-5.9. Detailed descriptions of the various forms cited in this figure are provided in the Appendix. Patients at participating sites who are enrolling in a linked "parent" treatment trial are approached by local research staff and invited to also enroll in 70501. To be eligible, patients must be able to read English on a computer screen (although for the PRO-CTCAE we will design an auditory-assitance administration option via earbuds, and in the furture we plan language translations, as described in section 5.c of this application). Once enrolled, patients undergo a

10-minute training session with local staff and may select a secure username and password. Patients then complete a "baseline" login sesion.

At the five following scheduled clinic visits, patients are approached by research staff and invited to login and self-report their adverse symptom events. A "Report" showing the longitudinal symptom reports of that patient is then printed and presented to the treating clinician for review during the visit. It is left to the discretion of clincian and patient whether to discuss or act on this information.

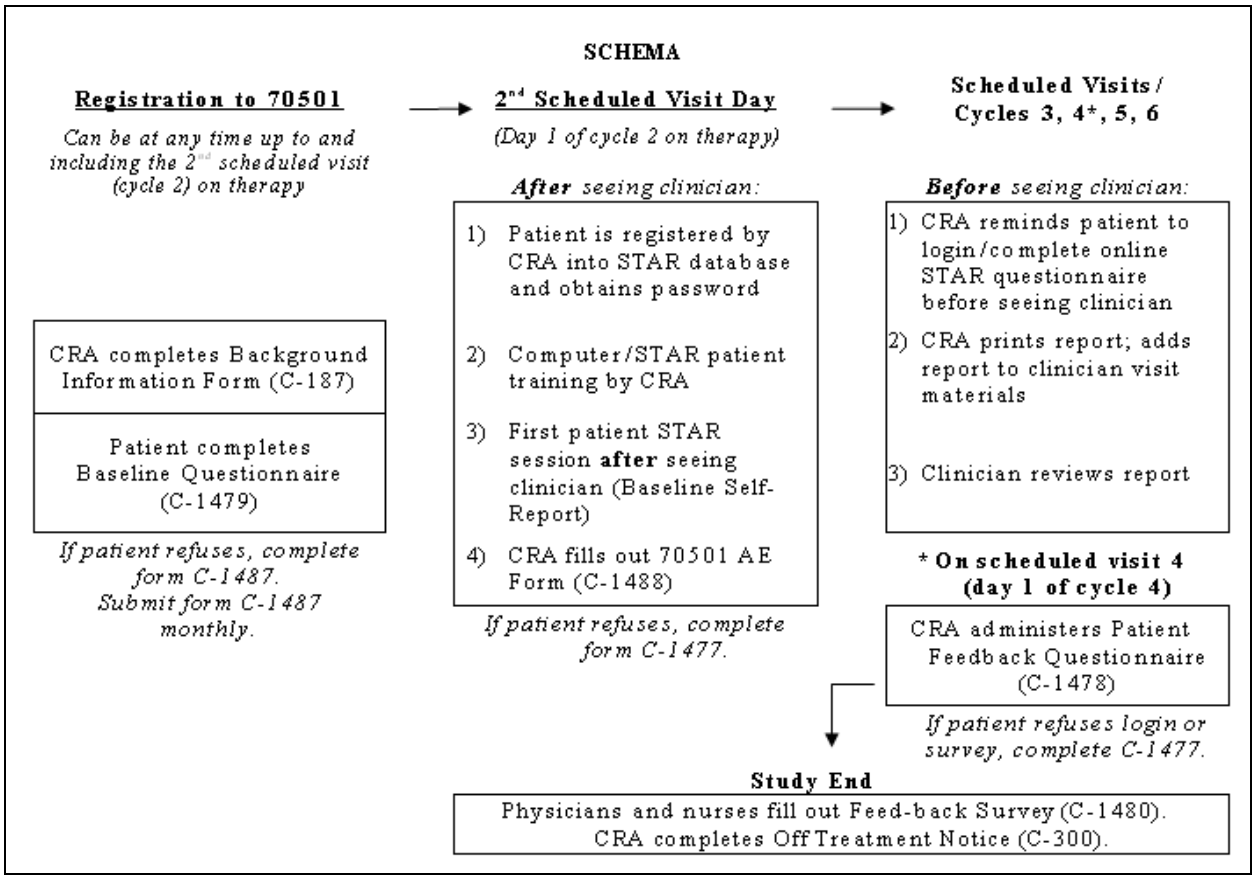



**Figure A-5.9. Schematic for CALGB 70501 (see full protocol in Appendix)**

*Data and Analysis*

For the analysis of feasibility, we track several metrics. First, we follow rates of patient willingness to enroll in 70501 when invited, and reasons for refusals. Then, we follow whether or not patients complete their scheduled logins at followup visits (ie, "adherence"), in order to ultimately tabulate the proportion of visits at which a given patient completes an electronic CRF, and the proportion of enrollees overall who adhere at any given sequential visit. We are interested in reasons patients may not login, and therefore a CRF is completed by local staff to document the reason(s) in cases of missed logins (ie, too ill, too busy, technical difficulty, etc). A staff CRF is included for mandatory clinician completion of the same CTCAE items that patients

are completing, in order to correlate the patient and clinician grading of analagous CTCAE items. And finally, there are patient and clinician feedback forms to assess acceptance of the system, satisfaction, and recommendations.

Ultimately, we will conclude that this is a data collection strategy worth pursuing further if 80% of approached patients agree to consent, and if a mean of 80% of enrolled patients login at followup visits. With a sample size of 175, the power to detect a difference in the proportion of patients completing online forms between 0.7 and 0.8 with an alpha of 5% (one-sided test for binomial proportions) is 85%, with correction for intraclass correlation.

*Recent Expansion of CALGB 70501*

CALGB 70501 was initially linked to treatment trial CALGB 90401, a randomized controlled trial of docetaxel in men with hormone-refractory metastatic prostate cancer. Because 90401 was close to completing accrual at the time 70501 opened, an amendment was written to link 70501 to four other treatment trials: two in breast cancer, and two in lung cancer (CALGB 30607, 30704, 40601, and 40603). This amendment has been internally approved in CALGB, and is currently being submitted to DCP (NCI Division of Cancer Prevention) for review. It is the amended version that is reporduced in the Appendix, with permission from CALGB to include it for demonstration purposes.

This amendment also expands the number of sites at which CALGB 70501 will be opened. CALGB 70501 was initially open at a limited numer of sites: Christiana Care CCOP, Memorial Sloan-Kettering Cancer Center, Missouri Baptist Medical Center, Northern Indiana Cancer Consortium, University of California at San Diego, University of Iowa, and Yale Cancer Center. In the amendment, we have expanded this to 25 sites, to which we will be sending wireless touchscreen tablet computers with built-in Verizon cards to transmit data to the secure central database without requiring a local network connection.

**5.h.2 Study Design for PRO-CTCAE Feasibility Assessment Fundamentally Differs from Standard Approaches to Measuring HRQL in Treatment Trials**

Initially, through a close working relationship with the 90401 National Study Chair, Dr. Kevin Kelly, Dr. Basch was able to integrate patient self-reporting of adverse symptom events into the flow of clinic visits during conduct of the treatment trial. This was a key accomplishment, because 70501 requires that patients self-report their toxicities prior to seeing a clinician at most visits, and that printed longitudinal reports of patients' toxicities be printed and presented to clinicians in real-time. More recently, Dr. Basch has worked closely with the PIs of the four additional linked treatment trials, as well as with the Breast and Respiratory committees at CALGB.

As such, the approach to this feasibility assesment fundamentally differs from classic health-related quality of life (HRQL) companions to treatment trials, for which

patients may report after visits are completed—and results are not generally reviewed by a clinician at the visit. Furthermore, this study differs from HRQL assessments because the primary metric is feasibility, and therefore multiple CRFs are required for completion by clinical staff to document system functionality, reasons for patient lack of adherence, and clinician acceptance. Finally, this design differs because patients are reporting symptoms which are potentially actionable in the context of a treatment trial if severe.

**5.h.3 Amendment of CALGB 70501 for PRO-CTCAE Feasibility Study**

As noted above, we believe that the fastest route to opening a cooperative group feasibility study of the PRO-CTCAE is to amend CALGB 70501 for assessment of this new adverse symptom assessment tool, and then open it not only across myriad sites in the CALGB (as we will be well established in numerous sites already from our ongoing work), but expand it into additional cooperative groups. Notably, investigators on this proposal hold leadership positions in multiple cooperative groups: CALGB (Schrag and Basch), NCCTG (Sloan), RTOG (Bruner and Trotti), and SWOG (Cleeland). Furthermore, this concept has already been presented and embraced by senior leadership of cooperative groups (see attached letters of support)

**5.h.4 Identification of Cooperative Group Trials for Linking PRO-CTCAE Feasibility Study**

As noted, the investigators participate in committee work in several cooperative groups. Drs. Basch, Schrag, and Abernethy are medical oncologists, and Dr. Trotti is a radiation oncologist, and therefore they have close collegial relationships with clinical investigators running trials in the cooperative groups. Therefore, the investigative team will be able to identify candidate treatment trials for linkage to the PRO-CTCAE feasibility protocol early in their development, in order to allow the PRO-CTCAE study to be linked and accruing from the time of a treatment trial opening. Notably, this is the approach that Dr. Basch has taken recently, and all four of the new trials to which 70501 will be linked are newly opening.

Our close ties in the cooperative groups will also allow us to educate clinical research staff, elicit feedback about the protocol at cooperative group meetings (as described in sections 3 and 10), and enlist local institutional allies to facilitate IRB review. This has also been a strategy followed by Dr. Basch with success, and he has presented this protocol at multiple CALGB committee meetings, including the Breast, Respiratory, Genitourinary (of which he is a liaison member), Quality of Life, Health Services (of which he is Vice Chair), and Clinical Research Associate committees. He has also presented this work as an invited presentation at a CALGB Group Meeting.

**5.h.5 Design of PRO-CTCAE Cooperative Group Feasibility Study**

Although the overall feasibility design and endpoints for the PRO-CTCAE cooperative group study are anticipated to be similar to the approach in CALGB 70501, there are several design element of the PRO-CTCAE system that are somewhat more sophisticated than the MSKCC STAR platform (as described in section 5.e), and hence merit special consideration:

- *Creating PRO-CTCAE CRFs:* Clinical investigators for the "parent" trials will be required to use the Form Builder/Publishing Tool to create the PRO-CTCAE electronic CRFs (system described in section 5.e). They will use the draft Training/Educational Materials to guide this process (described in the following section, 5.i). The rules-engine will be used during this process to set automated notifications which will determine when email alerts should be sent to clinicians for specific severity thresholds, or trigger expedited reporting forms. The schedule for required patient PRO-CTCAE logins will be linked the caBIG Study Calendar. The PRO-CTCAE items in the CRFs will include the mandatory "Core" items, "Diagnosis-Specific" items if applicable, and any "Optional" items the protocol of clinical investigator designates.
- At each specified visit, patients will be reminded to login by local clinical research staff. Wireless touchscreen tablet computers or touchscreen computer kiosks will be available at each participating site for this purpose (to be supplied by Drs. Basch and Abernethy). When they login, in addition to the items required in the PRO-CTCAE CRF, patients will have the option to select additional "Optional" items from a dropdown or checklist. These items will then be added to the required list of items for subsequent visits.
- For "checklist" items such as rash, in which patients only select Y/N for the presence/absence of that phenomenon, clinicians will be prompted and required to grade severity. This will either be by computer or paper (a design element that will have to be worked out depending on the results of usability testing in Task 5).
- If automated alerts are triggered by the rules-engine during the study, clinician actions will be tracked by research staff with an applicable form.
- When patients miss a scheduled login session, it will trigger an automated alert to study staff.
- For patients unwilling or unable to complete the online PRO-CTCAE CRF, or for sites with technical difficulties, generation of paper CRFs (a planned function of the PRO-CTCAE platform) will be an option.
- There will be direct transmission of data to CDUS and caAERS/AdEERS from the PRO-CTCAE platform.

For each of these areas of system functionality or user experience, we will track feasibility outcomes and barriers when they occur, via specific forms. We will follow clinician ability to use the Form Builder and specific difficulties; patient selection of additional items from the dropdown; clinician adherence with grading items reported by patients via the Y/N checklist; use of paper PRO-CTCAE CRFs instead of

electronic reporting; and success of PRO-CTCAE platform interoperability with other NCI information systems.

**5.h.6 Rationale for Selection of "Parent" Treatment Trials and Patient Populations for Linkage of PRO-CTCAE Study**

Selection of trials and patient populations for linkage of the PRO-CTCAE feasibility protocol, and selection of study sites, will be directed towards assuring a broad representation of patient characteristics and symptom prevalence. The rationale for selecting specific disease groups has previously been discussed in section 5.2.g. Selection of trials for linkage is also anticipated to depend in part on logistical issues in the cooperative groups.

**5.h.7 Timeline and Deliverables**

It is realistic to expect that we can develop a protocol for assessing the widespread feasibility of the PRO-CTCAE in the cooperative group setting quite quickly through a fairly straight forward amendment to CALGB 70501. However, we cannot guarantee the pace of amendment approval, IRB review at the participating sites, or accrual rates (which ultimately depend on rates of accrual into the treatment trials to which the PRO-CTCAE protocol will be linked). Nonetheless, given that 70501 took almost three years to move through the CALGB committee structure before approval, we feel that the best existing opportunity for opening a feasibility study of this type quickly is to amend the 70501 protocol, link it to additional studies, and move it into other cooperative groups, including RTOG, ECOG, and NCCTG. As noted above and in our letters of support, we have already gained support for this approach from senior leadership and committee members in several cooperative groups.

Early in Year 2 of the proposed contract period, we will plan to amend CALGB 70501 to include administration of the developed and refined PRO-CTCAE items. We will complete a draft of the amendment in Month 13, after which we will quickly refine the protocol in order to expedite review at CALGB and DCP. We will continue to update NCI staff on the progress of the protocol, as it makes its way through local IRBs. For sites in the CALGB where 70501 is already open, this will mean only an amendment review. However, for added sites in the CALGB or in other cooperative groups in which we opt to open the study, the process may be more exhaustive. Dr. Basch's team has extensive experience communicating with local research staff to facilitate IRB review and accrual. It is our hope that patients will start accrual in this PRO-CTCAE feasibility study at cooperative group sites prior to the Month 22 status report date, and we will aim for this to be the case. But again, this cannot be guaranteed. The proposed timeline for deliverables is shown below in table 5.h.19.

The number of sites at which to open the PRO-CTCAE feasibility study will be determined with cooperative group input at the time of protocol amendment (current 70501 amendment specifies 25 sites). Computers for this work will be provided to

participating sites by Drs. Basch and Abernethy, and will be returned at the completion of the study.

“Nesting” the feasibility assessment as a “correlative” study within treatment trial protocols would be an alternative strategy, in order to avoid double consent. However, this approach would require opening the study at all sites participating in the parent “treatment” trial, because per policy, nested correlatives cannot be limited access.

**Table 5.h.19. Schedule of Deliverables to NCI for Task 8**

| SPECIFIC DELIVERABLES | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| Draft of detailed design/protocol for feasibility study plan | Month 13 |
| Final detailed design/protocol for feasibility study plan | Month 14 |
| Status report #1 of implementation of feasibility study plan | Month 17 |
| Status report #2 of implementation of feasibility study plan | Month 20 |
| Status report #3 status of implementation of feasibility study plan | Month 22 |

## *5.i. Substantive Approach to Task 9: Create Training Materials*

| | |
|---|---|
| Task 9: | Create paper and electronic training manuals/educational materials and demonstrations for the PRO-CTCAE for both patients and clinical staff |
| Years of Contract: | Year 2 |
| Task Leader: | Dr. Vish Viswanath |
| Deputy Task Leader: | Dr. Jennifer Hay |

### 5.i.1. Overview

The Health Communications Core (HCC) at the Dana-Farber under the direct supervision of Dr. Vish Viswanath, will take the lead on planning and developing training materials for patients and clinicians, including a printed manual, a web-based training module, and a DVD. Providing patients and clinicians with these educational and reference materials will maximize accessibility and effectiveness of the overall system. Development of these materials will progress through six phases, with specific deliverables at the completion of each phase.

### 5.i.2. Specific Phases of Development

The six phases of development are, (1) *Planning, Requirements Analysis, and Project Definition*, which will begin during year one with HCC staff taking part in meetings and conference calls, and working closely with project staff to fully understand the content, functionality, and specifications of the system, in addition to understanding the audience's needs; (2) *Information Design,* in which the HCC will create outlines for the printed training materials, and the organizational structures for the web site and the DVD; (3) *Copywriting*, when the HCC will flesh out and fully develop the content for the educational materials; (4) *Visual Design and Graphic Production*, when HCC staff will create the visual style for the training materials and create all associated layouts and graphics, based on the graphic identity and style established for the system; (5) *Materials Production*, when the HCC will create final layouts for printed training materials and complete programming for the training web site and DVD; and (6) *Quality Assurance (QA) Testing*, when the HCC will do final proofing of the printed training materials, and proofing and testing of the training web site and DVD. Following the QA Testing phase, the printed materials and DVDs will be ready for printing/manufacturing, and the training web site will be ready for launch.

### 5.i.3. Deliverables

Specific deliverables to Dr. Basch's team for each phase will include, (1) detailed work plans and timelines, including plans for reviews and approvals; (2) up to three iterations of outlines for the printed materials, site maps for the web site, and flow charts for the DVD; (3) complete text drafts for each of the three formats, up to three iterations; (4) three iterations of graphic design work, beginning with simple mockups during the first round, progressing to fully shaped prototypes of representative pieces at the third round;

(5) based on the approved information design, content, and graphics, all content for the printed materials will be laid out, and proofs delivered for approval; all programming associated with the web site and DVDs will be completed, and prototypes made available for approval; and (6) the HCC will generate reports on the results of QA testing, and the resolution of any problems identified. Upon completion of the QA testing phase, the HCC will coordinate with project and task leaders to deliver all files and documents necessary for production and launch.

The schedule of deliverables to NCI staff is detailed in Table A.5-20.

**Table A.5-20. Specific Deliverables and Timeline for Task 9**

| SPECIFIC DELIVERABLE | DELIVERY DATE (MONTH OF CONTRACT) |
|---|---|
| 1st draft of training and educational materials for PRO-CTCAE system to facilitate implementation in validation studies | Month 13 |
| Revised draft of training and educational materials for PRO-CTCAE system to facilitate implementation in validation studies | Month 14 |
| Refined draft of training and educational materials for PRO-CTCAE system to facilitate implementation in validation studies | Month 20 |

## 6. Example Approaches to Fatigue and Neuropathy

### *6.a. Example Substantive Approach to Fatigue*

#### 6.a.1. Overview

*Importance of Assessing Fatigue in Cancer Trials.* Cancer-related fatigue has been consistently reported as the most prevalent subjective side effect of cancer therapy. Vogelzang et al. (Semin Hematol 1997) reported that 78% of a randomly selected group of cancer patients experienced fatigue during treatment. Although it is methodologically challenging to compare fatigue in cancer patients with fatigue in noncancer patients across studies, in published reports mean fatigue levels in cancer populations are generally significantly higher than in populations without cancer, for example, in analyses of large databases such as the fatigue items in the EORTC QLQ-C30 or Brief Fatigue Inventory (King, Qual Life Res 1996; Mendoza et al, Cancer 1999). Fatigue can be treatment or dose limiting, causing less than optimal delivery of potentially curative therapy. Therefore, it is essential that we develop a PRO-CTCAE fatigue item that is able to capture and grade this symptom in a way that is easy for patients to understand and report, adequately detects and measures the magnitude of the symptom, and is generalizable for use across a diverse patient population in the setting of cancer treatment trials. In the below sections, we describe the specific steps in our proposed approach to developing and evaluating a fatigue PRO-CTCAE item as an illustration of our overall approach to the PRO-CTCAE.

*Investigators' Experience Evaluating Fatigue.* The investigators have extensive past and ongoing experience in the evaluation of cancer-related fatigue. Drs. Cleeland and Mendoza developed and evaluated the Brief Fatigue Inventory (BFI), one of the most commonly used fatigue-specific instruments in cancer trials. Dr. Cleeland's team at MDACC also developed the MDASI, a multisymptom instrument that includes a fatigue item. This item was developed based on literature review and evaluation of measurement properties, using methods similar to those proposed in this application. Dr. Basch previously developed a patient adaptation of the CTCAE v3 fatigue item (ie, MSKCC STAR fatigue item). As previously described in this application, this item preserves the CTCAE v3's double-barreling of concepts (severity and interference) in the response options. Experience using this item in patients with various types of cancer has provided the investigators with the perspective to consider abandoning the double-barreling of CTCAE v3 symptom items in general for the PRO-CTCAE (including fatigue) in favor of a one-concept-per-item approach. Drs. Cleeland and Sloan lead the ASCPRO (Assessing Symptoms of Cancer using Patient-Reported Outcomes) fatigue initiative, in which Dr. Basch participates, which has included multiple meetings in Washington, DC, to determine optimal strategies for measuring fatigue. This initiative has brought together representatives of the NCI, the FDA, academia, and industry toward developing measures that are evidence-based and consistent with emerging standards for PROs. Dr. Basch also participates in the MAPI Values PROOF-C

Fatigue Consortium, which aims to bring industry representatives and the FDA together to characterize U.S. patients' perspectives on the definition of cancer-related fatigue and develop a measurement model to support a fatigue endpoint for labeling claims. Finally, the investigators are aware of the work of the PROMIS Fatigue Domain Team, and if awarded this contract would be pleased to work with NCI staff to determine the optimal strategy for capitalizing on PROMIS work for the PRO-CTCAE.

*Substantive Approach.* As described in the prior sections of this application, we propose a stepwise "production line" of research and technology processes to develop and evaluate PRO-CTCAE items such as fatigue, which will yield timely deliverables consistent with the requirements of the RFP. In the below sections, we describe these steps in terms of the specific questions requested in the RFP for the example symptoms. The investigators are aware that in CTCAE v4, the terms *fatigue*, *asthenia*, *lethargy,* and *malaise* may be separated, and we address this issue specifically below.

### 6.a.2. How the fatigue question would be asked of patients: question design

*Determination of What Concepts Are Appropriate for Measuring Fatigue in the PRO-CTCAE.*

- *Approach*

When developing any symptom item or instrument, it is fundamental to delineate and justify what concept(s) (eg, intensity, interference, bother, periodicity) are appropriate for measurement of that symptom in the context of interest. As described in the "Substantive Approach to Task 2: Item Development" section of this proposal, we propose a methodic approach to addressing this question for the PRO-CTCAE, which is intended to monitor myriad symptoms simultaneously in the context of cancer treatment trials. As described previously in Task 2, our first proposed step is an analysis of what concepts are currently measured for the symptom of interest in CTCAE v3, followed by a consensus process and literature review to determine whether it is reasonable to mirror those concepts in the analogous PRO-CTCAE item.

- *Analysis of Concepts in the Fatigue Item*

The fatigue item in CTCAE v3, as shown in Table A.6-1 asks clinical staff respondents to consider both severity (mild, moderate, severe) and interference (difficulty performing some ADL, interfering with ADL, disabling) during grading.

**Table A.6-1. Current CTCAE v3 Fatigue Item**

| CTCAE v3 TERM | GRADE 1 | GRADE 2 | GRADE 3 | GRADE 4 |
|---|---|---|---|---|
| **Fatigue** | Mild fatigue over baseline | Moderate or causing difficulty performing some ADL | Severe fatigue interfering with ADL | Disabling |

It is not clear whether clinical staff respondents consistently consider both "severity" and interference when grading CTCAE v3 symptoms such as fatigue, and such an evaluation is beyond the scope of this proposal. However, it is germane to this application whether patients are capable of considering both concepts in a single question. As previously discussed, we believe that this double-barreling approach "forces" patients into categories they may not see themselves as fitting. For example, if a patient feels that his or her fatigue is "moderate" in severity but also "disabling," he or she will be forced to choose between grade 2 and grade 4.

- *Nomenclature issues regarding use of the term "Severity" or "Intensity"*

As discussed previously in Section 5.b.6, an emerging issue in the symptom research community is whether there is a distinction between "intensity" and "severity." Please see that section for an overview of this methodologic issue. In general, in this application, when referring to the results of symptom research, we will use the term "intensity." However, in our discussion of creating PRO-CTCAE items which ask patients to globally rate the magnitude of an experience without specific reference to a concept (such as "interference" or level of bother), we feel it is more appropriate to use the term "severity." Unfortunately, due to ambiguity regarding the distinction between these terms at this time, in cases when PRO-CTCAE development is being discussed in the context of prior research or publications, we will use both terms or include one of the terms in parenthesis following the other. We intend to clarify this issue of nomenclature in discussions with NCI staff and key advisors at the time of contract commencement, towards clarity during the item development process.

- *Optional Study: Cognitive Interviews to Evaluate Patient Understanding of Double-Barreled Items: See Appendix 2*

*Determination of Number of Items Necessary to Measure Fatigue in the PRO-CTCAE*.

- *Approach*

When designing the PRO-CTCAE, we must balance the desire to aggregate as much information as possible about a patient's experience, with the necessity of minimizing questionnaire burden on patients. Since our intention is to measure multiple symptoms, it is not ideal to always administer multiple items for each symptom. We are cognizant that the purpose of the PRO-CTCAE is adverse event monitoring and not in-depth symptom characterization. As such, a single initial question-per-symptom approach will be optimal if it can be scientifically justified. Our approach would thus be to screen all applicable symptoms with a single question each, with an option in the PRO-CTCAE platform (toward possible future use) to designate triggering of additional items of particular interest toward further characterization (see "Substantive Approach to Task 5: Platform Development"). Which specific symptoms would trigger additional items would either be left to the discretion of the clinical investigator at the time of

creating the study-specific electronic CRF or could be built in as a default. We anticipate discussing the alternative models with NCI staff during the contract period.

Regardless, assuming that double-barreling of concepts is dispensed with and we adopt a one-concept-per-item approach, the question remains of what concept(s) should always be measured for selected symptoms in the PRO-CTCAE (ie, is it necessary to always measure more than one concept per symptom and hence more than one item per symptom). For fatigue, which is well characterized in the scientific literature, the concepts of greatest interest are severity and interference (as represented in the current CTCAE v3 fatigue item). The investigators concur that for fatigue (and for all PRO-CTCAE graded symptoms), evaluation of severity is essential and therefore must be included in the PRO-CTCAE. This is supported by the observation that the single-item score for "fatigue/tiredness at its worst" in the BFI is highly correlated with the total mean BFI score (Cleeland and Wang, Oncol 1999). The question then remains of whether additional concept items should also be elicited, beyond severity alone (that is, whether they provide sufficient additional information to merit the added administrative burden).

As previously discussed, based on existing evidence we believe that a single fatigue severity item will be sufficient for the purpose of initial adverse symptom event monitoring for most symptoms and furthermore that no additional concepts or items will be necessary for initial fatigue assessment in trials via the PRO-CTCAE. This assertion is supported by a prior factor analysis with a theory-driven factor model derived from a large sample with a 72-item fatigue bank covering a range of fatigue experiences from physical, mental, and social activities in which only one factor (fatigue "severity") was extracted (Lai et al, Qual Life Res 2007). However, discussion of whether additional items would be desired for triggering by a positive fatigue severity screen result above a certain cut-point score is an issue up for discussion with NCI staff.

- *Optional Retrospective Analysis of Existing Data Set: See Appendix 2*
- *Optional Study of Intensity vs. Interference: See Appendix 2*
- *Consideration of Automated Triggering of Additional Items*

As described above and in the "Substantive Approach to Task 5: Platform Development" section, we are building in an electronic system functionality to be able to trigger additional questions when a patient screens positive for a particular level of severity on an item. As such, this function will have the ability to automatically administer additional single items (eg, interference, frequency, or bother) or to trigger entire symptom questionnaires. This is a potential area in which harmonization with PROMIS is conceivable. Therefore, although the standard approach we recommend at the time of preparing this proposal is to

restrict the initial PRO-CTCAE questions administered to patients to severity items, we are building in an option to trigger additional items for specific symptoms of interest.

*Question Structure*.

- *Approach*

How should patients be instructed to rate their fatigue in the PRO-CTCAE? Limited guidance can be sought from CTCAE v3 instruction manual because it is designed for clinical staff rather than patients and does not specify any particular method by which symptom items should be evaluated or elicited (CTEP, CTCAE v3 Online Instructions and Guidelines, 2008). In the context of any given clinical trial, clinical staff must report CTCAE v3 items that are specifically required by the protocol at designated intervals and also report on any additional items they may select on an ad hoc basis. As such, it is implicit in the CTCAE that clinicians will understand the item names (ie, terms) and know how to evaluate them. This approach is not directly translatable to patient self-reporting because specific instructions are necessary to explain to patients what information we wish them to report.

- *Generic Question Formats*

As described in the "Substantive Approach to Task 2: Item Development" section, to develop a model for asking PRO-CTCAE questions, we previously analyzed the CTCAE and its use in clinical trials and existing symptom instruments. We concluded that there are three essential elements that must be included in the instructions/questions of the PRO-CTCAE: the concept being measured (eg, intensity, interference), the specific symptom term(s), and periodicity. Therefore, the structure of questions should include these elements. As previously detailed in that section, we can create a generic format for PRO-CTCAE symptom questions (Table A.6-2). The design of these draft item formats are based on review of existing validated measures created with patient input, consensus of the investigators based on prior item development experience, and consultation with health literacy experts. This generic format could be used in a model in which a separate question is presented for each symptom individually or as an instruction at the beginning of a listing of multiple symptom terms (ie, as a checklist). Final format will be evaluated and refined through the cognitive interviews proposed in the "Approach to Task 4" section of this proposal.

**Table A.6-2. Generic Question Formats for PRO-CTCAE Items**

| |
|---|
| **How [concept] has your [symptom term(s)] been since [period of recall]:** |
| **What is the worst [concept] your [symptom term(s)] has/have been since [period of recall]?** |

Studies proposed in this application are intended to provide information that will populate the bracketed fields in this generic format. As discussed above, "severity"

is proposed to populate the [concept] field. Regarding recall period, we will include a default choice in the PRO-CTCAE technology platform of "[since your last treatment]." This approach reflects a fundamental weakness of the current CTCAE reporting model as it relates to symptom elicitation: patients likely cannot accurately recall their symptom magnitudes throughout the entire time span since last treatment if it is longer than approximately a week (Broderick et al, Pain 2008). The decision about whether to require PRO-CTCAE reporting more frequently than every visit is an issue for NCI staff and is outside the explicit scope of the RFP. Nonetheless, we have included a preliminary study of PRO-CTCAE item recall in Task 7 of this proposal, which may provide the NCI with compelling preliminary data toward such decision making. Below, we discuss the proposed process for identifying specific item terms to correspond to each CTCAE v3 symptom (in this example, fatigue), and below we use this format to propose a draft PRO-CTCAE fatigue item.

*Identification of Existing Instruments for Use as Potential Sources of Terms or Items for the PRO-CTCAE.*

- *Approach*

  For each symptom identified for grading in the PRO-CTCAE (listed in the Task 2 section of this proposal), a literature search and review of existing instruments will be conducted to identify appropriate terms for inclusion in the severity item. For the purpose of this illustration, we will demonstrate how a fatigue severity item would be created. The first step in this process will be to conduct a literature search to identify all instruments measuring the symptom and concept of interest that have been demonstrated as valid and reliable in cancer populations. Source information on development of the instrument will be sought to ascertain if patient interviews or other methods were used to select (or reject) terms characterizing the symptom of interest. Terms will be selected for which there is the best evidence of correspondence to the symptom phenomenon of interest and that capture responses from as many patients with that symptom as possible while minimizing redundancy.

- *Fatigue Literature Review Overview*

  Hjollund et al. summarized assessment tools developed for measuring fatigue in either cancer and/or chronic diseases from 1975 to 2004 (Hjollund et al, Health Qual Life 2007). They found more than 250, with 150 used only once. Seventy-one scales designed specifically to measure fatigue were applied in 416 studies; 156 were multisymptom scales used in 670 studies. These tools varied in their coverage, structure, and frequency of appearance in published clinical research. They also varied in terms of whether they are standardized in administration and scoring and whether they meet psychometric properties of suitable reliability and validity. A systematic review of cancer symptom instruments published in 2006 noted 11 multisymptom assessment tools that include items to assess fatigue, tiredness, weakness, or lack of energy (Kirkova et al, J Clin Oncol 2006). These 11 instruments and pertinent literature citations are listed in Table A.6-3.

**Table A.6-3. Multisymptom Instruments That Elicit Fatigue, Tiredness, Weakness, or Lack of Energy**

| |
|---|
| SDS - Symptom Distress Scale (McCorkle, 1978) |
| RSCL – Rotterdam Symptom Checklist (de Haes, 2003) |
| ESAS - Edmonton Symptom Assessment System (Philip, 1998; Pautex, 2003) |
| MDASI - M.D. Anderson Symptom Inventory (Cleeland, 2000) |
| PSAR - Pain and Symptom Assessment Record (Bouvette, 2002) |
| The Canberra Symptom Score Card (Barresi, 2003) |
| CAMPAS-R - Cambridge Palliative Assessment Schedule (Ewing, 2004) |
| Symptom Monitor (Hoekstra, 2004) |
| The Symptom Experience Scale (Given, 2002) |
| A Computerized Symptom Assessment Instrument (Munro, 1989) |
| MSAS - The Memorial Symptom Assessment Scale (Portenoy, 1994) |
| PROMIS Global Items v1.0 (http://www.nihpromis.org) |

Previously, a review of the validity of fatigue instruments used to assess fatigue in oncology was published, noting several single-symptom multi-item questionnaires that have undergone evaluations of validity and reliability (Table A.6-4) (Schwartz, Pharmacotherapy 2002). Although not included in this review, we have added the PROMIS fatigue item bank to this table to be comprehensive.

**Table A.6-4. Single-Symptom Instruments That Elicit Fatigue**

| |
|---|
| PROMIS Fatigue Item Bank (http://www.nihpromis.org) |
| BFI - The Brief Fatigue Inventory |
| CFS - Cancer Fatigue Scale |
| FSI - Fatigue Symptom Inventory |
| MSFI - Multidimensional Fatigue Symptom Inventory |
| FACT/FACIT-F - The Functional Assessment of Cancer Therapy - Fatigue Subscale |
| LFS - Lee Fatigue Scale |
| MFI-20 - Multi-dimensional Fatigue Inventory |
| PFS - The Piper Fatigue Scale |
| SCFS – Schwartz Cancer Fatigue Scale |
| POMS - Profile of Mood States - Fatigue-inertia Subscale |

Two of these unidimensional fatigue scales are in common use in practice and research (based on frequency of use in published literature): the Brief Fatigue Inventory (Mendoza et al, Cancer 1999) and the Fatigue Symptom Inventory (Hann et al, Qual Life Res 1998; Hann et al, J Pain Symptom 1999; Hann et al, Qual Life Res 2000; Jacobsen et al, J Pain Symptom 1999). The Brief Fatigue Inventory was designed and validated to measure cancer-related fatigue in a variety of cancer types and treatments (Escalante et al, Cancer 2001; Wang et al, Cancer 2001; Wang

et al, J Clin Oncol 2002). Plain language wording of the instrument is intended to make it easily understood by patients with limited educational backgrounds and to make it amenable to language translation (Okuyama et al, J Pain Symptom 2003; Wang J Pain Symptom 2004). Nine items include severity of fatigue at its worst, at the current time, and at its usual level and the interference caused by fatigue on general activity, mood, walking ability, work (both vocational and daily chores), relationships with others, and enjoyment of life. A factor analysis producing a single underlying factor showed the measure's construct validity. Also in common use but not included in the Schwartz review is the Cancer-Related Fatigue Distress Scale (Holley, Oncol Nurs Forum 2000). A fourth measure, the Fatigue Severity Scale (FSS) (Krupp et al, Arch Neurol 1989) has recently been used in cancer populations (Rammohan et al, J Neurol Neurosurg Psychiatry 2002; Hjollund et al, Health Qual Life 2007). The FACT-F is less commonly used than the other instruments (Yellen et al, J Pain Symptom 1997).

In PROMIS Wave 1 testing, 129 fatigue items were administered with 98 items (including 15 legacy items that met psychometric criteria) calibrated using IRT analysis, 22 items retained in the pool but not calibrated, and 9 items excluded (PROMIS Website, 2008).

- *Identification of Prior Fatigue Items and Terms*

Table A.6-5 shows the questions and responses of items eliciting fatigue, tiredness, weakness, and lack of energy in multisymptom instruments that are used in oncology. Severity specifically is measured only in a subset of instruments.

**Table A.6-5. Fatigue, Tiredness, Weakness, or Lack of Energy Items From Multisymptom Instruments**

| QUESTION | SOURCE INSTRUMENT |
|---|---|
| "Have you, during the past week, been bothered by…tiredness?"<br>▪ Not at all/a little/quite a bit/very much | RSCL |
| "Have you, during the past week, been bothered by…lack of energy?"<br>▪ Not at all/a little/quite a bit/very much | RSCL |
| "During the past week, did you have any of the following symptoms? …Lack of energy."<br>▪ Yes/No<br>"If yes, how severe was it usually?"<br>▪ Slight/moderate/severe/very severe | MSAS |
| "How severe are your symptoms? People with cancer frequently have symptoms that are caused by their disease or by their treatment. We ask you to rate how severe the following symptoms have been in the last 24 hours. Please fill in the circle below from 0 (symptom has not been present) to 10 (the symptom was as bad as you can imagine it could be) for each item.<br>▪ Your fatigue (tiredness) at its worst? | MDASI |

| | |
|---|---|
| "Each of the following sections lists 4 different statements. Think about what each statement says, then place a circle around the one statement that most closely indicates how you have been feeling during the past 7 days."<br>▪ Fatigue: 1: I am usually not tired at all/2: I am occasionally rather tired/3: There are frequently periods when I am quite tired/4: I am usually very tired | SDS |
| "Please circle the number that best describes:"<br>▪ Not tired (0) → Worst possible tiredness (10) | ESAS |
| "How would you rate your pain on average?"<br>▪ None/mild/moderate/severe/very severe | PROMIS Global v1.0 |

Table A.6-6 shows the questions and responses of items eliciting fatigue severity within unidimensional multi-item instruments used in oncology.

**Table A.6-6 Fatigue Severity-Related Items From Single-Symptom Instruments**

| Question | Source Instrument |
|---|---|
| "Please rate your fatigue (weariness, tiredness) by circling one number that best describes your WORST level of fatigue during the past 24 hours."<br>▪ No fatigue (0) → As bad as you can imagine (10) | BFI |
| "Rate your level of fatigue on the day you felt most fatigued during the past week:"<br>▪ Not at all fatigued (0) → As fatigued as I could be (10) | FSI |
| "How would you describe the degree of intensity or severity of the fatigue which you are experiencing now?"<br>▪ Mild (1) → Severe (10) | PFS |
| "Describe how you are feeling RIGHT NOW"<br>▪ Fatigued: 1, Not at all/2, a little/3, moderate/4, quite a bit/5, extremely | POMS |
| "Please read each item and circle the number that indicates how much fatigue has made you feel in the past 2 to 3 days"<br>▪ Tired: 1, not at all/2, a little/3, moderately/4, quite a bit/5, extremely | SCFS |
| "I feel tired….."<br>▪ Not at all/a little bit/somewhat/quite a bit/very much | PROMIS Fatigue v1.0 *from:* FACIT-F |
| "I have energy….."<br>▪ Not at all/a little bit/somewhat/quite a bit/very much | PROMIS Fatigue v1.0 *from:* FACIT-F |
| "I feel listless (washed out)….."<br>▪ Not at all/a little bit/somewhat/quite a bit/very much | PROMIS Fatigue v1.0 *from:* FACIT-F |
| "I feel weak all over….."<br>▪ Not at all/a little bit/somewhat/quite a bit/very much | FACIT-F |

Terms used in well-developed severity items include *fatigue*, *weariness*, *tiredness*, and *lack of energy*. Review of available development literature finds these terms prevalent in patients' descriptions of symptoms, which map to the

experience/construct of fatigue and correlate with clinical measures related to the symptom. In an analysis of development work data for the MDASI, all of these terms appear to provide complementary information (ie, are not redundant; notably, *weakness* and *fatigue* were found to be redundant). Therefore, we conclude that it is reasonable for the PRO-CTCAE items to include all four terms: *fatigue*, *weariness*, *tiredness*, and *lack of energy*. Prior cognitive interviewing work and consultation with health literacy experts and patient advocates suggest that it is reasonable in a single question to bundle multiple terms that map to a single symptom phenomenon if they are connected with an "or" statement, as in the planned PRO-CTCAE. However, we will formally confirm this assertion in the planned cognitive interviewing study (described in the "Substantive Approach to Task 4: Cognitive Interviews" section of this proposal). During the contract period, we will also regularly consult with our patient advocate advisers, the patient advocate group in the CALGB CARE committee (which has agreed to review items for us), and our cultural/health literacy experts regarding the selection of terms.

- *Considerations for CTCAE v4*

The investigators are aware that in CTCAE v4, the terms *fatigue*, *asthenia*, *lethargy,* and *malaise* may be separated, with mapping to MedDRA terms accordingly. We propose that the PRO-CTCAE items that will include the four noted terms (*fatigue, weariness, tiredness*, and *lack of energy*) should map to *fatigue* in CTCAE v4.

**6.a.3. How the Fatigue Item would be scored: response option format**

*Categorization of Item by Degree of Subjective Content.* As detailed in the "Substantive Approach to Task 2: Item Development" section, we have previously developed an approach to categorizing CTCAE items based on degree of subjective content. This designation allows us not only to identify items in the CTCAE amenable to patient self-reporting but also to determine if an item is appropriate for patient grading of symptom magnitude (rather than simply using a checklist to report the presence/absence of a symptom). As shown in the table in that section, we have previously designated fatigue as category I (defined as *primarily subjective phenomena, without observable component*) and therefore amenable to patient grading of the magnitude of the symptom.

*Response Format Options*. As described in the "Substantive Approach to Task 2: Item Development" section, we propose a study to compare three alternative response option formats: 0-4 NRS, 0-10 NRS (using the same anchor terms as the BFI), and verbal descriptor scales. The potential advantages and disadvantages of each scale for the PRO-CTCAE are detailed in that section. We assume that double-barreled responses (analogous to current CTCAE v3 responses) will be rejected based on the above described cognitive debriefing study and because of the lengthy verbiage required for each such item as a PRO. Therefore, such response options are not envisioned for inclusion in this study (although we are open to including this option if surprised by cognitive interviewing results). This study will be

conducted in the same 120 patients as for the above described comparison of intensity vs. interference (single questionnaire for both assessments). We will select our response option format for the PRO-CTCAE based on an analysis of measurement error and a series of "patient preference" questions.

*Creation of PRO-CTCAE Draft Fatigue Severity Item.* Based on work and rationales described in the above sections, we anticipate creating a single fatigue item for the PRO-CTCAE that measures the magnitude of severity of fatigue, weariness, tiredness, and lack of energy. We propose to determine through a consensus process early in the contract period whether the 0-4 NRS, 0-10 NRS, or verbal descriptor scale will be used for the PRO-CTCAE. Therefore, for this example, we present both the 0-10 NRS and verbal descriptor scale options (we do not include the 0-4 option in this example, but the look would be similar). The potential construct for the PRO-CTCAE draft fatigue item is shown below in Figure A.6-1.

**Figure A.6-1. PRO-CTCAE Draft Fatigue Item with Two Alternative Response Format Options (Choice Pending Results of Planned Research)**

**PROPOSED QUESTIONS:**
*(Would be Refined in Cognitive Interviews)*

**How severe has your fatigue, weariness, tiredness, or lack of energy been since your last treatment?**

*or:*

**What is the worst severity your fatigue, weariness, tiredness, or lack of energy has been since your last treatment?**

**0-10 NRS RESPONSE OPTION:**

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **No fatigue** | | | | | | | | | | **As bad as you can imagine** |

**VERBAL DESCRIPTOR SCALE RESPONSE OPTION:**

| None | Mild | Moderate | Severe | Very severe |
|---|---|---|---|---|
| ☐ | ☐ | ☐ | ☐ | ☐ |

An alternative approach we will consider is a “symptom checklist” format in which instructions are provided to the respondent at the beginning of the questionnaire and then each item proceeds with the selected symptom terms only, adjacent to the response option scale. An example for the fatigue item is shown in Figure A.6-2. Completion of checklist style questionnaires is generally slightly faster than including a full question per item (as above), but patients may not pay as much attention to the content. Overall, however, these approaches provide relatively similar results. We anticipate that selection of format will be an early topic of conversation with NCI staff and key advisers, and the alternative approaches will be evaluated in the planned usability testing (described in the “Substantive Approach to Task 6: Usability Testing” section of this proposal).

**Figure A.6-2. Possible Alternative “Symptom Checklist” Style Draft Fatigue Item, with Both Response Format Options**

| | |
|---|---|
| **Fatigue, weariness, tiredness, or lack of energy** | **0 1 2 3 4 5 6 7 8 9 10**<br>No Fatigue — As bad as you can imagine |
| **Fatigue, weariness, tiredness, or lack of energy** | **None** ☐ **Mild** ☐ **Moderate** ☐ **Severe** ☐ **Very severe** ☐ |

We recognize that if a 0-10 NRS was adopted, there would be a scoring scale disparity between the PRO-CTCAE and CTCAE (which is graded 0-5). If NCI staff and FDA representatives express a preference during the contract period to keep the grading scales between the PRO-CTCAE and CTCAE compatible, this will be a strong incentive to keep the PRO-CTCAE scale limited to a 0-4 NRS or verbal descriptor scale. This will likely be a point of discussion among the investigators, NCI staff, and key advisers when the results of the study comparing response option formats are available. We also recognize that this item appears simple and similar in various regards to items in other available instruments. This is intentional: the development process for the PRO-CTCAE is intended to yield simple items that capitalize on many years of development work creating and evaluating symptom instruments.

### 6.a.4. How cultural literacy and minimization of bias of the Fatigue Item will be Ensured

As described in the “Substantive Approach to Task 3: Cultural Literacy” section, Dr. Vish Viswanath and his Core at Dana-Farber will provide input throughout the project to ensure that created items use plain language and are amenable to subsequent language translation, administration in children and their proxies, and use in patients with cognitive disability. In addition, two patient advocate advisers are included on this application, Cindy Georghegan and Deb Collyar. Deb Collyer will provide access to the CALGB CARE committee for feedback, which is constituted of patient

advocates. All planned studies will be administered across a diverse patient population, with included stratification by age, sex, ethnicity, and educational level. Comprehensive cognitive interviewing, as described in the "Substantive Approach to Task 4" section, will be used to ensure patient understanding and acceptance of items across a diverse patient population.

**6.a.5. How measurement properties of the Fatigue Item will be evaluated**

As described in the "Substantive Approach to Task 7: Measurement Properties" section of this proposal, we propose a prospective study in 300 patients from diverse cancer populations across our respective institutions. This study uses multiple design elements and analytic techniques to demonstrate the validity, reliability, sensitivity, clinical significance of score changes, and recall period of the PRO-CTCAE items. It is also designed to determine score cut points, as described below. Notably, clinical indicators pertaining to fatigue that will be used as anchors for this analysis will include interventions for anemia (blood transfusion or administration of erythropoiesis stimulating agents); referral to a fatigue clinic (relevant at 2 participating sites); initiation of a symptom-specific supportive medications (eg, modafinil, methylphenidate, or a corticosteroid); decrease of narcotic analgesics; or hospitalization, emergency department visit, or referral to hospice.

**6.a.6. How information would be used to grade an event: mandatory items and cut-point scores for automated notifications**

*Identification of Mandatory and Optional PRO-CTCAE Items for Reporting in Trials.* As described in the "Substantive Approach to Task 2: Item Development" section of this proposal, we will designate a group of "core" symptoms that always must be reported by every patient in every trial regardless of diagnosis. Fatigue will be included in this list. We will also designate "diagnosis-specific" items for the eight most common cancer diagnoses. Other items will be "optional," for either clinical investigators to designate for inclusion in PRO-CTCAE electronic CRFs at the time of study implementation or for patients to select from a dropdown list or checklist at the time of self-reporting during a trial. In this model, fatigue will be reported by every patient in every NCI-sponsored treatment trial at every scheduled visit.

*Determination of Symptom Cut-point scores for Triggering Clinician Alerts, Expedited Reporting, or Additional Items/Questionnaires*. As described previously in the "Substantive Approach to Task 7: Measurement Properties" section of this proposal, we will use established methods to determine cut-point scores of concern to trigger email alerts to clinical research staff and/or prepopulated expedited reporting forms linked to caAERS/AdEERS. In addition, these score thresholds may be considered for the triggering of additional items or questionnaires for selected symptoms.

*Flow of Data.* As we envision this system working, a clinical investigator will first populate a PRO-CTCAE CRF for a given trial. This CRF will include mandatory core items, diagnosis-specific items if applicable, and any additional PRO-CTCAE items

of interest for that study. The system will also be designed to able to accommodate additional questionnaires, and as these are added to the platform, there will be an opportunity to add these to the list of items for patient self-completion at visits. The investigator will use the rules engine to determine if and when alerts should be triggered to clinicians, trigger expedited reporting forms, or trigger additional symptom items. Linkage to the caBIG Study Calendar will allow for different CRFs to be administered at different visits during the study. During the trial, when patients attend visits they will login to the PRO-CTCAE system and self-report items in the PRO-CTCAE CRF. They will also have the option to add additional items from the checklist/dropdown list (which will then be subsequently elicited at follow-up visits).

**6.a.7. How information would be integrated into CTCAE v3 and v4**

Within caAERS, adverse event data reported by clinical staff and patient-reported data imported from the PRO-CTCAE platform would be aggregated. For adverse symptom events, the most likely scenario is that clinicians and patients will both be rating the same symptoms during trials. The extent to which the clinician grading approach may change is beyond the scope of this proposal, except that clinicians will be required to grade category III PRO-CTCAE items (such as rash) if patients report the presence of that phenomenon. Overall, symptom items in the PRO-CTCAE and CTCAE v3/4 may have different names for clinicians and patients, which will map 1-1 to each other. For example, the investigators are aware that in CTCAE v4, the terms *fatigue*, *asthenia*, *lethargy,* and *malaise* may be separated. We have previously addressed issues of nomenclature mapping to CTCAE v4 in this proposal. At this time, in general, the proposed studies of this application are designed to address specific adverse symptom items in the currently existing CTCAE v3 item bank. However, at the time of contract initiation, we will plan to meet with appropriate NCI staff at CTEP and caBIG/NCICB to discuss whether this paradigm will merit any changes given the transition from CTCAE v3 to v4. Regardless, it is our intention that PRO-CTCAE items will be mapped to MedDRA, which we envision will be achieved through full integration of the PRO-CTCAE with caAERS.

**6.a.8. How information would be conveyed electronically**

As described in detail in the “Substantive Approach to Task 5: Platform Development” section, all data will be directly entered into the PRO-CTCAE platform. Data could be stored in the clinical investigator’s site database, to which the software would be pointed at the time of CRF creation. During the trial, data will be accessible by query at the level of the individual patient, site, or entire study, depending on access privileges of a user. Rules-based notifications based on predetermined severity score cut points will have the capacity to trigger caAERS expedited reporting forms, which will be linked initially to AdEERS. Data will be directly exportable to CDUS. The software developer for this contract, SemanticBits, LLC, has extensive experience working with these systems, and these linkages will therefore be relatively straightforward to create. Most of these linkages have already been created within the caAERS system, which SemanticBits developed.

## *6.b. Example Substantive Approach to Neuropathy*

### 6.b.1. Overview

*Importance of assessing Sensory Neuropathy in cancer trials*. This section focuses on the assessment of sensory neuropathy, which is a discrete item in CTCAE v3 ("Neuropathy: sensory"). Assessment of motor neuropathy ("Neuropathy: motor" in CTCAE v3) is not amenable to PRO, and is hence beyond the scope of this proposal.

Treatment-related sensory dysfunction and discomfort are among the most common treatment-limiting factors and among the most distressing physical symptoms of cancer patients receiving chemotherapy (Alberts and Noel, Anticancer Drugs 1995; Casey et al, Brain 1973; Sandler et al, Neurology 1969; Weiden and Wright, N Engl Jrl Med 1972) and bortezomib (Garber, Science 2002). Sensory disturbances range from mild tingling to painful burning paresthesia, hyperalgesia, and allodynia (Forman, Oncology 1990; Postma et al, J Neuro Oncol 1995). Symptoms are usually symmetrical in a stocking-and-glove distribution and were previously considered generally refractory to treatment (Forman, Oncology 1990; Postma et al, J Neuro Oncol 1995), although recently drugs such as gabapentin, pregabalin, and amitriptyline have been suggested (Rao et al, Cancer 2007). Notably, this symptom often persists long after treatment and so also affects rehabilitation and the return of patients to productivity. Chemotherapy-induced neurotoxicity is frequently delayed in onset and may progress over time. In the nervous system, the greatest effects are on peripheral nerves, targeting the neuronal cell body, the axonal transport system, the myelin sheath, and glial support structures (Malik and Stillman, Curr Neurol Neurosci Rep 2008).

The most common chemotherapies associated with neuropathy are taxanes, platinums, and vinca alkaloids, although newer agents such as bortezomib and thalidomide are also associated with this adverse symptom. The incidence and severity of chemotherapy-induced symptom burden are positively correlated with the period and doses used, such that nearly all patients experience discomfort by a third treatment cycle with paclitaxel (Forsyth et al, J Neuro Oncol 1997; Iñiguez et al, 1998), cisplatin (Berger et al, Eur J Cancer 1997; Cavaletti et al, Cancer 1995; Quasthoff and Hartung, J Neurol 2002), or vincristine (Casey et al, Brain 1973; Postma et al, J Neuro Oncol 1993; Sandler et al, Neurology 1969), ultimately leading to discontinuation of therapy in approximately half of patients. This underlines the importance of regular systematic evaluation of sensory neuropathy severity in clinical trials.

*Challenges Specific to Measurement of Sensory Neuropathy in Patients With Cancer.*
A particular challenge related to the measurement of sensory neuropathy is that, like many symptoms, it is a complex multidimensional phenomenon that may be experienced differently by different patients. Therefore, it is essential to identify terms that capture the full constellation of patient experiences yet are specific enough to restrict measurement to the symptom of interest. For assessment of sensory neuropathy in cancer patients, a key issue that must be addressed is whether and how to include measurement of pain. Although in noncancer populations (such as those with diabetic peripheral neuropathy), pain assessment can be included in measures

(because the neuropathy is likely to be the cause of any detected pain), in cancer patients, there are many other potential causes of pain; therefore, inclusion of a pain measurement may reduce the specificity of an item. In the following sections, we address this challenge scientifically through literature review and mining of existing data sets.

*Investigators' Experience Evaluating Sensory Neuropathy.* The investigators have substantial past and ongoing experience in the evaluation of sensory neuropathy in patients with cancer. Dr. Sloan has pioneered an initiative in the NCCTG cooperative group to include patient self-reporting of sensory neuropathy symptoms in a trial of oxaliplatin in metastatic colon cancer. This drug has been associated with clinician-reported CTCAE grade 3 or 4 toxic symptoms in 15%-20% of patients receiving first-line oxaliplatin-based chemotherapy (Green et al, Ann Oncol 2004; Hochster, 2006; Tournigand, 2006) with high rates of patient refusal to continue drug therapy (Green et al, Ann Oncol 200; Saltz, 2007). A chart review by Dr. Sloan's group further suggests that clinician CTCAE grading routinely underreports the prevalence or severity of this symptom. Dr. Cleeland has also conducted substantial work in this area, having done prior research to correlate patient self-reported terms with the results of neurologic quantitative sensory testing (QST). Unlike fatigue, the physiologic component of neuropathy can be directly assessed, for example, via quantitative testing approaches such as QST. QST involves the use of calibrated stimuli (touch, heat, cold, pain) to determine the level of sensory involvement. Dr. Cleeland is currently PI of a large M.D. Anderson trial in which patients taking bortezomib, thalidomide, or oxaliplatin are followed up prospectively for a year to evaluate the emergence of treatment-related adverse symptoms, particularly sensory neuropathy, with correlations to QST. Dr. Basch's Working Group for the Development of a PRO Version of the CTCAE has previously reviewed existing neuropathy instruments and scientific literature to evaluate terms most commonly self-assigned by patients to describe the sensory neuropathy experience and reviewed these items with patient advocates and health literacy experts. Drs. Basch, Schrag, and Abernethy are medical oncologists and therefore have a substantial background in clinically evaluating patients with sensory neuropathy.

*Substantive Approach.* As described in the prior sections of this application, we propose a stepwise "production line" of research and technology processes to develop and evaluate PRO-CTCAE items such as sensory neuropathy, which will yield timely deliverables consistent with the requirements of the RFP. In the below sections, we describe these steps in terms of the specific questions requested in the RFP for the example symptoms.

### 6.b.2. How the "Sensory Neuropathy" question would be asked of patients: question design

*Determination of What Concepts Are Appropriate for Measuring Sensory Neuropathy in the PRO-CTCAE.*

- *Approach*

  As discussed previously in the fatigue example, when developing any symptom item or instrument, it is fundamental to delineate and justify what concept(s) (eg, intensity, interference, bother, periodicity) are appropriate for measurement of that symptom in the

context of interest. As described previously in the "Substantive Approach to Task 2: Item Development" section of this application, we propose a consensus process early in the contract period to address this question, based on best available evidence.

- *Analysis of Concepts in the CTCAE V3 Neuropathy (Sensory) Item*

The neuropathy (sensory) item in CTCAE v3, as shown in Table A.6-7 asks clinical staff respondents to consider both severity (asymptomatic, sensory alteration, paresthesia) and interference (interference with function, interference with ADL, disabling) during grading.

**Table A.6-7 Current CTCAE v3 Neuropathy (Sensory) Item**

| CTCAE v3 TERM | GRADE 1 | GRADE 2 | GRADE 3 | GRADE 4 |
|---|---|---|---|---|
| **Neuropathy (sensory)** | Asymptomatic; loss of deep tendon reflexes or paresthesia (including tingling) but not interfering with function | Sensory alteration or paresthesia (including tingling), interfering with function but not interfering with ADL | Sensory alteration or paresthesia interfering with ADL | Disabling |

It is not clear whether clinical staff or patient respondents are capable of discerning between interference with "function" vs. interference with "ADL" (as previously discussed in the "Rationale for Approach" section of this proposal). As noted in the fatigue example, we believe that this double-barreling approach "forces" patients into categories they may not see themselves as fitting, and should be abandoned in the PRO-CTCAE.

*Optional Study: Cognitive Interviews to Evaluate Patient Understanding of Double-Barreled Items*. See Appendix 2.

*Determination of Number of Items Necessary to Measure Sensory Neuropathy in the PRO-CTCAE.*

- *Approach*

When designing the PRO-CTCAE, we must balance the desire to aggregate as much information as possible about a patient's experience with the necessity to minimize questionnaire burden on patients. Since our intention is to measure multiple symptoms, it is not ideal to always administer multiple items for each symptom. We are cognizant that the purpose of the PRO-CTCAE is adverse event monitoring and not in-depth symptom characterization. As such, a single initial question-per-symptom approach is optimal. Our approach would thus be to screen all applicable symptoms with a single question each, with an option in the PRO-CTCAE platform to designate triggering of additional items of particular interest toward further characterization (see "Substantive Approach to Task 5: Platform Development"). Which specific symptoms would trigger additional items would either be left to the discretion of the clinical investigator at the

time of creating the study-specific electronic CRF or could be built in as a default. We anticipate discussing the alternative models with NCI staff during the contract period.

- *Consideration of Automated Triggering of Additional Items*

Please see discussion in the Example Approach to Fatigue.

*Question Structure*

- *Approach*
  Please see discussion in the Example Approach to Fatigue.

- *Generic Question Formats*
  Please see discussion in the Example Approach to Fatigue. Potential generic question formats are shown in Table A.6-8. Final format will be evaluated and refined through the cognitive interviews proposed in the "Approach to Task 4" section of this proposal.

**Table A.6-8 Generic Question Formats for PRO-CTCAE Items**

| |
|---|
| **How [concept] has your [symptom term(s)] been since [period of recall]:** |
| **What is the worst [concept] your [symptom term(s)] has/have been since [period of recall]?** |

*Identification of Existing Instruments for Use as Potential Sources of Terms or Items for the PRO-CTCAE.*

- *Approach*

For each symptom identified for grading in the PRO-CTCAE (listed in the "Substantive Approach to Task 2" section of this proposal), a literature search and review of existing instruments will be conducted to identify appropriate terms for inclusion in the severity item. For the purpose of this illustration, we will demonstrate how a sensory neuropathy severity item would be created. The first step in this process will be to conduct a literature search to identify all instruments measuring the symptom and concept of interest that have been demonstrated as valid and reliable in cancer populations. Source information on development of the instruments will be sought to ascertain if patient interviews or other methods were used to select (or reject) terms characterizing the symptom of interest. Terms will be selected for which there is the best evidence of correspondence to the symptom phenomenon of interest and that capture responses from as many patients with that symptom as possible while minimizing redundancy.

- *Overview of Literature and Instruments for Measuring Sensory Neuropathy*

Several instruments have been developed and evaluated to measure patient-reported sensory neuropathy, although the context has often been noncancer, for example, diabetic peripheral neuropathy, which has been demonstrated to exhibit different characteristics from chemotherapy-related neuropathy (Malik and Stillman, Curr Neurol Neurosci Rep 2008).In 2007, Bennett et al published a review of available measures of sensory neuropathy, including five different instruments (Bennett et al, Pain 2007).

Single-symptom instruments noted in this review, as well as in our literature review, are included in Table A.6-9, as well as whether they have been developed or specifically evaluated for use in cancer populations.

**Table A.6-9. Sensory Neuropathy Single-Item Scales**

| SCALE | DISEASE | REFERENCES |
|---|---|---|
| FACT-Ntx Subscale Items | Cancer | Calhoun, 2000; Calhoun, 2003 |
| Modified Total Neuropathy Score | Cancer | Wampler, M; Cavaletti, G |
| Neuropathic Pain Questionnaire (NPQ) | Noncancer specific | Backonja, 2004; Krause, 2003 |
| Neuropathy Total Symptom Score | Noncancer specific | Bastyr, 2005 |
| Neuropathic Pain Symptom Inventory | Noncancer specific | Bouhassira, 2004 |
| Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) | Noncancer specific | Bennett, 2007 |

In addition, a systematic review of cancer symptom instruments published in 2006 noted 3 multisymptom assessment tools that include items to assess sensory neuropathy (Kirkova et al, J Clin Oncol 2006). These instruments, and pertinent literature citations, are listed in Table A.6-10.

**Table A.6-10. Multisymptom Instruments That Elicit Sensory Neuropathy**

| |
|---|
| Rotterdam Symptom Checklist (de Haes, 2003) |
| The Memorial Symptom Assessment Scale (Portenoy, 1994) |
| M.D. Anderson Symptom Inventory (Cleeland, 2000) |

- *Identification of Sensory Neuropathy Symptom Terms*

As described above, a challenge in measuring sensory neuropathy is that it is a complex phenomenon that may be experienced differently in different individuals. Therefore, it is essential to identify terms that capture the full constellation of patient experiences yet are specific enough to restrict measurement to the symptom of interest. For assessment of sensory neuropathy in cancer patients, a key issue is whether to include measurement of pain. Because there are many other potential causes of pain besides neuropathy in cancer patients, inclusion of a pain measurement may reduce the specificity of an item. For this reason, we seek in the published literature and in existing data sets to which we have access approaches to selecting terms for measuring neuropathy that will maximize both sensitivity and specificity.

One example of a cancer-specific sensory neuropathy subscale is the FACT-Ntx, although it has not been widely used or evaluated and likely would not be accepted by the FDA as the basis for a labeling claim because of the lack of content validity or specificity of items (Calhoun et al, Proc ASCO 2000; Calhoun et al, Intl Jrl Gyn Cancer 2003). Nonetheless, it represents development work that can be considered during the

creation of the sensory neuropathy PRO-CTCAE item. The individual items in this subscale are shown in Table A.6-11.

**Table A.6-11. FACT-Ntx Subscale Items**

| |
|---|
| I have numbness or tingling in my hands |
| I have numbness or tingling in my feet |
| I feel discomfort in my hands |
| I feel discomfort in my feet |
| I have joint pain or muscle cramps |
| I feel weak all over |
| I have trouble hearing |
| I get a ringing or buzzing in my ears |
| I have trouble buttoning buttons |
| I have trouble feeling the shape of small objects when they are in my hand |
| I have trouble walking |

Review of the pertinent instruments and scientific literature reveals that available terms related to symptom characterization (rather than to interference with function) include *numbness*, *tingling*, *discomfort*, *pain*, *pricking*, *pins and needles*, *shooting*, *electric shocks*, and *hot or burning*. Location in the hands and feet is commonly included to increase specificity. To narrow the field of terms, we consulted an analysis by Dr. Cleeland's group at MDACC in which patients receiving taxanes or cisplatin and known to have peripheral neuropathy were asked to select terms to describe their sensory neuropathy. The patients chose six words on average (±2.1 SDs). Patients taking taxane and cisplatin selected similar descriptors. All patients (100%) in both groups chose *numbness* to describe the area of on-going sensation with the word *tingling* next most commonly chosen (81.8% for taxanes and 77.4% for cisplatin). Other items were found to be redundant information. Notably, pain did not add information to *numbness*, even at the high end of the severity spectrum when patients were experiencing pain. In other words, measurement of numbness and tingling was able to capture the severity related to the pain experience with peripheral neuropathy. Dr. Sloan, in his exploratory work administering PRO sensory neuropathy items in the NCCTG, has also noted that assessment of numbness and tingling captures the pain experience, thus obviating the need to separately ask about pain when measuring sensory neuropathy with a single item in this population. These data provide a rationale for a PRO-CTCAE severity monitoring item that asks patients to assess *numbness* and *tingling* but not *pain*. Ongoing work at MDACC and Mayo in these areas will provide additional data toward the development of this item during the contract period.

### 6.b.3. How the Sensory Neuropathy Item would be scored: response option format

*Categorization of Item by Degree of Subjective Content*. Please see discussion in the Example Approach to Fatigue.

*Response Format Options*. Please see discussion in the Example Approach to Fatigue.

### 6.b.4. Creation of PRO-CTCAE draft sensory fatigue severity item

Based on work described in the above sections and in the prior "Approach to Fatigue" discussion, assuming that planned studies yield the hypothesized results, we anticipate creating a single sensory neuropathy item for the PRO-CTCAE that measures the magnitude of severity of *numbness* and *tingling* and limits this evaluation to the locations of hands and feet. As described previously in the "Approach to Fatigue" section, we aim to determine through a consensus process early in the contract period the optimal response option format for PRO-CTCAE items (NRS vs. adjectival verbal descriptor scale). For this example, we present both the 0-10 NRS and verbal descriptor scale options (we do not include the 0-4 option in this example, but the look would be similar). The potential constructs for the PRO-CTCAE draft item are shown below in Figure A.6-3.

**Figure A.6-3 PRO-CTCAE Draft Sensory Neuropathy Item With Two Alternative Response Format Options (Choice Pending Results of Planned Research)**

**PROPOSED QUESTIONS:**
*(Would be Refined in Cognitive Interviews)*

**How severe has numbness or tingling in your hands or feet been since your last treatment?**
*or:*
**What is the worst numbness or tingling you have had in your hands or feet since your last treatment?**

**0-10 NRS RESPONSE OPTION:**

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| No numbness or tingling | | | | | | | | | | As bad as you can imagine |

**VERBAL DESCRIPTOR SCALE RESPONSE OPTION:**

| None | Mild | Moderate | Severe | Very severe |
|---|---|---|---|---|
| ☐ | ☐ | ☐ | ☐ | ☐ |

An alternative approach we will consider is a “symptom checklist” format in which instructions are provided to the respondent at the beginning of the questionnaire and then each item proceeds with the selected symptom terms only, adjacent to the response option scale. An example for the sensory neuropathy (numbness and tingling) item is shown in Figure A.6-4.

**Figure A.6-4. Possible Alternative Symptom Checklist Style Draft Fatigue Item, With Both Response Format Options**

| | |
|---|---|
| **Numbness or tingling in your hands or feet** | **0 1 2 3 4 5 6 7 8 9 10**<br>No Fatigue — As bad as you can imagine |
| **Numbness or tingling in your hands or feet** | None ☐ Mild ☐ Moderate ☐ Severe ☐ Very severe ☐ |

### 6.b.5. How cultural literacy and minimization of bias of the Sensory Neuropathy (numbness and tingling) Item will be ensured

Please see discussion in the Example Approach to Fatigue.

### 6.b.6. How measurement properties of the Sensory Neuropathy (numbness and tingling) Item will be evaluated

Please see discussion in the Example Approach to Fatigue.

Notably, clinical indicators for sensory neuropathy that will be used as anchors for this analysis will include dose reduction of chemotherapy associated with neuropathy (eg, taxanes, platinums, vinca alkaloids, bortezomib, thalidomide), initiation or increase of pain medications, initiation/increase in supportive medications commonly given for neuropathy (amitriptyline, gabapentin, pregabalin), or interference with sleep.

### 6.b.7. How information would be used to grade an event: mandatory items and cut-point scores for automated notifications

Please see discussion in the Example Approach to Fatigue.

### 6.b.8. How information would be integrated into CTCAE v3 and v4

Please see discussion in the Example Approach to Fatigue.

### 6.b.9. How information would be conveyed electronically

Please see discussion in the Example Approach to Fatigue.

## 7. Schedule: Timeline of Tasks and Deliverables

Previously, in sections A-5.a through A-5.i, we described individual proposed tasks, each accompanied by a schedule of deliverables to NCI staff. In Tables A.7-1, A.7-2, and A.7-3 on the following pages, we summarize these deliverables on a task-by-task basis, as well as the month of planned delivery from the date of authorization to proceed by the NCI Contracting Officer, and personnel responsible for discrete steps of each proposed task. These months are consistent with the schedule in *ARTICLE F.2 DELIVERIES* in the RFP. In several instances, we have added additional deliverables beyond those requested in the RFP, which we feel are logical to include in order to successfully achieve the overall objectives of the project.

As described elsewhere in this proposal, Dr. Basch and other members of the investigative team have track records of delivering high-quality products on rigid timelines. Dr. Basch has extensive past experience coordinating complex projects with numerous deliverables and partners. Prior to his career in oncology research, he built and ran a highly successful healthcare database company with numerous employees, clients, and aggressive timelines for deliverables. More recently, he oversaw development of the informatics and logistical infrastructure for the 10-institution DOD-funded Prostate Cancer Clinical Trials Consortium. As a faculty member at MSKCC, he has quickly built a multifaceted research program in an area directly pertinent to the proposed work. Therefore, he has an in-depth understanding of the facets involved with this work, as well as an appreciation of the necessary steps and support staff required to accomplish the objectives.

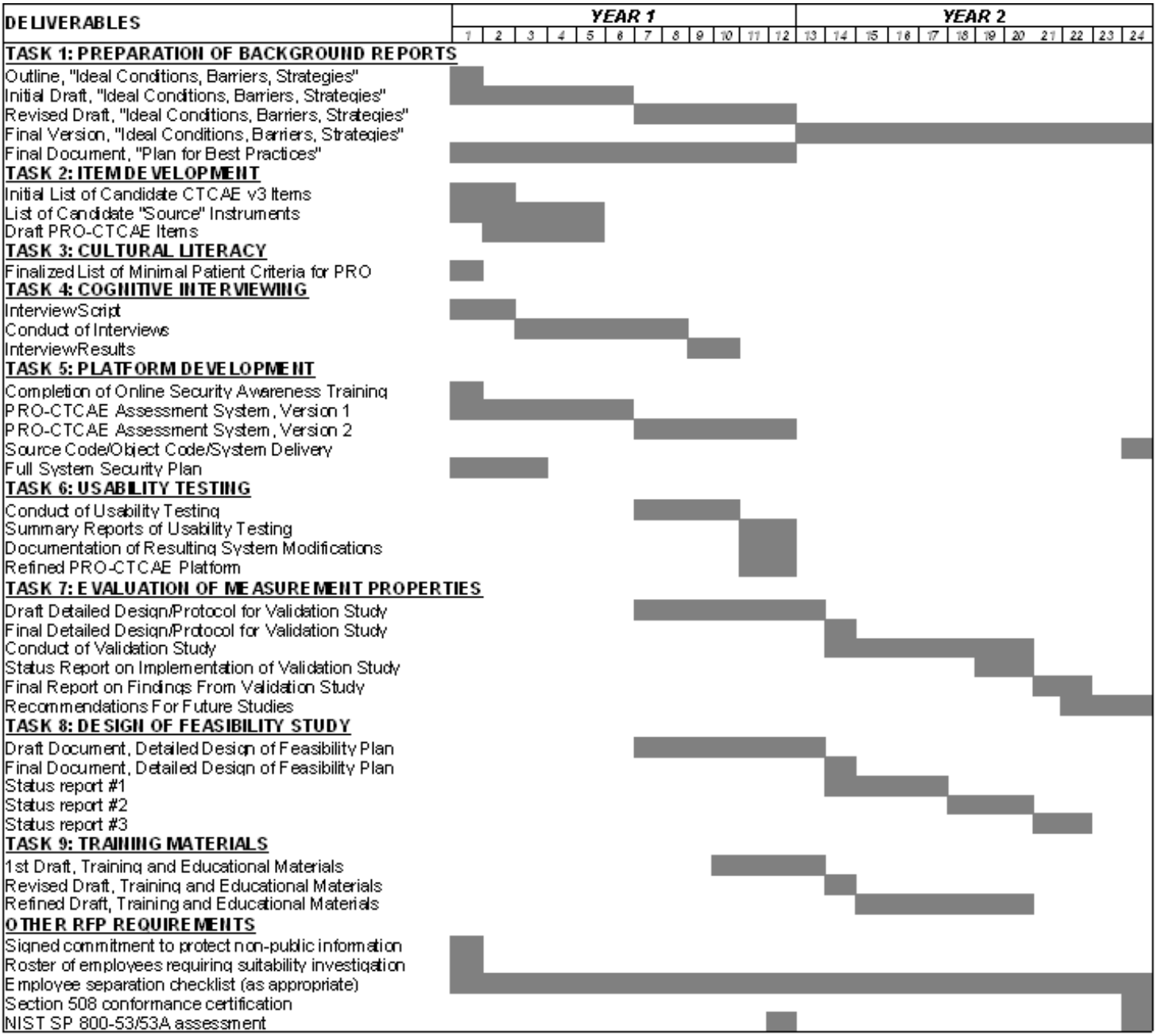
The RFP requests completion of a complex series of interrelated tasks on a short timeline, and Dr. Basch is uniquely qualified to successfully deliver product on time. He has the necessary organizational background, core skills and experience, and relationships across multiple disciplines. He has put together a strong administrative and data management team at MSKCC which will constantly communicate with all research and other pertinent personnel in order to assure adherence to timelines. Intentional redundancy in task leadership is built into the project to assure success. Our plan for communications between investigators and with NCI staff and advisors is further described in section B-10.

**Table A.7-1. Schedule of Deliverables by Month from Initiation of Contract**

| Task | Deliverable | Month |
|---|---|---|
| **Task 1: Background Reports** | Outline, "Ideal Conditions, Potential Barriers, Strategies to Implement the PRO-CTCAE" | 1 |
| | Initial Draft, "Ideal Conditions, Potential Barriers, Strategies to Implement the PRO-CTCAE" | 6 |
| | Revised Draft, "Ideal Conditions, Potential Barriers, Strategies to Implement the PRO-CTCAE" | 12 |
| | Final Version, "Ideal Conditions, Potential Barriers, Strategies to Implement the PRO-CTCAE" | 24 |
| | Final Document, "Plan for Best Practices in Community/Comprehensive Cancer Care Settings" | 12 |
| **Task 2: Item Development** | List of Candidate CTCAE v3 Items for PRO-CTCAE | 2 |
| | List of Candidate "Source" Instruments and Items for PRO-CTCAE | 4-5 |
| | Draft PRO-CTCAE Items | 5 |
| **Task 3: Cultural Literacy** | Finalized List of Minimal Criteria (Age, Educational Level, Cognitive Ability) for Collecting PROs | 1 |
| **Task 4: Cognitive** | Interview Script to Evaluate Draft PRO-CTCAE Items | 2 |
| | Interview Results (Evaluation of Draft PRO-CTCAE Items) | 10 |
| **Task 5: Platform Development** | Completion of Online Security Awareness Training by All Relevant Project Participants | 1 |
| | PRO-CTCAE Assessment System, Version 1 | 6 |
| | PRO-CTCAE Assessment System, Version 2 | 12 |
| | Source Code/Object Code/System Delivery | 24 |
| | Full System Security Plan | 3 |
| **Task 6: Usability Testing of PRO-CTCAE Platform in Patients and** | Summary Reports of Usability Testing | 12 |
| | Documentation of Resulting System Modifications | 12 |
| | Refined PRO-CTCAE Platform | 12 |
| **Task 7: Evaluation of Measurement Properties** | Draft Detailed Design/Protocol for Validation Study | 13 |
| | Final Detailed Design/Protocol for Validation Study | 14 |
| | Status Report on Implementation of Validation Study | 20 |
| | Final Report on Findings From Validation Study | 22 |
| | Recommendations For Future Studies | 22 |
| **Task 8: Feasibility Design** | Draft Document, Detailed Design of Feasibility Studies Plan | 13 |
| | Final Document, Detailed Design of Feasibility Studies Plan | 14 |
| | Status report #1, Implementation of the Validation/Feasibility Studies Plan | 17 |
| | Status report #2, Implementation of the Validation/Feasibility Studies Plan | 20 |
| | Status report #3, Implementation of the Validation/Feasibility Studies Plan | 22 |
| **Task 9: Training Materials** | 1st Draft, Training and Educational Materials for PRO-CTCAE Assessment System | 13 |
| | Revised Draft, Training and Educational Materials for PRO-CTCAE Assessment System | 14 |
| | Refined Draft, Training and Educational Materials for PRO-CTCAE Assessment System | 20 |

| Other RFP Requirements | |
|---|---|
| **Signed, Witnessed Commitment to Protect Non-Public Information** | Prior to month 1 |
| **Roster of Employees Requiring Suitability Investigations** | 15 days |
| **Employee Separation Checklist** | Within 7 calendar days of the termination/separation of an employee |
| **Section 508 Conformance Certification** | Completion date of contract |
| **NIST SP 800-53/53A Assessment** | 12 (resubmit annually) |

## Table A.7-2. Schedule of Deliverables by Month From initiation of Contract, in Gantt Format



## Table A.7-3. Personnel Responsible for Specific Steps of Each Proposed Task, and Responsibilities

| TASK 1: PREPARE BACKGROUND REPORT AND BEST PRACTICES DOCUMENT | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***RESPONSIBILITIES*** |
| Task oversight | Dr. Deb Bruner | As Task Leader, in collaboration with the PI and his team, Dr. Bruner will oversee the timeline and preparation of deliverables for this task. |
| Literature searches | Dr. Ethan Basch<br>Dr. Tito Mendoza<br>Dr. Charlie Cleeland<br>Dr. Jeff Sloan | The pertinent literature searches were previously conducted for preparation of the Technical Proposal, and were conducted primarily by Drs. Basch, Mendoza, Cleeland, Sloan, and Atkinson. These citations are included in Section E of the Technical Proposal. |
| Key informant interviews | Dr. Deb Bruner<br>Dr. Deb Schrag<br>Dr. Ethan Basch<br>Laura Sit | Interviews will be conducted by Drs. Bruner, Schrag, and Basch at cooperative group meetings, and by Ms. Sit by telephone (described in Section A.5.2 of the Technical Proposal). |
| Project management | Ms. Laura Sit | Ms. Sit, the MSKCC project manager, will provide project management support, organize interview files, and code interviews. |
| Preparation of draft, revised, and final report | Dr. Deb Bruner | Dr. Bruner will adapt content from the Technical Proposal and summarize interview results to prepare the background report. |
| Preparation of plan for best practices | Dr. Ethan Basch | Dr. Basch will draft the best practices document based on NCI input following award, with substantive input from the investigative team and advisors. |

| TASK 2: PRO-CTCAE ITEM DEVELOPMENT | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***ROLE/DESCRIPTION*** |
| Task oversight | Dr. Charlie Cleeland<br>Dr. Ethan Basch | As Task Leader, in collaboration with the PI and his team, Dr. Cleeland will oversee the timeline, study designs, and deliverables for this task. |
| Identify CTCAE items amenable to PRO | Dr. Ethan Basch | This has already been done, and is shown in Table A.5-5 of the Technical Proposal. |
| Identify potential source instruments/items | Dr. Tito Mendoza<br>Dr. Ethan Basch<br>Dr. Charlie Cleeland | Dr. Mendoza will direct this effort and will be assisted by Drs. Basch and Cleeland to assemble items from identified instruments/questionnaires (and supporting literature) for potential use in the PRO-CTCAE, as described in Section 5.b.8 of the Technical Proposal. |
| Create draft PRO-CTCAE items | Dr. Ethan Basch<br>Dr. Charlie Cleeland<br>Dr. Jeff Sloan | Drs. Basch, Cleeland, and Sloan will draft the initial PRO-CTCAE items as described in Section 5.b.10 of the Technical Proposal. This will be done with extensive input from other team members, advisors, and NCI staff. |
| Identification of ideal patient populations for evaluations | Dr. Deb Schrag<br>Dr. Ethan Basch | This has already been done, and is described in Section 5.b.12 of the Technical Proposal. |
| Determination of minimal patient criteria | Dr. Vish Viswanath | Described in Section 5.c.3 of the Technical Proposal. |
| Project management | Ms. Laura Sit | Ms. Sit will provide project management support and organize data/files. |

| TASK 3: ASSURE CULTURAL LITERACY & LACK OF BIAS | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***ROLE/DESCRIPTION*** |
| Task oversight | Dr. Vish Viswanath | As Task Leader, in collaboration with the PI and his team, Dr. Viswanath will oversee the timeline and deliverables for this task. |
| Assurance of health literacy | Dr. Vish Viswanath | As described in Section 5.c.2, Dr. Viswanath will interface with leaders of appropriate tasks to assure health literacy is considered during item development and evaluation. |
| Minimization of cultural bias | Dr. Vish Viswanath | As described in Section 5.c.3, Dr. Viswanath will interface with leaders of appropriate tasks to assure that potential cultural bias is minimized during item development and evaluation. |
| Determination of minimal patient criteria for PRO | Dr. Vish Viswanath<br>Dr. Jennifer Hay<br>Dr. Ethan Basch | As described in Section 5.c.3, Dr. Viswanath and members of his Health Communications Core at DFCI will work with Drs. Hay and Basch to develop minimal patient criteria for the collection of PRO, including consideration of age, education level, and cognitive ability, which will be reviewed with NCI. |

| TASK 4: CONDUCT COGNITIVE INTERVIEWS | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***ROLE/DESCRIPTION*** |
| Task oversight | Dr. Jennifer Hay | As Task Leader, in collaboration with the PI and his team, Dr. Hay will oversee the timeline, study design, and deliverables for this task. |
| Interviewing script design | Dr. Jennifer Hay | Dr. Hay will adapt a script that she and Dr. Basch have previously developed. |
| Cognitive interviews | Larry Shiman | Larry Shiman and his Data Technologies Core at DFCI will conduct and record the interviews. |
| Summary report | Larry Shiman | Larry Shiman and his staff will summarize/tabulate the interview results. |
| Analysis | Dr. Jennifer Hay | Dr. Hay will analyze the results of interviews towards item modification. |
| Report | Dr. Jennifer Hay<br>Dr. Ethan Basch | Drs. Hay and Basch will prepare a report of the interview results and recommended modifications to PRO-CTCAE items. |
| Project management | Ms. Laura Sit | Ms. Sit will provide project management support and organize files. |
| Data management | Ms. Rachel Jia | Ms. Jia will maintain and clean data as appropriate. |

| TASK 5: DEVELOP PRO-CTCAE SOFTWARE SYSTEM | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***ROLE/DESCRIPTION*** |
| Task oversight | Mr. Ram Chiukuri<br>Mr. Vinay Kumar<br>Mr. Marwan Shouery<br>Dr. Ethan Basch | Mr. Chilukuri and Mr. Vinay will oversee all system design on the SemanticBits side, whereas Mr. Shouery and Dr. Basch will closely monitor all progress. |
| Platform development | Mr. Ram Chilukuri<br>SB Programmers | Mr. Chilukuri will oversee the team of 1.5 programmer FTEs to develop the publishing tool, clinician interface, patient interface, alert system, study calendar integration, and caAERS integration. |
| System refinement | Mr. Ram Chilukuri<br>SB Programmers<br>Mr. Marwan Shouery | Mr. Chilukuri will oversee refinement of the system based on investigative team and NCI feedback, as well as based on results of cognitive interviews and usability testing. Mr. Shouery will QA this process. |
| Security, privacy, interoperability, regulatory compliance | Mr. Ram Chilukuri<br>SB Programmers<br>Mr. Marwan Shouery | Mr. Chilukuri will oversee assurance of compliance, and Mr. Shouery will QA this process. |
| Delivery of source code/system to NCI | Mr. Ram Chiukuri<br>Mr. Vinay Kumar | Mr. Chilukuri and Mr. Vinay will assure delivery of the system and code to NCI according to the specified schedule. |

| TASK 6: USABILITY TESTING | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***ROLE/DESCRIPTION*** |
| Task oversight | Dr. Amy Abernethy | As Task Leader, in collaboration with the PI and his team, Dr. Abernethy will oversee the timeline, study design, and deliverables for this task. |
| Study design | Dr. Amy Abernethy | Dr. Abernethy will design this study and write the protocol. |
| Study coordination/project management | Ms. Mary Miller<br>Ms. Alex Dupont<br>Ms. Laura Sit | Preparation of IRB documents, study logistics, accrual, distribution of wireless tablet computers, and forms will be coordinated by Ms. Miller and Ms. Dupont from Dr. Abernethy's team at Duke, with oversight by the MSKCC project manager, Ms. Sit. |
| Data/statistical support | Ms. April Coan<br>Ms. Rachel Jia | Data aggregation, data cleaning, and statistical analysis will be conducted by Ms. Coan from Dr. Abernethy's team, with oversight by Ms. Jia from MSKCC. |
| Analysis | Dr. Amy Abernethy<br>Dr. Ethan Basch | Drs. Abernethy and Basch will analyze the data. |
| Report/recommendations for system modifications | Dr. Amy Abernethy<br>Dr. Ethan Basch | Drs. Abernethy and Basch will prepare a report of findings, as well as recommendations for system modifications based on results. |

| TASK 7: EVALUATION OF ITEM MEASURMENT PROPERTIES | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***ROLE/DESCRIPTION*** |
| Task oversight | Dr. Charlie Cleeland<br>Dr. Jeff Sloan<br>Dr. Ethan Basch | The Task Leaders will collaboratively oversee the timeline, study design, and deliverables for this task, with Dr. Basch taking primary responsibility to assure compliance with timelines. |
| Identification of study populations | Dr. Deb Schrag | Dr. Schrag has previously analyzed cooperative group accrual data and published clinical trials to identify study populations which represent cooperative group trial enrollees, and which provide acces to patients experiencing symptoms of interest, as described in Section 5.g.2. |
| Study design and protocol preparation | Dr. Charlie Cleeland<br>Dr. Jeff Sloan<br>Dr. Ethan Basch<br>Dr. Tito Mendoza | Dr. Basch will assume primary responsibility for the drafting of the study protocol, with substantive input from Drs. Cleeland, Sloan, and Mendoza. |
| Project management | Ms. Laura Sit | Ms. Sit will provide project management support to coordinate IRB review at participating sites; training of local research assistants and clinical staff; distribution of wireless laptop computers; assurance of compliance with forms; and assurance of study timelines. |
| Data management/ statistical support | Ms. Rachel Jia<br>Dr. Thomas Atkinson | Central data collection at MSKCC will be coordinated by Ms. Jia. Ms. Jia and Dr. Atkinson will provide data cleaning and statistical support. |
| Analysis | Dr. Charlie Cleeland<br>Dr. Jeff Sloan<br>Dr. Ethan Basch | Drs. Cleeland, Sloan, and Basch will be primarily responsible for data analysis. |
| Report and recommendations | Dr. Charlie Cleeland<br>Dr. Jeff Sloan<br>Dr. Tito Mendoza<br>Dr. Ethan Basch | Drs. Cleeland, Sloan, Mendoza, and Basch will prepare a report of results, with recommendations for future assessments or system modifications. Dr. Basch will assume primary responsibility to assure timely preparation and delivery of the report and recommendations. |

| TASK 8: DESIGN FEASIBILITY STUDY | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***ROLE/DESCRIPTION*** |
| Task oversight | Dr. Deb Schrag<br>Dr. Ethan Basch<br>Dr. Deb Bruner<br>Dr. Andy Trotti | Drs. Schrag and Basch will assume primary responsibility for the timeline, study design, and deliverables of this task. |
| Identification of cooperative group trials amenable to linkage to PRO study | Dr. Deb Schrag<br>Dr. Ethan Basch<br>Dr. Deb Bruner<br>Dr. Andy Trotti | The study team will collaborate with project advisors, cooperative group committees, and cooperative group protocol editors to identify upcoming protocols in selected disease groups for linkage to the planned PRO-CTCAE companion/correlative study. |
| Amendment of CALGB 70501 to incorporate PRO-CTCAE system | Dr. Ethan Basch | Dr. Basch will amend CALGB protocol 70501, his current PRO study in the CALGB, to include administration of PRO-CTCAE items via the new PRO-CTCAE platform, as described in Section 5.h.5. |
| Preparation of reports | Dr. Deb Schrag | Dr. Schrag will prepare the reports of progress. |
| Project management | Ms. Laura Sit | Ms. Sit will provide project management support. |
| Data management | Ms. Rachel Jia | Data aggregation will likely be coordinated by CALGB, although statistical support will be provided at MSKCC by Ms. Jia. |

| TASK 9: CREATION OF TRAINING MATERIALS | | |
|---|---|---|
| ***SPECIFIC STEPS*** | ***PERSONNEL*** | ***ROLE/DESCRIPTION*** |
| Preparation of technical "user manual" for software installation/maintenance | Mr. Ram Chilukuri | Under Mr. Chilukuri's supervision, a technical manual aimed for use by IT professionals installing the PRO-CTCAE system at local institutions will be developed. Mr. Shouery and Dr. Basch will QA this process and manual. |
| Preparation of paper and electronic patient training materials | Dr. Vish Viswanath | As described in Section 5.i, Dr. Viswanath and members of his Health Communications Core at DFCI will create patient training materials, which will be available in paper versions, and electronically via download or CD. |
| Preparation of investigator paper and electronic training materials | Dr. Vish Viswanath | Dr. Viswanath and members of his Health Communications Core at DFCI will create instructional materials for investigators using the PRO-CTCAE system for clinical research. |
| Preparation of clinical staff paper and electronic training materials | Dr. Vish Viswanath | Dr. Viswanath and members of his Health Communications Core at DFCI will create instructional materials for clinical staff following patients enrolled in studies which use the PRO-CTCAE system. |

## E. LITERATURE CITED

Acquadro C, Berzon R, Dubois D, et al. Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. *Value Health*. 2003; 6(5):522-531.

Alberts DS, Noel JK. Cisplatin-associated neurotoxicity: can it be prevented? Anticancer Drugs 1995;6:369-383. PMID: 7670134

Anastasi A, Urbina S. *Psychological Testing*, 7th ed. Upper Saddle River, NJ: Prentice Hall, 1996.

Anderson KO: Role of cutpoints: why grade pain intensity? *Pain*. 2005;113:5-6.

Anderson KO, Syrjala KL, Cleeland CS. Assessment of cancer pain. In Turk DC, Melzack R, eds. *Handbook of Pain Assessment*, 2nd ed. New York: Guilford Press, 2001; pp. 579-600.

Armstrong T, Mendoza T. Validation of MDASI-Brain Tumor. *J Neuro-Oncol*. 2006;80:27-35.

Backonja M, Stacey B. Neuropathic pain Symptoms Relative to Overall Pain Rating. J Pain, 2004;5(9): 491-497

Barnes EA, Bruera E. Fatigue in patients with advanced cancer: a review.

Barresi MJ, Shadbolt B, Byrne D, et al: The development of the Canberra Symptom Score Card: A tool to monitor the physical symptoms of patients with advanced tumours. BMC Cancer 3:32, 2003

Basch E, Artz D, Dulko D, et al. Patient online self-reporting of toxicity symptoms during chemotherapy. *J Clin Oncol*. 2005;23(15):3552-3561.

Basch E. Does toxicity reporting by patients or clinicians predict clinical outcomes? Podium Presentation at 2008 ASCO Annual Meeting, Chicago, IL.

Basch E, Iasonos A, Barz A, et al. Long-term toxicity monitoring via electronic patient-reported outcomes in patients receiving chemotherapy. *J Clin Oncol*. 2007;25(34):5374-5380.

Basch E. Automated collection of patient symptom data during treatment. Invited lecture at 2005 group meeting of the Cancer and Leukemia Group B, Amelia Island, FL.

Basch E, Iasonos A, McDonough T, et al. Clinician versus patient self-reporting of symptoms during cancer treatment: a paired analysis using the National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE). *Lancet Oncol*. 2006;7(10):903-909.

Basch E, Artz D, Speakman J, et al. Design, implementation, and evaluation of a web portal for patient symptom monitoring. *J Am Med Inform Assoc*. 2007;14(3):264-268.

Basch E, Artz D, Dulko D, et al. Patient online self-reporting of toxicity symptoms during chemotherapy. Presented at the 2005 Annual Meeting of the American Society of Clinical Oncology, Orlando, FL.

Basch EM, Thaler HT, Shi W, Yakren S, Schrag D. Use of information resources by patients with cancer and their companions. *Cancer*. 2004;100(11):2476-2483.

Basch E. Methods used to assess and report pain-related endpoints in clinical trials. Invited presentation at FDA Oncology Division Advisory Committee (ODAC) Meeting, on behalf of the FDA Oncology Division. Rockville, MD, June 24, 2007.

Basch E, Barz A, Iasonos A, et al. Online reporting of toxicity symptoms by lung cancer patients during chemotherapy. Presented at 2006 annual U.S. meeting of the International Society for Pharmacoeconomics and Outcomes Research, Philadelphia, PA.

Basch E, Iasonos A, Speakman J, et al. Toxicity monitoring via electronic patient reported outcomes in end-stage chemotherapy patients. Presented at the 2007 annual meeting of the American Society of Clinical Oncology (ASCO), Chicago, IL.

Basch E, Iasonos A, Barz A, Kris MG, Scher HI, Schrag D. Long-term toxicity monitoring via electronic patient-reported outcomes in patients receiving chemotherapy. Presented at the 2007 ISOQOL Conference on Patient Reported Outcomes in Clinical Practice, Budapest, Hungary.

Basch E, McDonough T, Iasonos A, et al. Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events. Presented at 2006 annual meeting of the American Society of Clinical Oncology, Atlanta, GA.

Basch EM, Jia X, Barz A, et al. Does toxicity symptom reporting by patients or clinicians predict clinical outcomes? Presented at annual meeting of the American Society of Clinical Oncology, Chicago, IL, June 2, 2008.

Bastyr EJ, Price KL, Bril, V. Development and Validity of the Neuropathy Total Pain Score. Clin Ther 2005;27:1278-1294

Beck SL, Falkson G. Prevalence and management of cancer pain in South Africa. *Pain*. 2001;94(1):75-84.

Bennett CL, Nebeker JR, Yarnold PR, et al. Evaluation of serious adverse drug reactions: a proactive pharmacovigilance program (RADAR) vs safety activities conducted by the Food and Drug Administration and pharmaceutical manufacturers. *Arch Intern Med*. 2007;167(10):1041-1049.

Bennett MI, Attal N, Backonja MM, Baron R, Bouhassira D, Freynhagen R, Scholz J, Tölle TR, Wittchen HU, Jensen TS. Using screening tools to identify neuropathic pain. Pain. 2007 Feb;127(3):199-203.

Benoit A. Patient-reported outcomes and clinical practice. From Measurement instruments to decision tools:much more than a simple change in format. *Patient Reported Outcomes Newsletter*. 2006;36.

Bentzen SM, Dorr W, Anscher MS, Denham JW, Hauer-Jensen M, Marks LB, et al. Normal tissue effects: reporting and analysis. Semin Radiat Oncol 2003;13(3):189-202.

Berger T, Malayeri R, Doppelbauer A, et al. Neurological monitoring of neurotoxicity induced by paclitaxel/cisplatin chemotherapy. Eur J Cancer 1997;33(9):1393-1399. PMID: 9337680

Bland JM, Altman DG. Comparing two methods of clinical measurement: a personal history. *Internat J Epidemiol*. 1995;24(3)7-14.

Bogen K. The effect of questionnaire length on response rates: a review of the literature. In *1996 Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association, 1996; pp. 1020-1025.

Bouhassira D, Attal N, Fermanian J, Alchaar H, Gautron M, Masquelier E, Bouvette M, Fotherfill-Bourbonnais F, Perreault A: Implementation of the pain and symptom assessment record (PSAR). J Adv Nurs 40:685-700, 2002 77.

Breitbart W, McDonald MV, Rosenfeld B, et al. Pain in ambulatory AIDS patients. I: Pain characteristics and medical correlates. *Pain*. 1996;68(2-3):315-321.

Bretscher M, Rummans T, Sloan J, et al. Quality of life in hospice patients: a pilot study. *Psychosomatics*. 1999;40:309-313.

Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. *Pain*. 2008; Apr 30. [Epub ahead of print]

Broemeling LD, Cleeland CS. Clinical factors associated with cancer-related fatigue in patients being treated for leukemia and non-Hodgkin's lymphoma. J Clin Oncol. 2002 Mar 1;20(5):1319-28. PMID: 11870175

Brooks R, Rabin R, de Charro F, eds. The Measurement and Valuation of Health Status Using EQ-5D: A European Perspective. Norwell, MA: Kluwer Academic Publishers, 2003.

Brundage M, Feldman-Stewart D, Leis A, et al. Communicating quality of life information to cancer patients: a study of six presentation formats. *J Clin Oncol*. 2005:23(28):6949-6956.

Bruner DW. Should patient-reported outcomes be mandatory for toxicity reporting in cancer clinical trials? *J Clin Oncol*. 2007;25(34):5345-5347.

Bruner DW, Bryan CJ, Aaronson N, et al. Issues and challenges with integrating patient-reported outcomes in clinical trials supported by the National Institute-Sponsored Clinical Trials Networks. *J Clin Oncol*. October, 2007.

Bruner DW, Movsas B, Konski A, et al. Outcomes research in cancer clinical trial cooperative groups: the RTOG model. *Qual Life Res*. 2004 Aug;13(6):1025-1041.

Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol.1993;46(5):423-429.

Calhoun, E.A., Fishman, D.A., Roland, P.Y., Lurain, J.R., Chang, C.-H., & Cella, D. (2000). Validity and selective sensitivity of the FACT/GOG-Ntx. Proceedings of the American Society of Clinical Oncology, 19, 446a.

Calhoun, E.A., Welshman, E.E., Chang, C.H., Lurain, J.R., Fishman, D.A., Hunt, T., & Cella, D. Psychometric Evaluation of the Functional Assessment of Cancer Therapy/Gynecologic Oncology Group-Neurotoxicity (FACT/GOG-Ntx) questionnaire for patients receiving systemic chemotherapy. Submitted, International Journal of Gynecologic Cancer, April, 2003.

Cancer Institute-Common Toxicity Scale.J Peripher Nerv Syst. 2007 Sep;12(3):210-5.

Cancer-related fatigue: the scale of the problem. Oncologist. 2007;12 Suppl 1:4-10. Review. PMID: 17573451

Caraceni A, Mendoza TR, Mencaglia E, et al. A validation study of an Italian version of the Brief Pain Inventory (Breve Questionario per la Valutazione del Dolore). *Pain*. 1996;65(1):87-92.

Casey EB, Jellife AM, Le Quesne PM, Millet YL. Vincristine neuropathy. Clinical and electrophysiological observations. Brain 1973;96:69-86. PMID: 4348690

Cavaletti G, Bogliun G, Marzorati L, et al. Peripheral neurotoxicity of taxol in patients previously treated with cisplatin. Cancer 1995;75:1141-1150. PMID: 7850713

Cavaletti G, Frigeni B, Lanzani F, Piatti M, Rota S, Briani C, Zara G, Cleeland CS, Mendoza TR, Wang XS, et al: Assessing symptom distress in cancer patients: The M.D. Anderson Symptom Inventory. Cancer 89: 1634-1646, 2000

Cheung YB, Goh C, Thumboo J, Khoo KS, Wee J. Quality of life scores differed according to mode of administration in a review of three major oncology questionnaires. *J Clin Epidemiol*. 2006;59(2):185-191.

Cheung Y-B, Wong L-C, Tay M-H, et al. Order effects in the assessment of quality of life in cancer patients. *Qual Life Res*. 2004:13:1217-1223.

Chow E, Doyle M, Li K, et al. Mild, moderate, or severe pain categorized by patients with cancer with bone metastases. *J Palliat Med*. 2006;9(4):850-854.

Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol*. 1990;43(6):551-558.

Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med*. Singapore. 1994;23(2):129-138.

Cleeland CS, Syrjala KL. How to assess cancer pain. In Turk D, Melzack R, eds. *Pain Assessment*. New York: Guilford Press, 1992; pp. 360-387.

Cleeland CS, Wang XS. Measuring and understanding fatigue. Oncology 13(11A): 91-97, 1999.

Cleeland CS. Measurement of pain by subjective report. In Chapman CR, Loeser JD, eds. *Issues in Pain Measurement*. New York: Raven Press, 1989; pp. 391-403.

Cleeland CS. Assessment of pain in cancer: measurement issues. In Foley KM, Bonica JJ, Ventafridda V, eds. *Proceedings of the Second International Congress on Cancer Pain*. New York: Raven Press, 1990; pp. 47-55.

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

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.

Cleeland CS, Ladinsky JL, Serlin RC, Nugyen CT. (1988). Multidimensional measurement of cancer pain: comparisons of US and Vietnamese patients. *J Pain Symptom Manage*. 1988;3(1):23-27.

Cleeland CS, Mendoza TR, Wang XS, et al. Assessing symptom distress in cancer patients: the M. D. Anderson Symptom Inventory. *Cancer*. 2000;89(7):1634-1646.

Cleeland CS, Rotondi A, Brechner T, et al. A model for the treatment of cancer pain. *J Pain Symptom Manage*. 1986;1(4):209-215.

Cleeland CS, Serlin RC, Nakamura Y, Mendoza TR. Effects of culture and language on ratings of cancer pain and patterns of functional interference. In Jensen TS, Turner JA, Wiesenfeld-Hallin Z, eds. *Progress in Pain Research and Management*. Seattle, WA: IASP Press, 1997; pp. 35-51.

Cleopas A, Kolly V, Perneger TV. Longer response scales improved the acceptability and performance of the Nottingham Health Profile. *J Clin Epidemiol*. 2006;59(11):1183-90.

Cleveland WS, Grosse E. Computational methods for local regression. *Statistics Comput*. 1991;1:47-62.

Cohen J. A coefficient of agreement for nominal scales. *Educat Psychol Measure*.1960;20:37-46.

Comput Methods Programs Biomed. 1996 Nov;51(3):211-6. PMID: 8955589

Cornblath DR, Chaudhry V, Carter K, Lee D, Seysedadr M, Miernicki M, Joh T.Total neuropathy score: validation and reliability study. Neurology. 1999 Nov 10;53(8):1660-4. PMID: 10563609

Crane PK, Gibbons, LE, Ocepek-Welkson K, Cook K, Cella D, Narasimhalu K, Hays RD, Teresi JA. A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Qual Life Res*. 2007;16:69-84.

Daut RL, Cleeland CS. The prevalence and severity of pain in cancer. *Cancer*. 1982;50(9):1913-1918.

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.

de Haes JC, van Knippenberg FC, Neijt JP: Measuring psychological and physical distress in cancer patients: Structure and application of the Rotterdam Symptom Checklist. Br J Cancer 62: 1034-1038, 1990 Oxford University Press, 2003, pp 4-13

Dihle A, Helseth S, Paul SM, Miaskowski C.T he exploration of the establishment of cutpoints to categorize the severity of acute postoperative pain. *Clin J Pain*. 2006;22(7):617-624.

Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *J Adv Nurs*. 2003;42(1):57-63.

DuPen SL, DuPen 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.

Economics and Statistics Administration, National Telecommunications and Information Administration. A Nation Online: Entering the Broadband Age. Washington, DC: U.S. Department of Commerce, 2004.

Escalante CP, Grover T, Johnson BA, Harle M, Guo H, Mendoza TR, Rivera E, Ho V, Lee EL, Cleeland CS. A fatigue clinic in a comprehensive cancer center: design and experiences. Cancer. 2001 Sep 15;92(6 Suppl):1708-13. PMID: 11598891

Evans S, Fishman B, Spielman L, Haley A. Randomized trial of cognitive behavior therapy versus supportive psychotherapy for HIV-related peripheral neuropathic pain. *Psychosomatics*. 2003;44(1):44-50.

Ewing G, Todd C, Rogers M, et al: Validation of a symptom measure suitable for use among palliative care patients in the community: CAMPAS-R. J Pain Symptom Manage 27:287-299, 2004

Fairclough DL. Design and Analysis of Quality of Life Studies in Clinical Trials. Boca Raton: CRC Press; 2002.

Farley DO, Elliott MN, Short PF, et al: Effect of CAHPS performance information on health plan choices by Iowa Medicaid beneficiaries. *Med Care Res Rev*. 2002;59:319-336.

Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol*. 1990;43(6):543-549.

Ferrans CE. Definitions and conceptual models of quality of life. In: Lipscomb J, Gotay CC, Snyder C, eds. *Outcomes Assessment in Cancer: Measures, Methods, and Applications*. Cambridge: Cambridge University Press, 2005.

Ferraz MB, Quaresma MR, Aquino LR, Atra E, Tugwell P, Goldsmith CH. Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. *J Rheumatol*. 1990;17(8):1022-1024.

Flachenecker P, Kümpfel T, Kallmann B, Gottschalk M, Grauer O, Rieckmann P, Trenkwalder C, Toyka KV. Fatigue in multiple sclerosis: a comparison of different rating scales and correlation to clinical parameters. Mult Scler. 2002 Dec;8(6):523-6. PMID: 12474995

Fleiss JL. *Statistical Methods for Rates and Proportions*, 2nd ed. New York: Wiley, 1981.

Food and Drug Administration. *Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.* Rockville, MD: Food and Drug Administration, 2006.

Forman A. Peripheral neuropathy in cancer patients: Clinical types, etiology and presentation. Oncology 1990;4(2):85-89. PMID: 2167114

Forsyth PA, Balmaceda C, Peterson K, et al. Prospective study of paclitaxel-induced peripheral neuropathy with quantitative sensory testing. J Neuro-Oncol 1997;35:47-53. PMID: 9266440

Forsyth PA, Balmaceda C, Peterson K, et al. Prospective study of paclitaxel-induced peripheral neuropathy with quantitative sensory testing. J Neuro-Onocol 1997;35:47-53. PMID: 9266440

Fowler FJ Jr, Barry MJ, Lu-Yao G, Roman A, Wasson J, Wennberg JE. Patient-reported complications and follow-up treatment after radical prostatectomy. The National Medicare Experience: 1988-1990 (updated June 1993). *Urology*. 1993;42(6):622-629.

Freeman RW, Bleecker ML, Comstock GW, Brookmeyer RS. Validation of self-administered questionnaire for study of peripheral neuropathy. Am J Epidemiol. 1985 Feb;121(2):291-300. PMID: 2990198

Fromme EK, Eilers KM, Mori M, Hsieh YC, Beer TM. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. *J Clin Oncol*. 2004;22(17):3485-3490.

Gabrilove JL, Perez EA, Tomita DK, Rossi G, Cleeland CS. Assessing symptom burden using the M. D. Anderson Symptom Inventory in patients with chemotherapy-induced anemia: results of a multicenter, open-label study (SURPASS) of patients treated with darbepoetin-alpha at a dose of 200 mug every 2 weeks. *Cancer*. 2007;110(7):1629-1640.

Galer BS, Gianas A, Jensen MP. Painful diabetic polyneuropathy: epidemiology, pain description, and quality of life. *Diabetes Res Clin Pract*. 2000;47(2):123-128.

Gammaitoni AR, Galer BS, Lacouture P, Domingos J, Schlagheck T. Effectiveness and safety of new oxycodone/acetaminophen formulations with reduced acetaminophen for the treatment of low back pain. *Pain Med*. 2003;4(1):21-30.

Garber K. Cancer research. Taking garbage in, tossing cancer out? Science 2002;295(5555):612-613. PMID: 11809952

Ger LP, Ho ST, Sun WZ, Wang MS, Cleeland CS. Validation of the Brief Pain Inventory in a Taiwanese population. *J Pain Symptom Manage*. 1999;18(5):316-322.

Given B, Given CW, McCorkle R, et al: Pain and fatigue management: Results of a nursing randomized clinical trial. Oncol Nurs Forum 29:949-956, 2002

Goldberg RM, Sargent DJ, Morton RF, et al. Early detection of toxicity and adjustment of ongoing clinical trials: the history and performance of the North Central Cancer Treatment Group's real-time toxicity monitoring program. *J Clin Oncol*. 2002;20(23):4591-4596.

Golomb BA, McGraw JJ, Evans MA, Dimsdale JE. Physician response to patient reports of adverse drug effects: implications for patient-targeted adverse effect surveillance. *Drug Safety*. 2007;30(8):669-675.

Grassi M, Nucera A, Zanolin E, Omenaas E, Anto JM, Leynaert B; European Community Respiratory Health Study Quality of Life Working Group

Green E, Sargent D, Goldberg RM, et al: Detailed analysis of oxaliplatin-associated neurotoxicity in Intergroup trial N9741. Ann Oncol 15 (Suppl 3):70, abstr. 254O, 2004

Grunberg SM, Groshen S, Steingass S, Zaretsky S, Meyerowitz B. Comparison of conditional quality of life terminology and visual analogue scale measurements. *Qual Life Res*. 1996;5:65-72.

Gudex C, Dolan P, Kind P, Williams A. Health state valuations from the general public using the Visual Analogue Scale. Qual Life Res. 1996;5:521-531.

Guggenmoos-Holzmann I. The meaning of kappa: probabilistic concepts of reliability and validity revisited. *J Clin Epidemiol*. 1996;49(7):775-782.

Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health*; doi:10.1111/j.1524-4733.2007.00231.x

Gwede CK, Johnson DJ, Daniels SS. Assessment of toxicity in cooperative oncology clinical trials: the long and short of it. *J Oncol Manage*. 2002;11(2):15-21.

Hadi S, Fan G, Hird AE, Kirou-Mauro A, Filipczak LA, Chow E. Symptom clusters in patients with cancer with metastatic bone pain. *J Palliat Med*. 2008;11(4):591-600.

Haddock MG, Sloan JA, Bollinger JW, et al. Patient assessment of bowel function during and after pelvic radiotherapy: results of a prospective phase III North Central Cancer Treatment Group clinical trial. *J Clin Oncol*. 2007;25(10):1255-1259.

Hann DM, Denniston MM, Baker F. Measurement of fatigue in cancer patients: further validation of the Fatigue Symptom Inventory. Qual Life Res. 2000;9(7):847-54. PMID: 11297027

Hann DM, Garovoy N, Finkelstein B, Jacobsen PB, Azzarello LM, Fields KK. Fatigue and quality of life in breast cancer patients undergoing autologous stem cell transplantation: a longitudinal comparative study. J Pain Symptom Manage. 1999 May;17(5):311-9. PMID: 10355210

Hann DM, Jacobsen PB, Azzarello LM, Martin SC, Curran SL, Fields KK, Greenberg H, Lyman G. Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. Qual Life Res. 1998 May;7(4):301-10. PMID: 9610214

Harris Interactive. Almost three-quarters of all U.S. adults—an estimated 163 million—go online. Available at: www.harrisinteractive.com/harris_poll/index.asp?PID=569. Last updated May 12, 2005. Accessed January 17, 2006.

Hickok JT, Morrow GR, Roscoe JA, Mustian K, Okunieff P. Occurrence, severity, and longitudinal course of twelve common symptoms in 1129 consecutive patients during radiotherapy for cancer. *J Pain Symptom Manage*. 2005;30(5):433-442.

Hjermstad MJ, Fayers PM, Bjordal K, Kaasa S. Health-related quality of life in the general Norwegian population assessed by the European Organization for Research and Treatment of Cancer Core Quality-of-Life Questionnaire: the QLQ=C30 (+ 3). J Clin Oncol. 1998 Mar;16(3):1188-96. PMID: 9508207

Hjollund NH, Andersen JH, Bech P. Assessment of fatigue in chronic disease: a bibliographic study of fatigue measurement scales. Health Qual Life Outcomes. 2007 Feb 27;5:12. PMID: 17326844

Hoekstra J, Bindels PJE, van Duijn NP, et al: The symptom monitor: A diary for monitoring physical symptoms for cancer patients in palliative care—Feasibility, reliability and compliance. J Pain Symptom Manage 27:24-35, 2004

Hollen PJ, Gralla RJ, Kris MG, McCoy S, Donaldson GW, Moinpour CM. A comparison of visual analogue and numerical rating scale formats for the Lung Cancer Symptom Scale (LCSS): does format affect patient ratings of symptoms and quality of life? *Qual Life Res*. 2005;14(3):837-847.

Holley SK. Evaluating patient distress from cancer-related fatigue: an instrument development study. Oncol Nurs Forum. 2000 Oct;27(9):1425-31. PMID: 11058974

Huschka MM, Mandrekar SJ, Schaefer PL, Jett JR, Sloan JA. A pooled analysis of quality of life measures and adverse events data in north central cancer treatment group lung cancer clinical trials. *Cancer*. 2007;109(4):787-795.

Hyland ME, Sodergren SC. Development of a new type of global quality of life scale and comparison and preference for 12 global scales. *Qual Life Res*. 1996;5(5):469-480.

Iñiguez C, Larrodé P, Mayordomo JI, et al. Reversible peripheral neuropathy induced by a single administration of high-dose paclitaxel. Neurology 1998;51:868-870. PMID: 9748043

Iop A, Manfredi AM, Bonura S. Fatigue in cancer patients receiving chemotherapy: an analysis of published studies. Ann Oncol. 2004 May;15(5):712-20. Review. PMID: 15111337

Jacobsen PB, Donovan KA, Vadaparampil ST, Small BJ.Systematic review and meta-analysis of psychological and activity-based interventions for cancer-related fatigue.Health Psychol. 2007 Nov;26(6):660-7. Review. Erratum in: Health Psychol. 2008 Jan;27(1):42. PMID: 18020836

Jacobsen PB, Hann DM, Azzarello LM, Horton J, Balducci L, Lyman GH. Fatigue in women receiving adjuvant chemotherapy for breast cancer: characteristics, course, and correlates. J Pain Symptom Manage. 1999 Oct;18(4):233-42. PMID: 10534963

Jensen MP, Smith DG, Ehde DM, Robinsin LR: Pain site and the effects of amputation pain: further clarification of the meaning of mild, moderate, and severe pain. *Pain*. 2001;91:317-322.

Jensen MP, Ehde DM, Hoffman AJ, Patterson DR, Czerniecki JM, Robinson LR. Cognitions, coping and social environment predict adjustment to phantom limb pain. *Pain*. 2002;95(1-2):133-142.

Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain*. 1986;27(1):117-126.

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(5):309-318.

King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. Qual Life Res. 1996 Dec;5(6):555-67. PMID: 8993101

Kirkova J, Davis MP, Walsh D, et al. Cancer symptom assessment instruments: a systematic review. *J Clin Oncol*. 2006;24(9):1459-1473.

Klepstad P, Loge JH, Borchgrevink PC, Mendoza TR, Cleeland CS, Kaasa S. The Norwegian Brief Pain Inventory questionnaire: translation and validation in cancer pain patients. *J Pain Symptom Manage*. 2002;24(5):517-525.

Krause SJ, Backonja M. Development of a Neuropathic Pain Questonnaire. Clin J Pain 2003;19:306-314.

Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol. 1989 Oct;46(10):1121-3. PMID: 2803071

Lai JS, Crane PK, Cella D. Factor analysis techniques for assessing sufficient unidimensionality of cancer related fatigue. Qual Life Res. 2006 Sep;15(7):1179-90. Epub 2006 Sep 25. Erratum in: Qual Life Res. 2007 Jun;16(5):911. PMID: 17001438

Lance et al. The sources of four commonly reported cutoff criteria: what did they really say? *Organiz Res Methods*. 2006;9:202-220.

Langer MM, Hill CD, Thissen D, Burwinkle TM, Varni JW, DeWalt DA. Item response theory detected differential item functioning between healthy and ill children in quality-of-life measures. *J Clin Epidemiol*. 2008;61(3):268-276.

Larue F, Colleau SM, Brasseur L, Cleeland CS. Multicentre study of cancer pain and its treatment in France*. BMJ*. 1995;310(6986):1034-1037.

Laudico AV, Mendoza TR, Siguan SS, Cleeland CS. Measuring cancer pain intensity and its effect on daily functioning: validation of the Cebuano version of the Brief Pain Inventory (BPI-Ce). *Philippine J Surg Specialt*. 2002;57(3):94-99.

Lipscomb J, Reeve BB, Clauser SB, et al. Patient-reported outcomes assessment in cancer trials: taking stock, moving forward. *J Clin Oncol*. 2007;25(32):5133-5140.

Litwin MS, Lubeck DP, Henning JM, Carroll PR. Differences in urologist and patient assessments of health related quality of life in men with prostate cancer: results of the CaPSURE database. *J Urol*. 1998;159(6):1988-1992.

Locke DE, Decker PA, Sloan JA, et al. Validation of Single-Item Linear Analog Scale Assessment of Quality of Life in Neuro-Oncology Patients. *J Pain Symptom Manage*. 2007; [Epub ahead of print]

MacQueen KM, McLellan E, Kay K, Milstein B. Codebook development for team-based qualitative analysis. *Cult Anthrop Methods.* 1998; 10(2):31-36.

Marshall, GN, Hays RD, Nicholas R. Evaluating agreement between clinical assessment methods. *Internat J Methods Psych Res*. 1994;4:249-257.

Masson EA, Boulton AJ. The Neurometer: validation and comparison with conventional tests for diabetic neuropathy. Diabet Med. 1991;8 Spec No:S63-6. PMID: 1825961

McConahay T, Bryson M, Bulloch B. Defining mild, moderate, and severe pain by using the color analogue scale with children presenting to a pediatric emergency department. *Acad Emerg Med*. 2006;13(3):341-344.

McCorkle R, Quint-Benoliel J. Symptom distress, current concerns and mood disturbance after diagnosis of life-threatening disease. *Social Sci Med*. 1983;17(7):431-438.

McCorkle R, Young, K. Development of a symptom distress scale. *Cancer Nurs*. 1978;1(5): 373-378.

Mendoza TR, Anderson KO, Wang XS, Easley M, Mobley G, Cleeland CS. Phone it in or fill it in? Comparing symptom assessment using the interactive voice response system with the traditional paper and pencil method. Presented at the Annual Meeting of the International Society for Quality of Life Research, Amsterdam, Netherlands, Nov. 2001.

Mendoza TR, Chen C, Brugger A, et al. The utility and validity of the modified Brief Pain Inventory in a multiple-dose postoperative analgesic trial. *Clin J Pain*. 2004;20(5):357-362.

Mendoza TR, Wang XS, Cleeland CS, et al. The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. *Cancer*. 1999;85(5):1186-1196.

Mitchell AC, Fallon MT. A single infusion of intravenous ketamine improves pain relief in patients with critical limb ischaemia: results of a double blind randomised controlled trial. *Pain*. 2002;97(3):275-281.

Mitchell SA, Berger AM. Cancer-related fatigue: the evidence base for assessment and management. Cancer J. 2006 Sep-Oct;12(5):374-87. Review. PMID: 17034674

Moghtaderi A, Bakhshipour A, Rashidi H. Validation of Michigan neuropathy screening instrument for diabetic peripheral neuropathy. Clin Neurol Neurosurg. 2006 Jul;108(5):477-81. Epub 2005 Sep 16. PMID: 16150538

Moinpour CM, Feigl P, Metch B, Hayden KA, Meyskens FL Jr, Crowley J. Quality of life end points in cancer clinical trials: review and recommendations. *J Natl Cancer Inst*. 1989;81(7):485-495.

Mularski RA, Rosenfeld K, Coons SJ, et al. Measuring outcomes in randomized prospective trials in palliative care. *J Pain Symptom Manage*. 2007;34(1 Suppl):S7-S19.

Munro AJ, Potter S: A quantitative approach to the distress caused by patients treated with radical radiotherapy. Br J Cancer 74:640-647, 1996 J Nurs Stud 26:69-79, 1989

Murtagh FE, Addington-Hall JM, Higginson IJ. The value of cognitive interviewing techniques in palliative care research. *Palliat Med*. 2007;21(2):87-93.

Mystakidou K, Mendoza T, Tsilika E, et al. , Bellos, G. et al. Greek Brief Pain Inventory: validation and utility in cancer pain. *Oncology*. 2001;60(1):35-42.

Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika*. 1991;78:691-692.

National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Available at https://webapps.ctep.nci.nih.gov/webobjs/ctc/webhelp/Welcome_to_CTCAE.htm#Welcome_to_CTCAE_1.htm. Accessed December 21, 2005.

NCI, CTEP. Common Terminology Criteria for Adverse Events Version 3 - Instructions and Guidelines. https://webapps.ctep.nci.nih.gov/webobjs/ctc/webhelp/Welcome_to_CTCAE.htm#Welcome_to_CTCAE_1.htm (Last Accessed June 17, 2008).

Nielsen//NetRatings. U.S. Internet usage shown mature growth. Available at www.netratings.com/pr/pr_050318.pdf. Last updated March 18, 2005. Accessed January 17, 2006.

Nielsen//NetRatings. Three out of four Americans have access to the Internet, online population surges past 200 million mark for the first time. Available at www.nielsennetratings.com/pr/pr_040318.pdf. Last updated March 18, 2004. Accessed January 17, 2006.

Nunnally JC, Bernstein IH. *Psychometric Theory*, 3rd ed. New York: McGraw-Hill, 1994.

Oken M, Creech R, Tormey D, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol.* 1982; 5:649-655.

Okuyama T, Wang XS, Akechi T, Mendoza TR, Hosaka T, Cleeland CS, Uchitomi Y. Validation study of the Japanese version of the brief fatigue inventory. J Pain Symptom Manage. 2003 Feb;25(2):106-17. PMID: 12590026

Palos GR, Mendoza TR, Mobley GM, Cantor SB, Cleeland CS. Asking the community about cutpoints used to describe mild, moderate, and severe pain. *J Pain*. 2006;7(1):49-56.

Patrick DL, Ferketich SL, Frame PS, et al. National Institutes of Health State-of-the-Science Conference Statement: Symptom Management in Cancer: Pain, Depression, and Fatigue, July 15-17, 2002. *J Natl Cancer Inst*. 2003;95(15):1110-1117.

Paul SM, Zelman DC, Smith M, Miaskowski C: Categorizing the severity of cancer pain: further exploration of the establishment of cutpoints. *Pain*. 2005;113:37-44.

Philip J, Smith WB, Craft P, et al: Concurrent validity of the modified Edmonton Symptom Assessment System with the Rotterdam Symptom Checklist and the Brief Pain Inventory. Support Care Cancer 6:539-541, 1998

Portenoy RK, Thaler HT, Kornblith AB, et al: The Memorial Symptom Assessment Scale: An instrument for the evaluation of symptom prevalence, characteristics and distress. Eur J Cancer 30:1326-1336, 1994

Portenoy RK, Thaler HT, Kornblith AB, et al. Symptom prevalence, characteristics and distress in a cancer population. *Qual Life Res*. 1994;3(3):183-189.

Postma TJ, Benard BA, Huijgens PC, Ossenkoppele GJ, Heimans JJ. Long term effects of vincristine on the peripheral nervous system. J Neuro-Oncol 1993;15:23-27. PMID: 8384253

Postma TJ, Vermorken JB, Liefting AJM, Pinedo HM, Heimans JJ. Paclitaxel-induced neuropathy. Ann Onocol 1995;6:489-494. PMID: 7669713

PROMIS Website, Fatigue Domain Team, http://www.nihpromis.org/Version%2010%20Item%20Banks/Forms/AllItems.aspx (Last Accesssed, June 18, 2008)

Quasthoff S, Hartung H-P. Chemotherapy-induced peripheral neuropathy. J Neurol 2002;249:9-17. PMID: 11954874

Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.

Radbruch L, Loick G, Kiencke P, et al. Validation of the German version of the Brief Pain Inventory. *J Pain Symptom Manage*. 1999;18(3):180-187.

Rammohan KW, Rosenberg JH, Lynn DJ, Blumenfeld AM, Pollak CP, Nagaraja HN. Efficacy and safety of modafinil (Provigil) for the treatment of fatigue in multiple sclerosis: a two centre phase 2 study. J Neurol Neurosurg Psychiatry. 2002 Feb;72(2):179-83. PMID: 11796766

Rao RD, Michalak JC, Sloan JA, Loprinzi CL, Soori GS, Nikcevich DA, Warner DO, Novotny P, Kutteh LA, Wong GY; North Central Cancer Treatment Group.Efficacy of gabapentin in the management of chemotherapy-induced peripheral neuropathy: a phase 3 randomized, double-blind, placebo-controlled, crossover trial (N00C3). Cancer. 2007 Nov 1;110(9):2110-8.

Robinson TN, Patrick K, Eng TR, Gustafson D. An evidence-based approach to interactive health communication: a challenge to medicine in the information age. Science Panel on Interactive Communication and Health. JAMA 1998;280(14):1264-9.

Rogers AS. Adverse drug events: identification and attribution. *Drug Intell Clin Pharm*. 1987;21(11):915-920.

Rosal MC, Carbone ET, Goins KV. Use of cognitive interviewing to adapt measurement instruments for low-literate Hispanics. *Diabetes Educ*. 2003;29(6):1006-1017.

Rostaing S, Lanteri-Minet M, Collin E, Grisart J, Boureau F.Development and validation of the Neuropathic Pain Symptom Inventory. Pain. 2004 Apr;108(3):248-57

Rosenthal D, Mendoza T. Measuring head and neck cancer symptom burden: development of the MDASI head and neck module. *Head Neck*. 2007;10:923-931.

Roth SH, Fleischmann RM, Burch FX, et al. Around-the-clock, controlled-release oxycodone therapy for osteoarthritis-related pain: placebo-controlled trial and long-term evaluation. *Arch Intern Med*. 2000;160(6):853-860.

Rothman ML, Beltran P, Cappelleri JC, et al. Patient-reported outcomes: conceptual issues. *Value Health*. 2007;10(Suppl 2):S66-S75.

Sandler SG, Tobin W, Henderson ES. Vincristine-induced neuropathy. A clinical study of fifty leukemic patients. Neurology 1969;19:367-374. PMID: 5813374

Sargent DJ, Goldberg RM, Mahoney MR, et al. Rapid reporting and review of an increased incidence of a known adverse event. *J Natl Cancer Inst*. 2000;92(12):1011-1013.

Saxena A, Mendoza T, Cleeland CS. The assessment of cancer pain in north India: the validation of the Hindi Brief Pain Inventory--BPI-H. *J Pain Symptom Manage*. 1999;17(1):27-41.

Schiffmann R, Kopp JB, Austin HA, et al. Enzyme replacement therapy in Fabry disease: a randomized controlled trial. *JAMA*. 2001;285(21):2743-2749.

Schwartz, A. (2002). Validity of cancer-related fatigue instruments. Pharmacotherapy, 22 (11), 1433-1441.

Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res*. 2002;11:193-205.

Science Panel on Interactive Communication and Health. Wired for Health and Well-Being: the Emergence of Interactive Health Communication. Washington, DC: US Department of Health and Human Services, US Government Printing Office; April 1999.

Semenchuk MR, Sherman S, Davis B. Double-blind, randomized trial of bupropion SR for the treatment of neuropathic pain. *Neurology*. 2001;57(9):1583-1588.

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.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*.1979;86:420-428.

Shvartzman P, Friger M, Shani A, Barak F, Yoram C, Singer Y. Pain control in ambulatory cancer patients--can we do better? *J Pain Symptom Manage*. 2003;26(2):716-722.

Sloan JA. Statistical issues in the application of cancer outcome measures. In Lipscomb J, Gotay CC, Snyder C, eds. *Outcomes Assessment in Cancer: Measures, Methods, and Applications*. New York: Cambridge University Press, 2005; pp. 362-385.

Sloan J, Zhao C. Quality of life and genetics. *Curr Prob Cancer*. 2006;30:265-271.

Sloan JA, Aaronson N, Cappelleri JC, Fairclough DL, Varricchio C. Assessing the clinical significance of single items relative to summated scores. *Mayo Clin Proc*. 2002;77(5):479-487.

Sloan JA, Berk L, Roscoe J, et al. Integrating patient-reported outcomes into cancer symptom management clinical trials supported by the National Cancer Institute-sponsored clinical trials networks. *J Clin Oncol*. 2007;25(32):5070-5077.

Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. *J Clin Epidemiol*. 2005;58(12):1217-1219.

Sloan JA, Halyard MY, Frost MH, et al. The Mayo Clinic manuscript series relative to the discussion, dissemination, and operationalization of the Food and Drug Administration guidance on patient-reported outcomes. *Value Health*. 2007;10(Suppl 2):S59-S63.

Smith DJ, Huntington J, Sloan JA. Choosing the “correct” assessment tool. *Curr Prob Cancer*. 2005;29:272-282.

Smith MY, Egert J, Winkel G, Jacobson J. The impact of PTSD on pain experience in persons with HIV/AIDS. *Pain*. 2002;98(1-2):9-17.

Sneeuw KC, Sprangers MA & Aaronson NK 2002 The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. *Journal of Clinical Epidemiology* 55 1130–1143.

Sriwatanakul K, Kelvie W, Lasagna L, Calimlim JF, Weis OF, Mehta G. Studies with different types of visual analog scales for measurement of pain. *Clin Pharmacol and Ther*. 1983;34(2):234-239.

Stone P. The measurement, causes and effective management of cancer-related fatigue. Int J Palliat Nurs. 2002 Mar;8(3):120-8. Review. PMID: 11923740

Su L, Tucker R, Frey SE, et al. Measuring injection-site pain associated with vaccine administration in adults: a randomised, double-blind, placebo-controlled clinical trial. *J Epidemiol Biostat*. 2000;5(6):359-365.

Talcott JA, Rieker P, Propert KJ, et al. Patient-reported impotence and incontinence after nerve-sparing radical prostatectomy. *J Natl Cancer Inst*. 1997;89(15):1117-1123.

Teresi JA, Fleishman JA. Differential item functioning and health assessment. *Qual Life Res*. 2007;16:33-42.

Teresi JA. Overview of quantitative measurement methods: equivalence, invariance, and differential item functioning in health applications. *Med Care*. 2006;44(Suppl 3):S39-S49.

Thie NM, Prasad NG, Major PW. Evaluation of glucosamine sulfate compared to ibuprofen for the treatment of temporomandibular joint osteoarthritis: a randomized double blind controlled 3 month clinical trial. *J Rheumatol*. 2001;28(6):1347-1355.

Trotti A. The evolution and application of toxicity criteria. *Semin Radiat Oncol*. 2002;12(1 Suppl 1):1-3.

Trotti A, Bentzen SM. The need for adverse effects reporting standards in oncology clinical trials. *J Clin Oncol*. 2004;22(1):69-76.

Trotti A, Colevas AD, Setser A, Basch E. Patient-reported outcomes and the evolution of adverse event reporting in oncology. *J Clin Oncol*. 2007;25(32):5121-5127.

Trotti A, Colevas AD, Setser A, et al. CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. *Semin Radiat Oncol*. 2003;13(3):176-181.

Turk DC, Dworkin RH, Allen RR, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2003;106:337-345.

Uki J, Mendoza T, Cleeland CS, Nakamura Y, Takeda F. A brief cancer pain assessment tool in Japanese: the utility of the Japanese Brief Pain Inventory—BPI-J. *J Pain Symptom Manage*. 1998;16(6):364-373.

U.S. Department of Health and Human Services. What is usability? Available at http://www.usability.gov/basics/whatusa.html. Accessed June 1, 2008.

Validation study of the Chinese version of the Brief Fatigue Inventory (BFI-C). J Pain Symptom Manage. 2004 Apr;27(4):322-32. PMID: 15050660

Varni JW, Limbers C, Burwinkle TM. Literature review: health-related quality of life measurement in pediatric oncology: hearing the voices of the children. *J Pediatr Psychol*. 2007;32(9):1151-1163.

Varni JW, Limbers CA, Burwinkle TM. How young can children reliably and validly self-report their health-related quality of life?: an analysis of 8,591 children across age subgroups with the PedsQL 4.0 Generic Core Scales. *Health Qual Life Outcomes*. 2007;5:1.

Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol*. 2004;22(4):714-724.

Vileikyte L, Peyrot M, Bundy C, Rubin RR, Leventhal H, Mora P, Shaw JE, Baker P, Boulton AJ. The development and validation of a neuropathy- and foot ulcer-specific quality of life instrument. Diabetes Care. 2003 Sep;26(9):2549-55. PMID: 12941717

Vinik EJ, Hayes RP, Oglesby A, Bastyr E, Barlow P, Ford-Molvik SL, Vinik AI. The development and validation of the Norfolk QOL-DN, a new measure of patients' perception of the effects of diabetes and diabetic neuropathy.

Vogelzang NJ, Breitbart W, Cella D, et al. Patient, caregiver, and oncologist perceptions of cancer-related fatigue: results of a tripart assessment survey. The Fatigue Coalition. *Semin Hematol*. 1997;34(3 Suppl 2):4-12.

Wampler MA, Miaskowski C, Hamel K, Byl nancy, Rugo H, Topp K. The Modified Total Neuropathy Score. J Support Oncol 2006;4:W9-W16.

Wang XS, Janjan NA, Guo H, Johnson BA, Engstrom MC, Crane CH, Mendoza TR, Cleeland CS. Fatigue during preoperative chemoradiation for resectable rectal cancer. Cancer. 2001 Sep 15;92(6 Suppl):1725-32. PMID: 11598894

Wang XS, Giralt SA, Mendoza TR, et al. Clinical factors associated with cancer-related fatigue in patients being treated for leukemia and non-Hodgkin's lymphoma. *J Clin Oncol*. 2002;20(5):1319-1328.

Wang XS, Mendoza TR, Gao SZ, Cleeland CS. The Chinese version of the Brief Pain Inventory (BPI-C): its development and use in a study of cancer pain. *Pain*. 1996;67(2-3):407-416.

Ward S, Donovan HS, Owen B, Grosen E, Serlin R. An individualized intervention to overcome patient-related barriers to pain management in women with gynecologic cancers. *Res Nurs Health*. 2000; 23(5):393-405.

Watson K, Baranowski T, Thompson D. Item response modeling: an evaluation of the children's fruit and vegetable self-efficacy questionnaire. *Health Educat Res*. 2006;21(Supple 1):i47-i57.

Weiden PL, Wright SE. Vincristine Neurotoxicity. New Eng J Med 1972;286:1369-1370. PMID: 5027400

West CP, Huntington JL, Huschka MM, et al. A prospective study of the relationship between medical knowledge and professionalism among internal medicine residents. *Acad Med*. 2007;82(6):587-592.

Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. *Res Nurs Health*. 1990;13:227-236.

Wiklund I. Assessment of patient-reported outcomes in clinical trials: the example of health-related quality of life. *Fundam Clin Pharmacol*. 2004;18(3):351-63.

Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*. 2005;8(2):94-104.

Willis GB. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Sage Publications, 2004.

Willke RJ, Burke LB, Erickson P. Measuring treatment impact: a review of patient-reported outcomes and other efficacy endpoints in approved product labels. *Control Clin Trials*. 2004;25(6):535-52.

Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. J Pain Symptom Manage. 1997 Feb;13(2):63-74.

Yu DT, Chae J, Walker ME, Fang ZP. Percutaneous intramuscular neuromuscular electric stimulation for the treatment of shoulder subluxation and pain in patients with chronic hemiplegia: a pilot study. *Arch Phys Med Rehab*. 2001;82(1):20-25.

Yun YH, Mendoza TR, Heo DS, et al. Development of a cancer pain assessment tool in Korea: a validation study of a Korean version of the Brief Pain Inventory. *Oncology*. 2004;66(6):439-444.

Zelman DC, Hoffman DL, Seifeldin R, Dukes EM. Development of a metric for a day of manageable pain control: derivation of pain severity cut-points for low back pain and osteoarthritis. *Pain*. 2003;106:35-42.

Zelman DC, Gore M, Dukes E, Tai KS, Brandenburg N. Validation of a modified version of the brief pain inventory for painful diabetic peripheral neuropathy. J Pain Symptom Manage. 2005 Apr;29(4):401-10.

Zelman DC, Gore M, Dukes E, Tai KS, Brandenburg N. Validation of a modified version of the Brief Pain Inventory for painful diabetic peripheral neuropathy. J Vasc Nurs. 2005 Sep;23(3):97-104. PMID: 16125633

# APPENDIX 2:
## OPTIONAL STUDIES NOT INCLUDED IN CURRENT TECHNICAL PROPOSAL OR BUDGET

### 1. Prospective Questionnaire Study to Evaluate Extent of Added Information From "Interference" to "Intensity"

- *Background*. If the current double-barreled structure of CTCAE symptom items is abandoned (based on consensus and/or results of the above cognitive interviewing study) and the PRO-CTCAE is designed to include only a single concept per item, the question still remains as to whether it is necessary to separately measure more than one concept for each symptom to adequately characterize those symptoms. In other words, is it sufficient only to measure symptom intensity (in the setting of screening for adverse events, or do we believe that there is enough additional information provided by *also* measuring interference or frequency that it is worth the additional questionnaire burden to ask more than one question per symptom. Thus, the choices we are faced with for concepts in PRO-CTCAE items are as follows: (1) single item per symptom asking about intensity (or severity) only or (2) multiple items per symptom, with the first question asking about intensity (or severity), and if that screens positive, then a follow-up question is triggered about interference (for category I items) or frequency (for category II items).

- *Study Design.* We will administer an anonymous waiting room paper questionnaire in 120 patients with cancer diagnoses specified in Table A.5-6 who are likely to be symptomatic at three of the participating sites (Mayo, MSKCC, MDACC). We will evaluate approximately 45 different symptoms (ie, category I and II symptoms) with included items for both intensity and interference, but each questionnaire will include items for only 10 discrete symptoms, with selection of items such that we are likely to elicit responses from patients who are experiencing the symptoms of interest. The response option scale for both items will be 0-10 to maximize our analytic approaches (described below). Survey formats will involve randomization of item orders among questionnaires to ensure that order effect does not confound the results. Surveys will be administered to patients and collected by clinical research associates/data managers at each site and will be submitted to the Coordinating Center (MSKCC) for tabulation. Notably, these interviews will be conducted off-protocol via an IRB waiver and therefore will not require development of a full IRB protocol or site IRB review. The investigators have previously negotiated this designation for such interviews with site IRBs, which will save us time and effort in moving this work forward. The rationale accepted by site IRBs for receiving waivers is that these interviews represent minimal risk to respondents and no PHI (Protected Health Information) is being collected (only age, educational level, ethnicity, sex, and cancer type will be self-reported by respondents). The investigators have extensive past experience administering this type of waiting room questionnaire.

Notably, this questionnaire will also include items for the below-described analysis of alternative response option scales. Therefore, the total number of items in each questionnaire will be 45, anticipated to take no more than 15 minutes to complete.

- *Primary Analytic Approach (Correlation Coefficients).* The intraclass correlation coefficient (ICC) and kappa coefficients will be used to measure agreement and test measurement equivalence. The ICC provides a means to assess the reliability of scores from an instrument administered across multiple raters (Shrout and Fleiss, Psychol Bul 1979). The ICC takes into account both relative position in the group of scores and the amount of deviation above or below the group mean (Anastasi and Urbina, Psychol Test 1996). The kappa coefficient is a commonly used measure of agreement that calculates the proportion of responses in agreement in relation to the proportion of responses that would be expected by chance alone (Cohen, 1960). Fleiss (1981) suggests that kappa coefficients of less than 0.40 are poor, 0.40 to 0.59 are fair, 0.60 to 0.74 are good, and greater than 0.74 are excellent. For ICC results, we will use the standards for acceptable levels of reliability, specifically 0.70 for group comparisons and 0.85 to 0.95 for individual comparisons (Nunnally and Bernstein, Psychom Theory 1994).

- *Secondary Analytic Approach (Bland and Altman).* The use of correlation coefficients or measures of agreement alone is necessary but not sufficient because they might not be sensitive to systematic mean differences between groups and as a result tend to overestimate agreement. Hence, we will further examine equivalence of the alternative methods of measurement using methods developed by Bland and Altman (Intl Jrl Epidem 1995), which combine simple graphical techniques with hypothesis testing for measurement equivalence. Several examples of applications of these measurement equivalence procedures have been published, including work by the analytical team at Mayo under the leadership of Dr. Sloan (Marshall et al., Intl Jrl Psych Res 1994; Sloan et al., J Clin Epidem 2005; Smith et al., Cur Prob CA 2005). The basic approach here is to compute the average intrapatient difference with the alternative assessment modes (via a simple paired *t* test) and correlate the individual intrapatient differences vs. the individual's average on the two scores, which allows for the determination of whether differences in the mode of administration are subject to floor or ceiling effects or vary depending on the level of symptom severity reported (via graphical inspection of a bivariate scatterplot and a formal hypothesis test using a Fisher's transformation z-statistic for a correlation of zero between the intrapatient difference and average).

- *Exploratory Analytic Approach (Regression Analysis).* As a further comparison of whether intensity or interference of a symptom provides a greater impact on patient well-being, we will use a stepwise polytomous logistic regression modeling process with a single-item self-reported global measure of health state as the dependent variable to see what if any difference in the amount of variance is accounted for by symptom interference and intensity. The polytomous logistic regression model is

used here to account for the ordinal nature and small number of response categories. We may have to collapse response categories, depending on the performance status distribution. We will directly test, using standard nonlinear modeling processes (ie, chi-square testing), whether having both intensity and interference in the model adds to the model's ability to predict a patient's global health state. We will compare generalization of the coefficient of determination $R^2$ developed for discrete models among a null model, a model with intensity item only, and a model with both intensity and interference items (Nagelkerke, Biometrika 1991). The investigators have previously used this approach in related work.

- *Power Calculation*. A sample size that provides 38 observations for each PRO-CTCAE symptom of interest would provide 90% power to evaluate whether the correlation coefficient between the two concept items is at least 0.4 (at least fair agreement) if the true correlation is 0.75 using a two-sided test.

- *Interpretation of Results.* If the observed Pearson correlations of interference and intensity are generally above 0.74 (demonstrating that the two concepts share at least half of their variance), and the secondary and exploratory analyses support this finding, we will conclude that the two domains are sufficiently redundant, such that for the purpose of adverse event screening, the added questionnaire burden associated with evaluating both concepts is not justified (Nunnally and Bernstein, Psychom Theory 1994; Lance et al, Org Res Meth 2006). If this is found to be the case, the investigators will be inclined to include only a single intensity item for each PRO-CTCAE symptom. However, if the items are found to have less agreement than hypothesized, the investigators will plan to discuss the results with NCI staff to determine the degree of interest in evaluating *both* intensity (or severity) and interference for each symptom in the PRO-CTCAE. (Notably, even if only severity is included in the initial PRO-CTCAE, an advantage of creating an electronic infrastructure for administering items is that we will still have an opportunity in the future to modify the system to automatically trigger follow-up items that evaluate other concepts on a trial-by-trial basis).

- *Timeline and Deliverables to NCI.* The questionnaire for this study will be designed by Drs. Basch, Sloan, Cleeland, and Mendoza with input from the investigative team and NCI staff during month 1 of the contract period. Because we are able to do this work under IRB waivers, we will not need to develop a full protocol or wait for local IRB approval. Thus, questionnaire administration can begin immediately on completion of the questionnaire design. Questionnaires will be administered by clinical research associates/data managers at each site and will be tightly coordinated by the Project Manager, Laura Sit, at the Coordinating Center (MSKCC). Questionnaires will be administered between month 2 and month 4, which based on past experience of the investigators is sufficient time. Compilation and analysis of the results and presentation of this work to NCI staff for review will occur in month 5.

This information will be used for creation of the draft PRO-CTCAE items later that same month.

*Proposed Retrospective Data Mining Study to Correlate Interference With Intensity*

- *Study Design.* The investigators have access to two large existing questionnaire data sets in which patients with various cancer diagnoses were asked to respond to items regarding both intensity and interference. This retrospective analysis will measure the Pearson correlation coefficients, with a similar design to the above prospective study design.

- *Data Sets.* Drs. Cleeland and Mendoza have a data set at MDACC from prior questionnaire development work in 640 patients for the MDASI (Cleeland et al, Cancer 2000). Specifically, 52 distinct symptom items have been included with assessments of intensity and interference, which overlap with all of the proposed PRO-CTCAE core items and 70% of the disease-specific items. Data are available from administration in 640 patients with diverse cancer diagnoses. Drs. Bill and Barbara Given at Michigan State University similarly have an NCI-funded data set in which 353 patients were asked to report both intensity and interference. Each patient was asked to respond to 17 symptom intensity items and for each symptom item was then asked how much that symptom interferes with daily activities, enjoyment of life, social relationships, and emotions. Dr. Mendoza has previously collaborated with Dr. Given in analyzing these data, and Dr. Given has graciously agreed to provide his data set for this analysis.

- *Analytic Approach.* The same analytic approaches will be used as described above for this data set, except that for the fourth approach, instead of using the global single item as an anchor, we will use the general health state item from the SF-36 (which was administered to participants), with analysis via a simple regression model.

**2. Study of what format is most appropriate for PRO-CTCAE response options (eg, 0-10 NRS, 0-4 NRS, or verbal descriptor scale)**

- *Questionnaire.* A paper survey can be designed for administration in 120 patients which includes symptom items measured via three alternative scale options: 0-4 NRS, 0-10 NRS, and a verbal descriptor scale with five rankings (none, mild, moderate, severe, very severe). We will evaluate approximately 45 different symptoms in total across the population (ie, category I and II symptoms), but any given questionnaire will only need to include 15 discrete symptoms towards our planned analyses. The order of symptoms and order of formats within each symptom will be randomly assigned among questionnaires. On completion of these items, patients will be presented with all four formats on a single page and asked a single question to ascertain their preference among the four (first choice and rank order). A final question will ask them for their estimate of how much they believe the format matters to the way in which they reported their score (not at all, a little, a

moderate amount, a lot). The investigators have extensive past experience with the design and administration of similar waiting room questionnaires.

- *Patients and Questionnaire Administration.* The questionnaire will be administered in 120 patients with cancer diagnoses, at three of the participating sites. Surveys will be administered to patients and collected by clinical research associates/data managers at each site and will be submitted to the Coordinating Center (MSKCC) for tabulation. These questionnaires will be conducted off-protocol via IRB waivers and therefore will not require development of a full IRB protocol or site IRB review.

- *Analysis*. The purpose of this investigation is to verify our hypothesis that the response option format does not incur substantial measurement error. If substantial error is observed, we will choose the optimal format based on the results of these representative symptoms. It is unrealistic to have a different response format for different symptoms. The Bland and Altman approach (previously described) will be used to compare pairs of response formats per symptom. Additionally, an analysis of variance model will be constructed to assess the effect of order of items on the results. Analysis of the "patient preference" questions will be conducted to ascertain the extent to which patients feel that bias is introduced by response format (ie, do patients who feel there is no difference demonstrate no difference in the scores they provide?; do patients who say there is a difference actually demonstrate a difference?).

- *Power Calculation*. For the power test per the Bland and Altman approach, a sample of 38 observations will have >90% power to detect a moderate correlation (0.5) between the difference in the scales and the average of the scales vs. a weak correlation (0.1) using a two-sided test. The Investigators have extensive past experience with this analytic approach.

- *Interpretation.* We will discuss with NCI staff and advisers the choice of response option format, also taking into account the advantages and disadvantages of each. The selected format will be used for creation of the draft PRO-CTCAE items.

**3. Optional study design to evaluate patient understanding of double-barreled concepts in response options and preferences for various response formats**

*Rationale for Study.* This study is designed to evaluate the extent to which patients might "silently misinterpret" double-barreled response options if they were included in the PRO-CTCAE (ie, responses that require respondents to evaluate both intensity and interference to answer a single question). As described in the "Rationale for Approach" and "Task 2" sections of the Technical Proposal, the investigators are predisposed against the use of double-barreling because we suspect that patients will misinterpret the response options, that double-barreled items are not easily translated or adapted for use in children or patients with cognitive disabilities, and because such items take longer

to administer than items with NRS or VAS response options (and therefore cause excess questionnaire burden for respondents). Yet, the current symptom items in CTCAE v3 use double-barreled response options, and therefore the investigators believe that the burden is on us to demonstrate scientifically why we propose to abandon this traditional CTCAE format.

*Study Design and Patient Population.* We will conduct cognitive interviews in 30 patients with various cancer types receiving chemotherapy and/or radiotherapy at two of our five study sites (DFCI and M.D. Anderson), including 15 patients at main hospitals and 15 in community satellite clinics. The recruitment strategy will ensure that we include 15 patients with some education beyond high school and 15 with a high school education or less. Because this is a fundamental structural issue related to item design, we do not believe that our findings will vary by patient characteristics, but nonetheless have designed the sample to provide some variation by sex, race, and ethnicity.

Notably, these interviews will be conducted off-protocol via an IRB waiver and therefore will not require development of a full IRB protocol or site IRB review. The investigators have previously negotiated this designation for cognitive interviews with site IRBs at DFCI and M.D. Anderson, which will save us time and effort in moving this work forward. The rationale accepted by site IRBs for receiving waivers is that these interviews represent minimal risk to respondents and no PHI is being collected (only age, educational level, ethnicity, sex, and cancer type will be self-reported by respondents).

*Enrollment.* Patients will be initially approached by clinical staff at each location. Patients will be screened for educational level for stratification purposes at this time, and private interview space at each clinic area has been previously arranged (a financial incentive is budgeted for sites for each patient identified and provision of interview space). A financial incentive for patients ($25 gift certificate for each patient) will be budgeted.

*Questionnaire*. The questionnaire will be developed by Dr. Hay and her staff, in consultation with the investigative team and NCI staff. Interviews will be conducted by a trained interviewer on Larry Shiman's staff. The overall approach will involve each patient first responding to a paper questionnaire that includes double-barreled items (ie, current MSKCC "STAR" patient adaptations of CTCAE v3 items). This will be followed by retrospective semistructured "probing" by the interviewer to determine reasons a patient selected particular item responses. As a secondary analysis, we will include items with 0-10 NRS, 0-4 NRS, and verbal descriptor scales in the questionnaire and will include probes to ascertain if patients prefer a specific response format. All interviews will be audio recorded.

*Analysis.* For the evaluation of double-barreling, we will tabulate the proportion of patients who answered each item predominantly based on concept 1 (eg, symptom intensity) vs. concept 2 (eg, interference or frequency, depending on the item) vs. a combination of both concepts. This will provide descriptive evidence regarding the

proportion of patients who consider responses in each way.  For the evaluation of response format preference, we will tabulate patient preferences and reasons for these preferences.