Clinical Outcome Assessments (COA) Qualification Program
DDT COA #000079: PROMIS® Physical Function in Oncology
Qualification Plan

Section 1: Proposed Plan for COA Qualification
1.1 Introduction and Overview

This Qualification Plan is submitted to support the Food and Drug Administration's consideration of one of two custom PROMIS® Physical Function short forms – a version with the PROMIS standard "no recall" period (PROMIS Short Form v2.0 – Physical Function 8c**)** or a custom 7-day recall period version (PROMIS Short Form v2.0 – Physical Function 8c 7-Day) as a qualified Clinical Outcome Assessment Drug Development Tool (COA DDT) in the context of advanced oncology. Henceforth, unless otherwise specified, all references to **PROMIS Short Form v2.0 – Physical Function 8c** refer to both the no-recall and 7-day versions. Specifically, we are submitting the **PROMIS Short Form v2.0 – Physical Function 8c** as a qualified COA DDT in adult patients with advanced solid tumors or hematologic malignancies. The proposed context of use for this COA is as follows: (1) adult patients with advanced solid tumors or hematologic malignancies of any primary site or origin (i.e., Stage III-IV), including unknown primary site; (2) performance status range 0-3; (3) receiving active anti-cancer therapy; (4) symptomatic or asymptomatic; and (5) without limitation regarding language or background/culture of the patient.

The items under consideration for inclusion in the custom short form were identified in 2016-2018 from a larger set of items in the Patient Reported Outcome Measurement Information System® (PROMIS) Physical Function item bank, an assessment tool that was designed using well-specified instrument development protocols. This item bank allows for assessment across many disease states and a broad range of physical function ability. The items included in the **PROMIS Short Form v2.0 – Physical Function 8c** were selected based on applicability across cultures and backgrounds as well as their performance statistics in a sample of advanced cancer patients. The scores produced from this cancer specific custom short form can be mapped directly onto the general (non-cancer specific) PROMIS Physical Function metric.

The rationale for assessment based on the PROMIS Physical Function metric stems from the physical function limitations experienced by many cancer patients at various points in the progression of disease, during treatment, and in remission. Given that decrements in physical function represent both increased mortality risk and quality of life impairment, a precise and valid measurement tool is needed in order to specify the effects of cancer treatments on physical function. Although clinician-rated performance status measures are used widely to characterize physical function in individuals with advanced cancer, they are limited by low inter-rater reliability and lack of precision required to detect meaningful improvement or worsening. Consequently, measures of physical function that incorporate the patient's perspective are integral to accurately assessing the impact of interventions, highlighting the need for reliable and valid PRO measures of physical function that could be incorporated into clinical trials. While several widely used PRO measures assess some aspects of physical function in addition to other symptoms (e.g., SF-36), the PROMIS Physical Function item bank was developed to allow for the assessment of functioning across a wide range of activities and levels of ability. This represents an important advance in measuring physical function from the patient perspective. The PROMIS Physical Function item bank has been used in a number of cancer populations, including diverse solid tumors with multiple primary sites, non-Hodgkin's lymphoma, gynecologic tumors, and head and neck tumors. The Oncology and Hematology Division of the U.S. Food and Drug Administration (FDA) recently suggested that physical function be measured as an endpoint in clinical trials.[1]

The **PROMIS Short Form v2.0 – Physical Function 8c** contains a selection of items that cover a range of physical ability and include example activities that are applicable to advanced cancer patients across cultures and

backgrounds. Items considered for inclusion in the **PROMIS Short Form v2.0 – Physical Function 8c** were pulled from a larger set of items in the PROMIS-Cancer Physical Function (PROMIS-Ca Physical Function) item bank. As part of the DDT Qualification submission process, 31 items were selected from the PROMIS-Ca Physical Function item bank for further evaluation. Selection of the 31 items was achieved through a collaboration process and meeting including Northwestern University researchers, pharmaceutical industry representatives, and the FDA, including the COA Qualification Review Team (QRT) and the Office of Oncology and Hematology Products (OHOP). What follows below includes background on the development of the general PROMIS Physical Function item bank and the PROMIS-Ca Physical Function item bank (see Sections 1.1.1 – 1.1.2), and the item selection process for the **PROMIS Short Form v2.0 – Physical Function 8c** to be submitted as a qualified COA DDT.

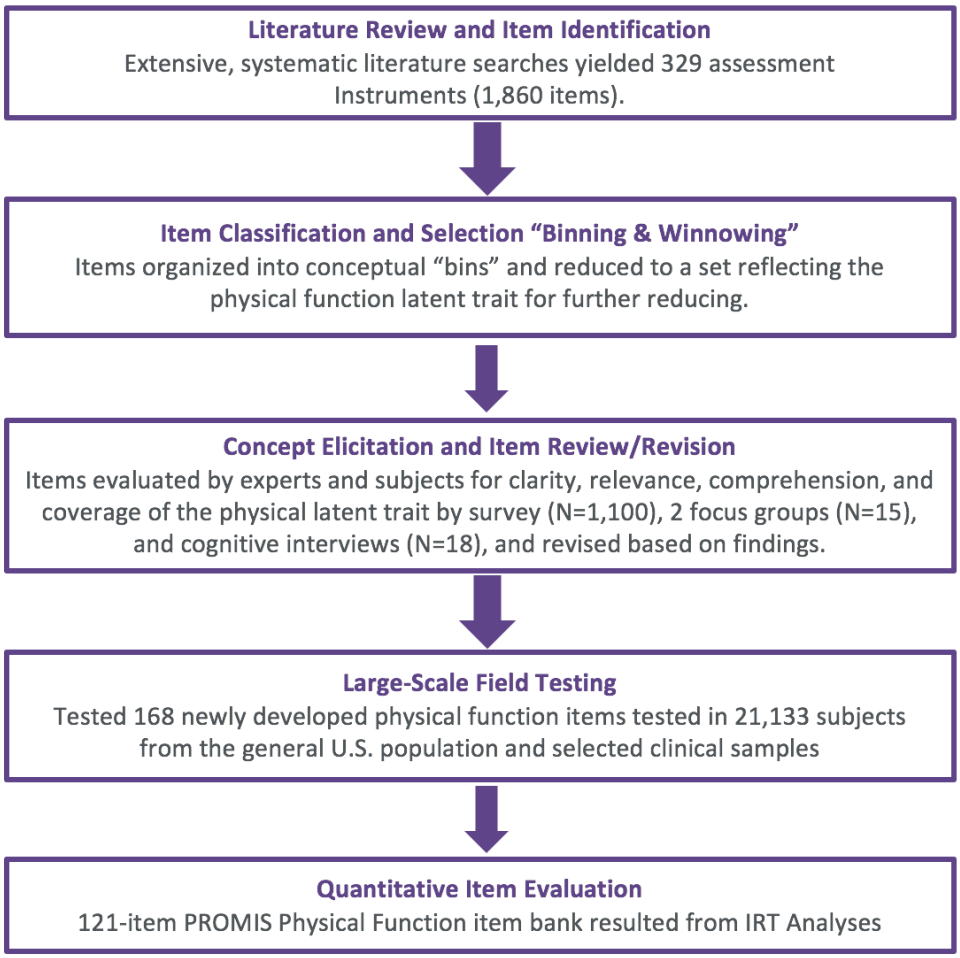
1.1.1 Background on the Development of the PROMIS® Physical Function Item Bank

The PROMIS Physical Function item bank comprises 165 items that assess a large range of physical ability and target the subdomains of mobility, upper extremity, and central body function. **Given that the items that constitute the PROMIS Short Form v2.0 – Physical Function 8c are a subset of the PROMIS Physical Function and PROMIS-Ca Physical Function item banks that produce scores on the same metric, description of the development of the two item banks is provided here as background for description of the PROMIS Short Form v2.0 – Physical Function 8c developed as part of the submission process of DDT# 079.**

The PROMIS Physical Function item bank is an assessment tool developed using well-specified instrument development protocols that allow for assessment across many disease states and a broad range of physical function ability. Development of the item bank was based on several aims, including the development of measures that: (1) assess a wide range of physical function ability across many disease states,[2] (2) capitalize on the benefits of inclusion in the PROMIS measurement framework (particularly standardization and uniformity across instruments), and (3) reflect cutting edge methods for instrument development and administration. To this end, in 2005-2008 a team of experts in the assessment of physical function at the Stanford PROMIS Primary Research Site (James Fries, PI, Bonnie Bruce, Deborah Ambrosini, and Barathi Lingala) collaborated to achieve these aims by following the protocols for PROMIS measurement development, a multistep, mixed-methods approach to develop the physical function item bank.[3,4]

As shown in Figure 1, development of the item bank followed a mixed-methods approach involving a comprehensive literature review, identification and development of a large pool of candidate items, item reduction through binning and winnowing, qualitative research to elicit aspects of the physical function concept, review and revision of the reduced item set,[4] large-scale quantitative data collection based on administration to a representative sample, and intensive psychometric analyses.[5] Through this process, PROMIS investigators achieved a 165-item physical function item bank appropriate for use across many disease states and a broad range of physical function ability. Methods for the development of the PROMIS® Physical Function item bank are detailed below.

Figure 1. Development of the PROMIS Physical Function Item Bank



1.1.1.2 Literature review

The team of experts responsible for developing the PROMIS Physical Function item bank conducted a comprehensive literature review to explore the physical function domain, identify its subordinate concepts, and identify all existing measures of physical function. Given the extensive prior research on physical function (across a variety of disease states), PROMIS experts concluded that the concept of physical function had been sufficiently researched and articulated to enable an efficient and incisive distillation of the full range of content covered by this domain. Following completion of the literature review, PROMIS experts defined physical function as "the ability to carry out various activities that require physical capability, ranging from self-care (basic activities of daily living) to more vigorous activities that require increasing degrees of mobility, strength, and/or endurance".[4] Further, PROMIS investigators identified four distinct but conceptually-related subdomains, namely mobility or lower extremity function, back and neck (central) function, upper extremity function, and complex activities that involve more than one subdomain (instrumental activities of daily living).[5]

1.1.1.3 Item identification

The literature review described in Section 1.1.1.2 resulted in the identification of 329 assessment instruments of physical function used in a range of fields including oncology, cardiology, gastroenterology, rheumatology, neurology, ophthalmology, pediatrics, and physical medicine and rehabilitation.[4] Of the 329 instruments identified, half (n = 165) were considered relevant as they contained items that assessed one or more of the four primary subdomains of physical function.[5] Instruments were excluded due to irrelevant content, such as a focus on pain, fatigue, or satisfaction, rather than physical function. From the 165 instruments deemed relevant to physical function, PROMIS investigators identified a pool of 1,860 candidate items including those from legacy measures such as the HAQ-DI and the SF-36. Ninety-three percent (n = 1,728) of the 1,860 items represented one or more

of the physical function subdomains. Most of the items (n = 1,578) assessed *capability* (e.g., "Are you able..."). The bulk of the remaining items assessed *performance* ("Did you..."). All 1,860 items were retained in a descriptive file denoting each item's intellectual property status, origin, and other notes regarding the item utility and/or limitations for consideration during subsequent phases of item reduction.

1.1.1.4 Binning and winnowing

Next, the 1,860 candidate items were sorted into bins based on item content; 71 different bins were created in total.[5] The largest proportion of items assessed *walking* (17%; n = 309) and *dressing* and *grooming* (7%; n = 133). Other large bins included *eating*, *gripping*, and *hygiene*. Subsequent winnowing of the items eliminated 551 that were redundant or duplicative, 447 that were condition-specific or not generalizable, 329 that were confusing or unclear, and 332 that were inconsistent with the physical function definition. Several additional items were eliminated for miscellaneous or idiosyncratic reasons. For example, items were eliminated if they described unusually strenuous tasks or included social interaction as such items might be confounded with non-health-related factors.

1.1.1.5 Item Review and revision

Further evaluation of item inclusiveness, clarity, comprehension, recall period, and response options were then conducted using reporter input via focus groups, mailed questionnaires, and cognitive interviews.[5]

1.1.1.5.1 Focus groups

To elicit concepts and identify gaps in coverage of the physical function domain, three focus groups were conducted by the Stanford PROMIS Primary Research Site. Focus group participants were recruited from flyers in clinics, listservs, and through letters of invitation to general medical rehab patients.[4,5] Focus group participants (N=15) ranged in age from 31 to 80 years, were 80% female, and 93% white. Participants were asked open-ended questions to permit the free exchange of ideas, and themes were subsequently identified through qualitative analysis of participant discussion. The emergent themes were cross-referenced against content in the item pool to confirm that all of the primary themes were covered.

1.1.1.5.2 Mailed questionnaires

In addition to focus groups, further input was obtained via questionnaires mailed to randomly selected subjects from the PROMIS Primary Research Site ARAMIS study cohort affiliated with Stanford University (N=1,100). These participants were drawn from rheumatoid arthritis, osteoarthritis, and healthy aging cohorts; they were 87% white, 53% female, and had an average education of 16 years. Each participant was asked to rate 30-item subsets of the remaining 1,860 candidate items for clarity (from "totally clear" to "not clear at all"), for personal importance and relevance (from "very important" to "not important at all"), and for comprehensibility by being asked to describe the meaning of each item in their own words. Respondents were also asked to rate their preferences regarding other features of the items.[4] One of these related to the item-**recall period**. More than 10 different recall periods had been used in the instruments from which the original pool of items was drawn. More than half of the 1,860 items (n = 975) did not specify a recall period or were prefaced with "the present", "today", or "now." The remaining items used recall periods that ranged from "in the last 2 days" to "in the past year." Questionnaire participants generally preferred the use of the present tense. Another feature rated by questionnaire participants related to **response options**.[4] The original pool of 1,860 items used 262 uniquely worded response scales. These ranged from asking about the presence or absence (i.e., yes/no), degree of severity (none to extreme), frequency (never to always), or ability (without any difficulty to unable to do). The number of response options ranged from two (yes/no) to 101 (0 to 100 on a visual analog scale [VAS]). Questionnaire participants preferred having four or five response options. Based on results from the mailed questionnaires, 37 items were identified for further evaluation because they had been rated as "unclear" by more than 5% of the participants and/or they had been rated as "totally clear" by fewer than 80% of respondents.

1.1.1.5.3 Cognitive interviews
Each of the 37 items selected for further assessment was evaluated by three to six participants in the cognitive interviews (with intentional variability by gender, age, race, and education). Cognitive interviews were conducted by the Stanford group to assess relevance and comprehension of the PROMIS Physical Function items. Using a semi-structured interview protocol, investigators conducted phone interviews with 18 individuals drawn from the PROMIS study cohort of rheumatoid arthritis, osteoarthritis, and healthy aging participants.[4] Cognitive interview participants ranged in age from 48 to 93 years, were 67% female, and 67% white.

Taken together, results from the focus groups, questionnaires, and cognitive interviews informed item elimination and item revision to enhance clarity and comprehension.[4] Additional changes were made in order to minimize variation in item attribution, to make use of a present tense (no recall) language, and to reduce the number of response scales from 262 to two (from "without any difficulty" to "unable to do" and from "not at all" to "cannot do"). **At the conclusion of the item review and revision phase, 168 items remained in the PROMIS Physical Function item bank that underwent evaluation and field testing in various clinical samples as described in Section 1.1.1.5.4, and notably cancer, which is relevant to this submission and detailed as part of this Qualification Plan in Section 1.1.2 & 1.1.3.**

1.1.1.5.4 Large-scale field testing
The 168 remaining candidate items were field tested by 21,133 subjects.[5] Ninety-three percent of the sample (n = 19,601) was recruited by Polimetrix, a polling firm based in Palo Alto, California; the remainder was recruited by the Stanford PROMIS Research Site study cohort and the North Carolina PROMIS Research Site study cohort (n = 1,532). The Polimetrix sample was designed to reflect Year 2000 United States Census demographics. Two-thirds of these subjects (n = 13,250) were drawn from the general United States population. The remaining Polimetrix subjects were recruited from clinical samples of individuals with cancer (n = 1,754), chronic obstructive pulmonary disease (n = 1,214), psychiatric illness (n = 1,193), heart disease (n = 1,156), osteoarthritis (n = 918), rheumatoid arthritis (n = 557), spinal cord injury (n = 531), and other conditions (n = 560). Overall, the field testing sample had a median age of 50 years, was 52% female, 82% white, and 97% had a high school education or more. In addition to the candidate items, subsets of the field-testing participants were administered 30 items from two legacy tools (20 items from the Health Assessment Questionnaire-Disability Index [HAQ-DI][6] and 10 items from the SF-36 Health Survey Physical Functioning [PF-10][7]) or one physical function item from the PROMIS Global Health measure.[5]

Data from the field-testing sample were used for a range of quantitative analyses.[5] Factor analyses suggested that, despite the preliminary identification of four physical function subdomains, these factors were highly correlated. The candidate items were also evaluated for skewness, unidimensionality and local dependence, differential item functioning (by disease group), and monotonicity. Based on these analyses, 44 items were subsequently removed for various reasons including poor item fit, copyright issues, and evidence of variable item functioning.

The remaining 124 items were then *calibrated* and placed on the same metric by applying a unidimensional model based on item response theory (IRT)[5]. IRT is a family of mathematical models that estimate unique properties for each item response category relative to a single dimension that is measured by all the items.[8,9] These properties ("parameters") are based on how likely people with different levels of the measured trait are to endorse categories of an item. The application of IRT is both convenient and accurate, for it permits users to administer any subset of items from a bank and compare scores on a common metric. Because IRT assumptions for unidimensionality are stringent,[10] the items in PROMIS banks are highly inter-correlated, yielding a single dimension that explains the large majority of variance in person-level scores. PROMIS Physical Function items were calibrated using the graded response model (GRM).[11] The GRM estimates both item location (level of health) and discrimination (ability to distinguish people at different levels of health) parameters. These calibrations were then used in simulation studies in order to describe properties of the PROMIS Physical Function item bank (and computer adaptive tests [CATs] that can be developed from it) as well as to construct fixed-length short forms.

Development of the short forms was also informed by the strength of associations with the legacy measures (HAQ-DI and SF-36 PF) administered to the same sample. The PROMIS PF bank is highly correlated with these two legacy measures (absolute value of Pearson r 0.80 - 0.88).[3]

1.1.1.6 Subsequent versions of the PROMIS Physical Function item bank

The PROMIS Physical Function item bank has undergone two minor revisions since it was initially developed. The original version (v1.0) included 124 items that were tested and validated in a large diverse sample (including both a mixed-disease subsample [n = 8,633] and a general population subsample that was representative of the U.S. [n = 8,493]).[2] Subsequent efforts to translate the item bank into several languages other than English led to minor modifications to 19 items (e.g., metric equivalents were added to references for distances and weights such as "over 10 pounds / 5 kg"); these changes represent version 1.1. Version 1.2 resulted from the elimination of three items (PFA7, PFC20, and PFC34) due to the potential for proprietary restrictions in their use. The current version (2.0) incorporated the Global06 item from the PROMIS Scale v1.2 Global Health, three items from the PASTOR (Pain Assessment Screening Tool and Outcomes Registry) project, and the reintroduction of eight items from the first version of the Physical Function item bank, because they had acceptable measurement properties once the PASTOR items were added.

Scores across the PROMIS Physical Function versions can be compared to each other. On a set of common items (e.g., between v1.2 & v2.0), we observed a maximum difference in an impaired sample of ¼ of a T score point using old and new parameters. Thus, the Physical Functions scores on v1.0, v1.1, v1.2 and v2.0 are comparable.

1.1.2 Development of the PROMIS-Ca Physical Function item bank

Not long after the PROMIS Physical Function item bank (v1.0) was developed for universal application, a separate (but overlapping) team of experts – the Cancer PROMIS Supplement (CaPS) working group -- was funded by the National Cancer Institute (NCI) to identify (and create, as needed) physical function items with particular relevance to cancer.[12] Like the general item bank, this process followed PROMIS standard procedures, including the use of concept elicitation focus groups and interviews, expert input, and large-scale field-testing. These procedures are described in detail below.

1.1.2.1 Concept elicitation

Two focus groups with cancer patients were conducted to elicit health concerns and concepts most relevant to their condition. The focus group sessions comprised 21 patients, divided between a public health care setting (Chicago Cook County Hospital; n=10) and a private nonprofit health care provider (Evanston-Northwestern Healthcare; n=11). Following completion of the focus groups, 40 individual patient interviews were conducted to elicit definitions and ensure coverage of cancer-relevant physical function concerns. Table 1 outlines characteristics of the focus group and concept elicitation interview participants.

Table 1. Concept elicitation participant characteristics from the development of cancer-specific PROMIS Physical Function metric[13]

	Focus Groups (2)	Patient Interviews
Sample size	21	40
Gender	8 male / 13 female	11 male / 29 female
Age	36 to 80; mean 56	28-77; mean 48
Time since diagnosis	Within 1 year: 7 1 to 5 years: 9 beyond 5 years: 5	Within 1 year: 21 1 to 5 years: 11 beyond 5 years: 8
Cancer type	Breast: 7 colorectal: 6 head & neck: 2	Breast: 12 colorectal: 8 lung: 6

	prostate: 2 melanoma: 1 liver: 1 Hodgkin's lymphoma: 1	head & neck: 3 non-Hodgkin's lymphoma: 3 prostate: pancreatic: 1 stomach: 1 urological: 1 gynecological: 3 Hodgkin's lymphoma: 1 neuroendocrine: 1
Metastases	No: 9; Yes: 8; unsure: 4	No: 23; Yes: 16; unsure: 1
Recurrence	No: 9; Yes: 9; unsure: 3	No: 30; Yes: 9; unsure: 1
On treatment	No: 3; Yes: 17; unsure: 1	No: 6; Yes: 33; unsure: 1

Findings of the concept elicitation focus groups and interviews revealed that overall, the PROMIS Physical Function Item Bank provided good coverage of the cancer-specific physical function concerns. However, three new items were created in light of the cancer-specific qualitative findings.

1.1.2.2 Cognitive interviews

As described in Section 1.1.2.1, three new items were developed based on findings from the concept elicitation focus groups and interviews. These three items were subsequently evaluated for clarity and comprehension via cognitive interviews with five individuals diagnosed with cancer recruited from Evanston-Northwestern Healthcare.

1.1.2.3 Expert input

Following concept elicitation phase and cognitive interviews, an expert panel of oncologists, pharmacists, psychologists, and CaPS researchers with expertise in PRO development reviewed the pool of 171 items (168 general physical function candidate items and the three newly developed cancer-specific items) for their relevance to cancer.

Based upon expert input and psychometric analyses of the field-testing data for the general PROMIS Physical Function item bank, 37 items were excluded from the cancer-specific item pool due to irrelevance or probable infrequency of concern in cancer patients (e.g., grip problems) and 73 additional items were excluded for psychometric reasons or conceptual redundancy.

1.1.2.4 Field testing

Field testing of the remaining 61 cancer-specific candidate items proceeded with a large cancer convenience sample (any cancer type, stage, or diagnostic status; N = 521, see Table 2 for sample characteristics). Participants were administered all 61 items. Cancer-specific field testing data were consequently merged with the data obtained from the general population participants in the original development of PROMIS and tested for differential item functioning (DIF) by sample. The majority of items showed only trivial DIF by sample, supporting the use of identical parameters for both the general and cancer populations. After comprehensive psychometric analyses of the field-tested items, 14 of the 61 items were dropped based on psychometric concerns specific to the cancer sample, such as non-trivial DIF, local dependency, ceiling effects, and/or low discrimination values, leaving 47 items for calibration of the PROMIS Physical Function Cancer item bank. Two items were subsequently deleted due to intellectual property concerns, yielding a PROMIS-Ca Physical Function item bank with 45 items specifically relevant to cancer.

Table 2. Demographic and clinical characteristics cancer field-testing sample

Sample size	521
Sex	140 male / 380 female (1 missing)
Age	18-80; mean = 55
Cancer type	breast: 177 urological: 96 gynecological: 56 colorectal: 46 non-Hodgkin's lymphoma: 27 lung: 21 Hodgkin's lymphoma: 11 head and neck: 10 leukemia: 9 brain: 8 multiple myeloma: 7 prostate: 7 esophageal: 4 melanoma: 4 testicular: 4 bone/muscle: 2 pancreatic: 2 gastric: 2 non-melanoma skin: 1 other: 27
Metastases	No: 310; Yes: 154; unsure: 57
Recurrence	No: 335; Yes: 141; unsure: 45
On treatment	No: 202; Yes: 309; unsure: 10

1.1.2.5 Subsequent versions of the PROMIS-Ca Physical Function item bank

As described in Section 1.1.1.5, two minor revisions were made to the general PROMIS Physical Function item bank that are also reflected in the PROMIS-Ca Physical Function item bank. Specifically, the original version of the PROMIS-Ca Physical Function item bank (v1.0) included 45 items that were tested and validated in a large diverse sample. However, subsequent efforts to translate the items led to minor modifications (v 1.1). The current version (v1.2) resulted from the elimination of three items (PFA7, PFC20, and PFC34) due to the potential for proprietary restrictions in their use.

Of the 165 items currently available in the PROMIS Physical Function bank (v2.0), 38 had been identified in 2007 as particularly relevant to cancer by the Cancer PROMIS Supplement (CaPS) working group.[12] The CaPS group also identified an additional seven items as useful for the assessment of cancer, but these items are not currently included in the PROMIS Physical Function bank (v2.0).

1.1.3 PROMIS-Ca Physical Function short forms

1.1.3.1 PROMIS short form v1.2 Physical Function 10b

The PROMIS® Short Form v1.2 Physical Function 10b (also known as the PROMIS Physical Function 10-item Short Form for Cancer) was subsequently identified in 2011[14] for the purpose of estimating important (meaningful) differences in physical function specifically for advanced-stage cancer patients. These 10 items were

a carefully selected subset of the items identified by the PROMIS CaPS working group; all of these items are also in the PROMIS Physical Function item bank (v1.2).[14] Once again, the item selection process involved review by multidisciplinary panels of clinical experts working in oncology (i.e., psychologists, nurses, physicians, pharmacists); this time input was obtained regarding whether the items were clinically relevant for use in the assessment of patient concerns regarding physical function, particularly in terms of content coverage and identifying cases in need of intervention. The PROMIS Physical Function 10-item Short Form for Cancer was originally submitted in the Initial Briefing Package for FDA DDT #079 by Northwestern researchers on May 9, 2016 (see Section 1.1.4).

1.1.3.2 PROMIS short form v1.2 Physical Function 16

After the establishment of the PROMIS-Ca item bank, the Georgetown PROMIS site (A. Potosky, PI) conducted a large-scale psychometric evaluation of a subset of the PROMIS Physical Function measures in a diverse, population-based sample of 4,840 cancer patients, which informed the creation of a custom 16-item PROMIS Physical Function short-form. The reliability (internal consistency) of the 4a, 6b, 10a, and custom 16-item PROMIS Physical Function short forms was high, with Cronbach's coefficient alphas of 0.92-0.96 (see Table 3).[13] Mean standard errors for the IRT-based scores were 2.2-3.9. In another dataset that combined two observational studies of advanced cancer patients receiving chemotherapy, Cronbach's alpha for the 10-item PROMIS Physical Function Short Form for Cancer was 0.93 and the mean standard error was 2.5.[14] In the subset of 114 patients who reported that their physical functioning was about the same 6-12 weeks later, the intraclass correlation coefficient for test-retest reliability was 0.75 (95% CI: 0.66 – 0.82). Inter-rater reliability is not applicable as this assessment is a "self-reported" instrument.

PROMIS Physical Function demonstrated differences between clinically distinct known groups defined by clinical and self-reported variables.[13] Cancer patients diagnosed with advanced cancer and those who received chemotherapy reported significantly lower physical function scores. Lung cancer patients had the lowest physical function scores, while men with prostate cancer had the highest mean physical function scores. Patients with a greater number of self-reported comorbidities reported worse physical function scores.

Table 3. Internal Consistency Reliability of 4, 6, 10, and 16 item PROMIS Physical Function Short Forms

PROMIS Physical Function Short Form	Jensen et al., 2015 (N = 4,840)	
	Cronbach's alpha	Mean SE
PF10b-Cancer	–	–
PF4a	0.92	3.9
PF6b	0.94	3.1
PF10a	0.92	2.7
PF16	0.96	2.2

1.1.4 Overview of COA submission history

On October 28, 2015, at the encouragement of the U.S. Food and Drug Administration (FDA), a team of researchers from the Department of Medical Social Sciences (MSS) at Northwestern University Feinberg School of Medicine submitted a Letter of Intent to the FDA Drug Development Tools (DDT) Clinical Outcome Assessment (COA) qualification program. The Northwestern team proposed the PROMIS Short Form v1.2 Physical Function 10b (described in Section 1.1.3.1) as a COA for use in adults with advanced solid tumors. The Center for Drug Evaluation and Research acknowledged receipt of the Letter of Intent (LOI) on November 3, 2015 by requesting the submission of an Initial Briefing Package (IBP). The IBP (DDT# 079) was prepared and submitted to the Clinical Outcome Assessment (COA) Qualification Program on May 9, 2016. Following the submission of the IBP, the Northwestern team convened a group of 10 pharmaceutical industry partners to support and advise on the further development and submission of the qualification package.

On October 17, 2016, the Northwestern team received a letter from the FDA detailing comments on the IBP submission, in which the FDA identified two specific concerns about the measure:

1. Concern that some of the examples of physical activities may not be applicable to all backgrounds/cultures (i.e., vacuuming, yard work, bowling, or playing golf). The FDA recommended that items be applicable to the majority of patients, independent of their cultures/backgrounds (i.e., bathing and dressing (item PFA6), shopping (item PFA53), lifting (item PFB28r1), walking, and preparing simple meals).
2. Concern that the recall period is not clearly specified on the PROMIS Physical Function Short Form 10b. The FDA expressed concern that, without a clearly-specified recall period, the treatment benefit may not be reliably evaluated in an interventional study, as patients may use various reference time points to answer the questions.
    a. The FDA requested information on how patients qualitatively describe the time periods they are considering when they answer this item bank without a directed recall period and requested information about qualitative study subjects' demographics and cultures be provided.

Following a November 16, 2016 industry partner meeting, the Northwestern team submitted a response letter on December 8, 2016 outlining the proposal of a panel survey study and international cognitive interview study to address FDA comments on the IBP. During a teleconference held on December 15, 2016, the Qualification Review Team (QRT) offered to work closely with the submission team by reviewing and providing input on the planned procedures, items for potential replacement, and materials for the proposed qualitative and quantitative studies.

Following the December meeting with the FDA QRT team, the Northwestern study team generated a "CUE Sheet" containing all items from the PROMIS Physical Function item bank, including cancer calibration statistics and items recommended for consideration by the Northwestern team based on item performance statistics. The CUE sheet was submitted to the FDA QRT for review on January 20, 2017. The FDA then recommended items expected to be applicable across cultures from the items recommended by the Northwestern team. On February 23, 2017, the FDA QRT sent the Northwestern study team a modified CUE sheet, which identified the FDA QRT's recommended items for inclusion in the custom short form for cognitive interview evaluation. In light of the FDA's recommendations and input from industry partners, the Northwestern study team combined the 10 additional items selected by the FDA QRT and items from the PROMIS Physical Function 10b (proposed in IBP submission) and the 16-item custom short form used in a large cancer population reference study to create a PROMIS Physical Function 31-item subset. The panel survey study and international cognitive interview study evaluated the recall period and applicability of these 31 items that were jointly selected by Northwestern and the FDA QRT.

The 21st Century Cures Legislation passed on December 13, 2016, which necessitated changes to the DDT COA Qualification process. In response to the legislation, on September 28, 2017, the FDA QRT sent a draft outline of the newly required Qualification Plan to the Northwestern team in order to comply with and implement the new regulations from the 21st Century Cures Legislation. The Northwestern team and pharmaceutical partners reorganized submission content from the IBP, incorporated results from the recent panel survey study, and outlined the methods of the international cognitive interview study (data collection ongoing at the time) into the Qualification Plan outline provided by the FDA QRT. The Northwestern team submitted the draft Qualification Plan for DDT# 079 to the FDA QRT for review on December 21, 2017.

Meanwhile, on May 4, 2018 the Northwestern team submitted an interim study report describing methods and results of the panel survey study and international cognitive interview study for review by the FDA QRT.

During a teleconference held on July 13, 2018, the FDA QRT, Northwestern team, and pharmaceutical partners discussed recall period, preliminary item reduction, plans for final item reduction, and restructuring of the Qualification Plan. Conclusion of the teleconference yielded three decisions/actions items:

1. The custom short form submitted in the Final Qualification Package will include a 1-week recall period.
2. The Northwestern and pharmaceutical partner submission team will reformat the Qualification Plan chronologically and include a detailed plan for reducing the 31-item set to a 7-10 item short form and outline the development of the PROMIS item bank in chronological order under Section 1.
3. Schedule a face-to-face meeting with the FDA, Northwestern representatives, pharmaceutical representatives, along with expert and patient stakeholders to participate in the selection of 7-10 items (of 31 candidate items) for a PROMIS Physical Function COA to be submitted for use as a DDT COA in advanced oncology.

On September 30, 2018, the Northwestern team submitted a revised Qualification Plan to support the Food and Drug Administration's consideration of a custom PROMIS® Physical Function short form as a qualified COA DDT in the setting of advanced cancer treatment for stakeholder review in advance of the upcoming face-to-face item selection meeting.

On Monday, October 22, 2018, representatives from Northwestern University, FDA, industry, independent clinicians, and two patient stakeholders convened a half-day in-person meeting in the Washington, DC area. By the end of the meeting, stakeholders collaboratively selected an 8-item **PROMIS Short Form v2.0 – Physical Function 8c** for qualification. At the conclusion of the meeting, the FDA confirmed that no additional research would be required for the submission of the Qualification Plan and requested an amendment to the September 30th 2018 version of the Qualification Plan detailing the item selection and reduction process, which was submitted to the agency on December 6, 2018. Subsequently, on December 17, 2018, NU submitted a report of differential item function (DIF) analyses of measure time recall period context (e.g., 7-day recall period or no recall period) using data collected as a part of the agency's recommended research.

After review of the revised QP and DIF analyses, the FDA submitted two information requests: 1) revised QP that contains an updated quantitative analysis plan for the newly selected items (December 20, 2018); 2) additional supplementary quantitative analyses of the new instrument (January 11, 2019). On February 11, 2019, the Northwestern team submitted a third version of the QP that incorporated the requested information.

On May 13, 2019, the QRT and Northwestern team participated in a teleconference to discuss further required revisions to the QP, and on July 4, 2019, the Northwestern team received a QRT Reviewability Memorandum, which outlined two pathways to DDT COA Qualification: "Limited Context of Use" and "Broader Context of Use," specifying the QP and FQP requirements for each path. At this time, **we are submitting the PROMIS Short Form v2.0 – Physical Function 8c for <u>Limited context of use</u>** and have revised this QP to suit the requirements from the reviewability memorandum. The requirements for the limited context of use approval outlined in the FDA memorandum are:

- The protocol for the recall study.
- The description of the analyses that are to be conducted on the recall study data set, explicitly noting things like which populations the analyses will be performed on and how missing data will be handled. Analyses should include essential psychometric properties, such as reliability and construct validity. These analyses should use the short form proposed for qualification. If other forms will also be used, this should be clearly listed and explained.
- The plan to calibrate the 7-day recall instrument based on the data from the recall study. This should include data from the general population and oncology population, and, if available, additional 7-day recall data.
- In addition to the DIF analyses, the proposed methods are to directly compare the parameters from these two recall time period-related models for the items that will be in the tool proposed for qualification (at a minimum).

The remainder of this qualification plan documents the considerable evidence supporting the use of the *PROMIS Physical Function* metric and development of the **PROMIS Short Form v2.0 – Physical Function 8c** completed as part of the Drug Development Tools (DDT) Qualification submission process. At this time, we are submitting the **PROMIS Short Form v2.0 – Physical Function 8c** for the Limited Context of Use DDT COA Qualification pathway.

1.2 Concept of Interest for Meaningful Treatment Benefit

The concept proposed as an indicator of treatment benefit is physical function among adults with advanced or metastatic cancer. Physical function, defined as "one's ability to carry out various activities that require physical capability, ranging from self-care (activities of daily living) to more vigorous activities that require increasing degrees of mobility, strength, or endurance" represents an important aspect of well-being across the physical health continuum.[3,15-17] There are several subdomains of physical function which contribute to one's ability to engage in daily activities, from basic to instrumental activities of daily living, and including mobility (lower extremity function), dexterity (upper extremity function), and axial (neck and back) function.[2]

Physical function may be affected by numerous acute and chronic health conditions, including advanced or metastatic cancer. Many people diagnosed with advanced or metastatic cancer have symptoms and associated limitations in physical function at the time of diagnosis, and many will develop such limitations at some point after diagnosis because of disease progression, treatment side effects, or both. Indeed, more than half of cancer survivors, even long after treatment, experience decrements in physical function.[18] These negative effects on physical function often extend long beyond completion of treatment and have been associated with increased mortality risk and quality of life impairment.[19] For example, recent findings from a study that employed patient-reported outcome (PRO) measures as part of symptom screening in a gynecologic oncology clinical practice revealed a lower level of physical function in women with gynecologic cancer compared to the general population and reported that impaired physical function was the most common reason for clinical alerts and referrals.[20]

Given the high proportion of advanced or metastatic cancer patients who experience decrements in physical functioning at some point in their disease course (and often at multiple points),[18] it is common for individuals in this population to cite physical function limitations as an important concern.[21] This is particularly true for patients undergoing treatments that are intended to extend the patient's life at the expense of reduced quality of life. One of the more vulnerable periods for physical function impairment is during the receipt of chemotherapy or other interventions for advanced disease, as both the disease and the treatment itself can exert deleterious effects.[22-29] This circumstance implies that evaluations of advanced and metastatic interventions depend, in part, on the need for precise and valid measurement of physical function to determine whether interventions can, by virtue of their symptom-inducing effect on the cancer, produce an overall benefit to the physical function of the patient.

1.3 Context of Use

1.3.1 Target study population

The target study population is as follows:

1. Adult patients (≥ 18 years of age) with advanced solid tumors or hematologic malignancies of any primary site or origin (i.e., Stage III-IV), including unknown primary site;
2. Performance status range 0-3;
3. Receiving active anti-cancer therapy;
4. Symptomatic or asymptomatic; and
5. Without limitation regarding language or background/culture of the patient.

1.3.2 Target study design

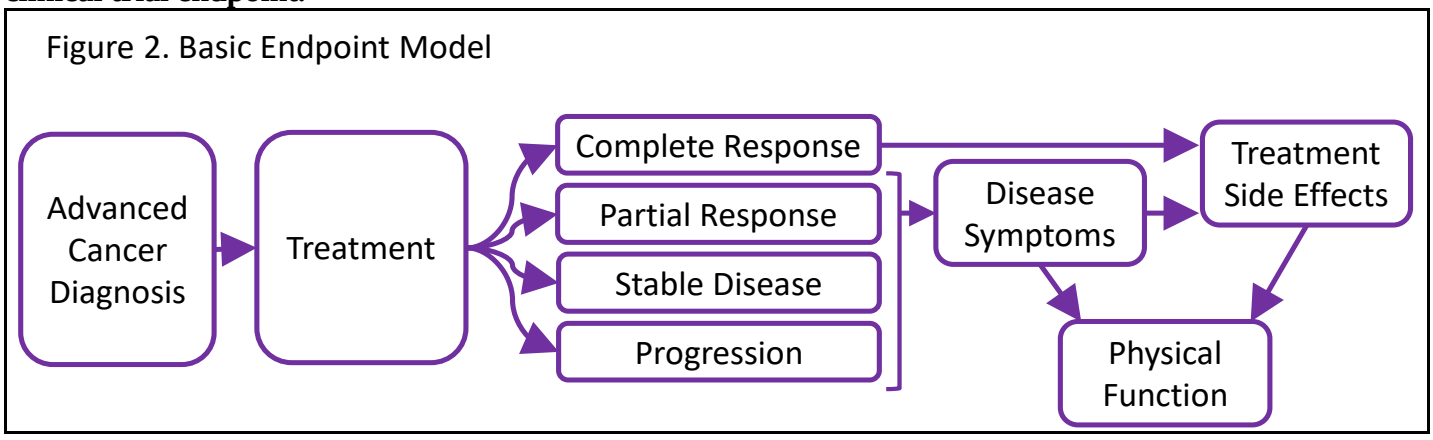
The target study design will be a comparison of an experimental treatment to a control treatment (placebo control or comparator). We expect that the proposed physical function outcome will typically be used as a secondary endpoint in randomized clinical trials in oncology, as a "common pathway" indicator of the combined impact of disease-related and treatment-related symptoms. The anticipated study design would have patients at baseline with

a mixture of disease-related symptoms and, perhaps, some residual treatment-related side effects (associated with prior surgery, radiation, or chemotherapy). Physical function, using the 8-item **PROMIS Short Form v2.0 – Physical Function 8c** developed as part of this DDT COA submission, would be measured along with these symptoms at baseline and then repeatedly while on study. Assessments would likely occur at intervals approximately one month apart and continue for variable lengths of time (depending upon the expected progression-free interval of the patient population). Typically, assessment much beyond the anticipated median survival of the experimental arm is less informative, as the sources that influence symptoms and physical function become increasingly variable over time. Baseline (pre-randomization) assessment should be followed by assessment tailored to the natural history of the disease and expected trajectory of physical function, as driven by other disease outcomes (i.e., expected symptoms, side effects, response, progression and survival). Generally, extending the final planned PF assessment more than a few months beyond the expected median PFS of the experimental group has limited value for the purpose of comparing treatment arms. For example, in a case where median progression free survival (PFS) in the comparator arm is six months, with an expectation of 8-9 month PFS in the experimental treatment, follow up assessments could be monthly for six months and quarterly out to one year post randomization. Thus, a typical assessment plan might include baseline measurement followed by subsequent assessments each month (or every three weeks for 21-day chemotherapy cycles) for 6-12 months (depending on the expected time to progression). This could be followed by quarterly assessments until the anticipated median survival time of the experimental group has passed.

1.3.3 Target study objectives

The proposed physical function outcome can be affected by both disease and treatment. In a typical endpoint model, we anticipate that disease symptoms and treatment-related adverse events would be the most proximal PRO endpoints, often linked to radiographic progression and ultimately survival. This physical function endpoint would most likely be positioned as a secondary PRO endpoint slightly more distal than disease symptoms and treatment side effects and very much influenced by them. In cases where disease- and treatment-related symptoms move in different directions and are difficult to disentangle, this physical function outcome can serve as a useful adjudicating endpoint. In such a case, the physical function endpoint can help evaluate overall clinical benefit by comparing treatments on the extent to which *overall* physical function is affected by one or another treatment. In other cases, it can help determine if consistent differences between treatment arms are associated with benefits to patients in terms of important everyday functions.

Figure 2. (Basic Endpoint Model) illustrates where physical function fits in the overall context of use as a clinical trial endpoint.



1.4 Critical Details of the Measure to the Degree Known

1.4.1 Reporter

PROMIS Physical Function is a self-report assessment tool. Individual patients rate themselves on statements ("items") about physical function. Patients choose the response option that most accurately describes their ability to perform the specific physical activity (or activities) described in each statement.

1.4.2 Item content or description of the measure

The **PROMIS Short Form v2.0 – Physical Function 8c** is an 8-item fixed length short form derived from the PROMIS Physical Function item bank. It has a recall period of the past seven days and includes a 5-point scale with three sets of response options. Six capability-framed questions are administered with one set of response options, "Without any difficulty" to "Unable to do," and two disability framed-questions are administered with two sets response options, "Not at all" to "Cannot do" and "No difficulty at all" to "Can't do because of health". Overall, the content of the eight items reflects the following concepts: activities of daily living, mobility, and global impact of physical functioning. Scores on the **PROMIS Short Form v2.0 – Physical Function 8c** are reported on a T score metric (mean = 50 and SD = 10), with higher scores reflecting better physical functioning.

1.4.3 Mode of administration & data collection method

All of the items in the **PROMIS Short Form v2.0 – Physical Function 8c**, *PROMIS-Ca Physical Function item bank,* and *PROMIS Physical Function item bank* are intended to be self-administered (i.e., they do not require an interviewer). Several modes of data collection are possible using the **PROMIS Short Form v2.0 – Physical Function 8c**, including paper-and-pencil; electronic data collection via tablet computer, smartphone, or similar device; and data capture via telephone using interactive voice response.

The specific data collection method will be dependent upon future clinical trial methodology and researchers' preference, as previous research has found no differential effect by method of data collection on items. For example, Bjorner et al.[30] compared four modes of administration using two non-overlapping parallel forms, each containing eight items from the PROMIS Physical Function item bank. The PROMIS items were administered to 923 adults with COPD, depression, or rheumatoid arthritis. The four modes included 1) paper-and-pencil, 2) automated telephone interview using interactive voice response, 3) computer connected to the internet, and 4) personal digital assistant (PDA). In difference score analyses, no significant mode differences were found, and all confidence intervals were within the pre-specified minimally important difference threshold of 0.2 SD. Parallel forms reliabilities were very high (ICC=0.85-0.93). Only one across-mode intra-class correlation (ICC) was significantly lower than the same-mode ICC. Tests of validity showed no differential effect by mode of administration, though participants preferred screen interface over the other modes.

1.5 Description of the Involvement of External Expertise, Including Scientific Communities or Other International Regulatory Agencies, if applicable (i.e., Working Group, Consortia)

PROMIS Physical Function item bank development. The team that developed the original PROMIS Physical Function item bank was led by James F. Fries (developer of the Health Assessment Questionnaire), John Ware (developer of the MOS SF-36), Ron D. Hays (developer of several MOS scales), Matthias Rose, and Jakob Bjorner, as well as several other investigators from across the PROMIS network. The following external experts and agencies provided input and expertise during the development and validation phases of the PROMIS Physical Function item bank: The U.S. Food and Drug Administration (two meetings held at FDA in 2006 and 2008), the OMERACT group, the National Cancer Institute (through science officer participation at steering committee meetings and CaPS project meetings at NCI in 2006), the World Health Organization and its International Classification of Functioning, Disability, and Health (coordinated through Bedirhan Ustun), and many consultants with considerable expertise in the measurement of physical functioning and health.

PROMIS-Ca Physical Function item bank development. The National Cancer Institute provided supplementary funding to an overlapping set of investigators with the PROMIS Network to investigate and evaluate PROMIS item bank validity in cancer patients and survivors across the continuum of care. Domain expert and patient input were obtained to in the process of evaluating and in an attempt to enhance the core set of PROMIS measures.[12]

PROMIS Short Form v2.0 – Physical Function 8c development. With regard to the development of the DDT #079 submission, several individuals and organizations have provided input at various stages:

After submission of the LOI and IBP for DDT #079, in which the Northwestern team submitted the PROMIS Physical Function 10b for qualification, the Northwestern team engaged both the agency's COA QRT team and convened a working group of 10 pharmaceutical companies and respective representatives as collaborators in the DDT submission process, including FDA recommended research, and selection of 31 existing items from the PROMIS Physical Function v2.0 item bank as candidates for the development of a custom short form most appropriate for use in advanced cancer patients across cultures and backgrounds.

As described in Section 1.1.4, a collaborative item reduction process was performed to reduce the pool from 31 candidate items to a subset of eight items. The in-person collaborative item selection process took place during a half-day meeting held on October 22, 2018 in the Washington D.C. area. Meeting attendees included Northwestern representatives, FDA representatives, NIH representatives, industry representatives, oncology clinicians, patients, and care partners.

Section 2: Executive Summary

This Qualification Plan follows the FDA's Clinical Outcome Assessment (COA) template posted publicly under Section 507 of the FD&C Act. Section 1 presents an overview of physical function in cancer and the clinical trial settings in which the COA would be used, the concept of interest for meaningful treatment benefit, the context of use (e.g., target population, study design, study objectives, endpoints), as well as the chronological history of research regarding the development of the general and cancer-specific PROMIS Physical Function item banks through selection of 31 candidate PROMIS Physical Function items in collaboration members of the FDA and representatives from 10 pharmaceutical companies**.** The Qualification Plan provides evidence of content validity and face validity both historically and through a recent international cognitive interview study with patient reporters across a variety of cultures and backgrounds collected as part of this DDT COA submission process to inform the reduction of 31 items to the 8-item **PROMIS Short Form v2.0 – Physical Function 8c**. The Qualification Plan also details quantitative and qualitative results regarding patient evaluation of the standard PROMIS "No Recall" time context. The Qualification Plan also presents the process for item reduction from 31 to eight items and the draft conceptual framework. The Qualification Plan details quantitative analysis plans leveraging the cross-sectional recall period study as evidence to support the **PROMIS Short Form v2.0 – Physical Function 8c,** including descriptive statistics, internal consistency reliability, convergent and discriminant validity, and known-groups validity. The **PROMIS Short Form v2.0 - Physical Function 8c** is being submitted for Qualification under Limited Context of Use, with only cross-sectional analyses planned. Per agreement with FDA and as detailed in the FDA Reviewability Memorandum, longitudinal analyses will be conducted after approval for limited context of use. The Qualification Plan describes information about PROMIS standards for translations and indicates the items in the **PROMIS Short Form v2.0 – Physical Function 8c**. The Qualification Plan concludes with questions to the FDA's Center for Drug Evaluation and Research (CDER).

References:

1. Kluetz PG, Slagle A, Papadopoulos E, et al. Focusing on Core Patient-Reported Outcomes in Cancer Clinical Trials: Symptomatic Adverse Events, Physical Function, and Disease-Related Symptoms. *Clinical Cancer Research.* 2016.
2. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *Journal of Clinical Epidemiology.* 2008;61(1):17-33.
3. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of clinical epidemiology.* 2010;63(11):1179-1194.
4. Bruce B, Fries, J. F., Ambrosini, D., Lingala, B., Gandek, B., Rose, M., & Ware Jr, J. E. Better assessment of physical function: item improvement is neglected but essential. *Arthritis Research and Therapy.* 2009;11(6):R191.
5. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware Jr JE. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *Journal of Clinical Epidemiology.* 2014;67(5):516-526.
6. Fries JF SP, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum.* 1980;23(137):e45.
7. Ware JE, Jr., Sherbourne CE. The MOS 36-item short form health survey (SF-36): I. Conceptual framework and item selection. *Med Care.* 1992;30:473-483.
8. Reeve BB, & Fayers, P. *Applying item response theory modeling for evaluating questionnaire item and scale properties.* New York, NY2005.
9. Revelle W, & Condon, D. M. . Reliability from alpha to omega: a tutorial. osf.io/e685p. Accessed.
10. Reeve B, Hays R, Bjorner J, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care.* 2007;45(5):S22-S31.
11. F. S. *Estimation of Latent Ability Using a Response Pattern of Graded Scores.* Richmond, VA.: Psychometric Society;1969.
12. Garcia SF, Cella D, Clauser SB, et al. Standardizing Patient-Reported Outcomes Assessment in Cancer Clinical Trials: A Patient-Reported Outcomes Measurement Information System Initiative. 2007;25(32):5106-5112.
13. Jensen RE, Potosky AL, Reeve BB, et al. Validation of the PROMIS physical function measures in a diverse US population-based cohort of cancer patients. *Quality of Life Research.* 2015;24(10):2333-2344.
14. Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *Journal of Clinical Epidemiology.* 2011;64(5):507-516.
15. Haley SM, Coster WJ, Binda-Sundberg K. Measuring physical disablement: the contextual challenge. *Physical Therapy.* 1994;74(5):443-451.
16. Haley SM, McHorney CA, Ware JE. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *Journal of clinical epidemiology.* 1994;47(6):671-684.
17. Fries J, Bruce B, Bjorner J, Rose M. More relevant, precise, and efficient items for assessment of physical function and disability: moving beyond the classic instruments. *Annals of the rheumatic diseases.* 2006;65(suppl 3):iii16-iii21.
18. Ness KK, Wall MM, Oakes JM, Robison LL, Gurney JG. Physical performance limitations and participation restrictions among cancer survivors: a population-based study. *Annals of epidemiology.* 2006;16(3):197-205.

19. Izano M, Satariano WA, Hiatt RA, Braithwaite D. The impact of functional limitations on long-term outcomes among African-American and white women with breast cancer: a cohort study. *BMJ open.* 2013;3(10):e003232.
20. Wagner LI, Schink J, Bass M, et al. Bringing PROMIS to practice: brief and precise symptom screening in ambulatory cancer care. *Cancer.* 2015;121(6):927-934.
21. Cella D, Rosenbloom SK, Beaumont JL, et al. Development and Validation of 11 Symptom Indexes to Evaluate Response to Chemotherapy for Advanced Cancer. *Journal of the National Comprehensive Cancer Network.* 2011;9(3):268-278.
22. Cella DF, Orofiamma B, Holland JC, et al. Relationship of psychological distress, extent of disease, and performance status in patients with lung cancer. *Cancer.* 1987;60:1661-1667.
23. Cella D, Eton D, Hensing TA, Masters GA, B. P. Relationship between symptom change, objective tumor measurements, and performance status during chemotherapy for advanced lung cancer. *Clinical Lung Cancer.* 2008;9(1):51-58.
24. Lilenbaum RC, Cashy J, Hensing T, Yount S, Cella D. Prevalence of poor performance status in lung cancer patients: implications for research. *Journal of Thoracic Oncology.* 2008;3(2):125-129.
25. Aitchison M, Bray C, Van Poppel H, et al. Adjuvant 5-flurouracil, alpha-interferon and interleukin-2 versus observation in patients at high risk of recurrence after nephrectomy for renal cell carcinoma: results of a phase lll randomized European Organization for Research and Treatment of Cancer (Genito-Urinary Cancers Group)/National Cancer Research Institute trial. *European Journal of Cancer.* 2014;50(1):70-77.
26. Hahn EA, Glendenning GA, Sorensen MV, et al. Quality of life in patients with newly diagnosed chronic phase chronic myeloid leukemia on imatinib versus interferon alfa plus low-dose cytarabine: results from the IRIS Study. *J Clin Oncol.* 2003;21(11):2138-2146.
27. Kirchheiner K, Nout R, Czajka-Pepl A, et al. Health related quality of life and patient reported symptoms before and during definitive radio (chemo) therapy using image-guided adaptive brachytherapy for locally advanced cervical cancer and early recovery- a mono-institutional prospective study. *Gynecologic Oncology.* 2015;136(3):415-423.
28. Reni M, Bonetto E, Cordio S, et al. Quality of life assessment in advanced pancreatic adenocarcinoma: result from a phase lll randomized trial *Pancreatology.* 2006;6(5):454-463.
29. Ringash J, Au H, Siu L, et al. Quality of life in patients with K-RAS wild-type colorectal cancer. *Cancer.* 2014;120(2):181-189.
30. Bjorner JB, Rose M, Gandek B, Stone AA, Junghaenel DU, Ware JE. Method of administration of PROMIS scales did not significantly impact score level, reliability, or validity. *Journal of Clinical Epidemiology.* 2014;67(1):108-113.