PHYSICAL FUNCTION
MEASURE DIFFERENCES

A brief guide to differences between the PROMIS® Physical Function instruments:

ADULT	PEDIATRIC	PARENT PROXY
PROMIS Bank v1.0 - Physical Function*	PROMIS Pediatric Bank v1.0 – Mobility*	PROMIS Parent Proxy Bank v1.0 – Mobility*
PROMIS Short Form v1.0 – Physical Function 4a*	PROMIS Pediatric Short Form v1.0 – Mobility 8a*	PROMIS Parent Proxy Short Form v1.0 – Mobility 8a*
PROMIS Short Form v1.0-Physical Function 6a*	PROMIS Pediatric Bank v1.0 – Upper Extremity*	PROMIS Parent Proxy Bank v1.0 – Upper Extremity*
PROMIS Short Form v1.0-Physical Function 8a*	PROMIS Pediatric Short Form v1.0 – Upper Extremity 8a*	PROMIS Parent Proxy Short Form v1.0 – Upper Extremity 8a*
PROMIS Short Form v1.0 – Physical Function 10a*	PROMIS Pediatric Bank v2.0 – Mobility	PROMIS Parent Proxy Bank v2.0 – Mobility
PROMIS Short Form v1.0 – Physical Function 12a	PROMIS Pediatric Bank v2.0 – Mobility (recommended)	PROMIS Parent Proxy Bank v2.0 – Mobility (recommended)
PROMIS Short Form v1.0 – Physical Function 20a*	PROMIS Pediatric Bank v2.0 – Mobility (screen-to-CAT)	PROMIS Parent Proxy Bank v2.0 – Mobility (screen-to-CAT)
PROMIS-Ca Bank v1.0 – Physical Function*	PROMIS Pediatric Short Form v2.0 – Mobility 8a	PROMIS Parent Proxy Short Form v2.0 – Mobility 8a
PROMIS-Ca Bank v1.1 – Physical Function	PROMIS Pediatric Bank v2.0 – Upper Extremity	PROMIS Parent Proxy Bank v2.0 – Upper Extremity
PROMIS Bank v1.0 – Physical Function for Samples with Mobility Aid Users	PROMIS Pediatric Bank v2.0 – Upper Extremity (recommended)	PROMIS Parent Proxy Bank v2.0 – Upper Extremity (recommended)
PROMIS Short Form v1.0 – Physical Function Samples with Mobility Aid Users 11a	PROMIS Pediatric Bank v2.0 – Upper Extremity (screen-to-CAT)	PROMIS Parent Proxy Bank v2.0 – Upper Extremity (screen-to-CAT)
PROMIS Bank v1.1 - Physical Function*	PROMIS Pediatric Short Form v2.0 – Upper Extremity 8a	PROMIS Parent Proxy Short Form v2.0 – Upper Extremity 8a
PROMIS Bank v1.2 – Physical Function*		
PROMIS Short Form v1.2 – Physical Function 6b*		
PROMIS Short Form v1.2 – Physical Function 8b*		
PROMIS Bank v1.2 – Mobility*		
PROMIS Bank v1.2 – Upper Extremity*		
PROMIS Bank v2.0 – Mobility*		
PROMIS Bank v2.1 - Mobility		
PROMIS Bank v2.0 - Physical Function		
PROMIS Bank v2.0 - Physical Function (recommended)		
PROMIS Bank v2.0 - Physical Function (screen-to-CAT)		
PROMIS Short Form v2.0 – Physical Function 4a		
PROMIS Short Form v2.0 – Physical Function 6b		
PROMIS Short Form v2.0 – Physical Function 8b		
PROMIS Short Form v2.0 – Physical Function 8c		
PROMIS Short Form v2.0 – Physical Function 8c 7-day		
PROMIS Short Form v2.0 – Physical Function 10a		
PROMIS Short Form v2.0 – Physical Function 10b		
PROMIS Short Form v2.0 – Physical Function 20a		
PROMIS Short Form v2.0 – Physical Function 24a (PROMIS HAQ)		
PROMIS Bank v2.0 – Upper Extremity*		
PROMIS Short Form v2.0 – Upper Extremity 7a*		
PROMIS Bank v2.1 – Upper Extremity		
PROMIS Bank v2.1 – Upper Extremity (recommended)		
PROMIS Bank v2.1 – Upper Extremity (screen-to-CAT)		
PROMIS Short Form v2.0 – Upper Extremity 7a*		
PROMIS Short Form v2.1 – Upper Extremity 7a		

*retired measure

ABOUT PHYSICAL FUNCTION

PROMIS Physical Function instruments measure self-reported capability rather than actual performance of physical activities. This includes the functioning of one's upper extremities (dexterity), lower extremities (walking or mobility), and central regions (neck, back), as well as instrumental activities of daily living, such as

running errands. A single Physical Function capability score is obtained from a short form. Each Physical Function instrument is appropriate for the adult general population and adults with chronic health conditions. The forms are universal rather than disease-specific. Each form assesses current function rather than function over a specified time period.

Physical Function instruments are available for adults (ages 18+), pediatric self-report (ages 8-17) and for parents serving as proxy reporters for their child (youth ages 5-17). Adult instruments (18+) include: Physical Function, Mobility, Upper Extremity, Physical Function – Cancer, and Physical Function for Samples with Mobility Aid Users instruments. Pediatric and parent proxy instruments were developed for each Physical Function sub-domains of Mobility and Upper Extremity.

MOBILITY

Focuses on activities of physical mobility such as getting out of bed or a chair to activities such as running. The adult Mobility measures include selected items from the full Physical Function item bank.

UPPER EXTREMITY

Focuses on activities that require use of the upper extremity including shoulder, arm, and hand activities. Examples include writing, using buttons, or opening containers. The adult Upper Extremity measures include selected items from the full Physical Function item bank.

PHYSICAL FUNCTION FOR SAMPLES WITH MOBILITY AID USERS

This item bank and short form include screening items about one's ability to stand and walk. Based upon one's response, some items may be skipped. These measures are intended for samples that may include those who use mobility aids like wheelchairs. It is not restricted for use to only those that use mobility aids.

For complete list of all PROMIS definitions, go to: http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/list-of-adult-measures

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing Physical Function: short forms and computer adaptive tests (CATs). When administering a short form, instruct participants to answer all of the items (i.e., questions or statements) presented. With a CAT, participant responses guide the system's choice of subsequent items from the full item bank (165 items in total in adult bank). Although items differ across respondents taking CAT, scores are comparable across participants.

Some administrators may prefer to ask the same question of all respondents or of the same respondent over time, to enable a more direct comparability across people or time. In these cases, or when paper administration is preferred, a short form would be more desirable than CAT. This guide provides information on all Physical Function short form and CAT instruments.

CAT: A minimum number of items (e.g., 4) must be answered in order to receive a score for Physical Function CAT. The response to the first item will guide the system's choice of the next item for the participant. The participant's response to the second item will dictate the selection of the following question, and so on. As

additional items are administered, the potential for error is reduced and confidence in the respondent's score increases. CAT will continue until either the standard error drops below a specified level (e.g., on the T-score metric 3.0), or the participant has answered the maximum number of questions (e.g., 12), whichever occurs first. For some CATs, specifically "recommended" and "screen-to-CAT" there are additional stopping rules. These include stopping when the standard error isn't improving much or if a respondent is asymptomatic. For details on the exact stopping rules for Physical Function CATs, see below.

CAT versus Short Form: Whether one uses a short form or CAT, the score metric is Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of physical function represented by all items in the item bank. When choosing between CAT and a short form, it is useful to consider the demands of computer-based assessment, and the psychological, physical, and cognitive burden placed on respondents as a result of the number of questions asked.

Figure 1 illustrates the correlations (strength of relationship) of the full v1.0 Physical Function bank with CAT and with v1.0 short forms of varying length. The correlation of CAT scores with the full bank score is greater than a short form of any length. A longer CAT or longer short form offers greater correlation, as well as greater precision. When evaluating precision, not all questions are equally informative. The flexibility of CAT to choose more informative questions offers more precision.

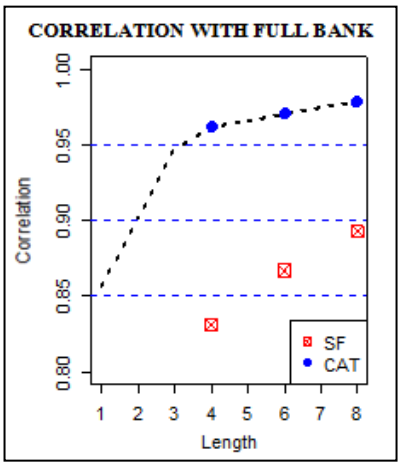

Figure 1

VERSION DIFFERENCES

Some PROMIS domains have multiple versions of instruments (i.e. v1.0, v1.1, v2.0). Generally, **it is recommended that you use the most recent version available which can be identified as the instrument with the highest version number.** In most cases, an instrument that has a decimal increase (v1.0 to v1.1) retains the same item-level parameters as well as instrument reliability and validity. In cases where a version number increases by a whole number (e.g., v1.0 to v2.0), the changes to the instrument are more substantial.

Adult

Physical Function: The PROMIS adult Physical Function v1.0 item bank included 124 items. Later, during translation into multiple languages, some items were modified (e.g., metric equivalents to measurements such as "over 10 pounds/5kg" were added). These 19 modifications resulted in the creation of a v1.1 item bank. Later, v1.2 was created by eliminating three items due to restrictions in use. This also resulted in replacing items in two short forms. PROMIS short form v1.0 – Physical Function 6a was replaced by PROMIS short form v1.2 – Physical Function 6b. One item is different. PROMIS short form v1.0 – Physical Function 8a was replaced by PROMIS short form v1.2 – Physical Function 8b. One item is different. The Mobility v1.2 and Upper Extremity v1.2 banks include the modifications for improved translations. Calibrations are identical in v1.0, v1.1, and v1.2 adult measures. The PROMIS-Ca v1.0 measure was also replaced by a v1.1 with items removed due to restrictions in use.

The Physical Function item bank v2.0 has 165 items. All item parameters were changed to match the Rose et al. (2014) publication. New items were added from three sources: 1) Global06 was added from the PROMIS Global-10; 2) 35 items were added from the PASTOR (Pain Assessment Screening Tool and Outcomes Registry) project; and 3) eight items originally excluded from v1.0 were re-introduced because they had acceptable measurement

properties once the PASTOR items were added. One of PASTOR's subprojects developed new content to measure "elite" physical function. Finally, two items were substituted out for similarly-worded items with better measurement properties. The Mobility v2.0 item bank was also updated to match the Rose et al. (2014) item calibrations, but no new content was added to it.

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

Mobility: The Mobility v2.1 item bank consists of 44 items, all of which are from the PROMIS Physical Function v2.0 item bank. Items include content that require lower extremities (e.g., walking, climbing stairs). There are more items in Mobility v2.1 than in the earlier Mobility v2.0 item bank. Consequently, it covers a wider range of function including both lower and higher levels of mobility. Scores obtained from Mobility v2.1 and Mobility v2.0 are comparable. However, because the range of function covered by v2.1 is larger, individual scores at the very high and very low end may be different. Version 2.1 is recommended over v2.0.

Upper Extremity: The Upper Extremity v1.2 item bank was revised to v2.0. First, additional items from Physical Function v2.0 item bank that were sufficiently related to upper extremity function were added. All v1.2 items were retained. This allowed for the creation of a 46-item bank and a 7-item short form. Second, all items were re-calibrated so that the scores reflected only upper extremity function. Moving Upper Extremity to its own metric improved measurement properties for individuals with known or suspected upper extremity limitations (though it remains centered on the USA general population). Conversely, the v1.2 items were calibrated to reflect overall physical function. One item (PFM16r) from the Upper Extremity v2.0 item bank and short form was modified to improve understandability by English speakers and improve translations. This change resulted in the creation of the Upper Extremity v2.1 measures. The v2.0 and v2.1 measures are otherwise identical.

Calibrations from Upper Extremity v1.0 – v1.2 cannot be compared with Upper Extremity v2.0 – v2.1. Calibrations are identical in PROMIS Upper Extremity v2.0 and v2.1 instruments, and therefore scores between PROMIS Upper Extremity v2.0 and v2.1 can be compared to each other. For users interested in comparing scores over time and desiring to switch from Upper Extremity v1.2 to Upper Extremity v2.1, we recommend rescoring the v1.2 administrations via HealthMeasures Scoring Service (https://www.assessmentcenter.net/ac_scoringservice) using the v2.1 calibrations. This will treat the previous administrations as a custom short form from the v2.1 item bank and allow score comparability over time.

Standard, Recommended, and Screen-to-CAT Stopping Rules: The standard, recommended, and screen-to-CAT Physical Function and Upper Extremity computer adaptive tests are based on the exact same item banks, but utilize different stopping rules. The PROMIS Bank v2.0 – Physical Function and PROMIS Bank v2.1 – Upper Extremity measures are administered by default as computer adaptive tests using the following standard stopping rules:

- Minimum number of items administered = 4
- Stop when one of these occurs:
    - 12 items are administered OR
    - Standard error is below 0.3 on the theta metric (3.0 on the T-score metric)

The PROMIS Bank v2.0 – Physical Function (recommended) and PROMIS Bank v2.1 – Upper Extremity (recommended) measures use the following stopping rules:

- Minimum number of items administered = 4
- Stop when one of these occurs:
    - 8 items are administered OR
    - Standard error is below 0.3 on the theta metric (3.0 on the T-score metric) OR
    - Standard error changes by less than 0.01 on the theta metric (0.1 on the T-score metric)

The PROMIS Bank v2.0 – Physical Function (screen-to-CAT) and PROMIS Bank v2.1 – Upper Extremity (screen-to-CAT) measures use the following stopping rules:

- If the response to the first item is the "healthiest" response, then stop.
- If the response to the first item is NOT the "healthiest" response, proceed with the "recommended" CAT stopping rules.

Pediatric and Parent Proxy
For Pediatric and Parent Proxy Mobility and Upper Extremity function, v2.0 measures replaced v1.0. The v2.0 measures 1) changed from using response scores of 0-4 to use 1-5 (item IDs amended with an "r") and 2) added new items (item IDs start with 7000). The calibrations between v1.0 and v2.0 are identical as is the item content on short forms.

Rose, M., Bjorner, J.B., Gandek, B., Bruce, B., Fries, J.F., & Ware Jr, J.E. (2014). The PROMIS Physical Function Item Bank Was Calibrated to a Standardized Metric and Shown to Improve Measurement Efficiency. Journal of Clinical Epidemiology, 67(5), 516-526. http://dx.doi.org/10.1016/j.jclinepi.2013.10.024

Standard, Recommended, and Screen-to-CAT Stopping Rules: The standard, recommended, and screen-to-CAT pediatric and parent proxy Mobility and Upper Extremity computer adaptive tests are based on the exact same item banks, but utilize different stopping rules. The PROMIS Pediatric Bank v2.0 – Mobility, PROMIS Pediatric Bank v2.0 – Upper Extremity, PROMIS Parent Proxy Bank v2.0 – Mobility, and PROMIS Parent Proxy Bank v2.0 – Upper Extremity measures are administered by default as computer adaptive tests using the following standard stopping rules:

- Minimum number of items administered = 5
- Stop when one of these occurs:
    - 12 items are administered OR
    - Standard error is below 0.4 on the theta metric (4.0 on the T-score metric)

PROMIS Pediatric Bank v2.0 – Mobility (recommended), PROMIS Pediatric Bank v2.0 – Upper Extremity (recommended), PROMIS Parent Proxy Bank v2.0 – Mobility (recommended), and PROMIS Parent Proxy Bank v2.0 – Upper Extremity (recommended) measures use the following stopping rules:

- Minimum number of items administered = 5
- Stop when one of these occurs:
    - 12 items are administered OR
    - Standard error is below 0.4 on the theta metric (4.0 on the T-score metric) OR
    - Standard error changes by less than 0.01 on the theta metric (0.1 on the T-score metric)

PROMIS Pediatric Bank v2.0 – Mobility (screen-to-CAT), PROMIS Pediatric Bank v2.0 – Upper Extremity (screen-to-CAT), PROMIS Parent Proxy Bank v2.0 – Mobility (screen-to-CAT), and PROMIS Parent Proxy Bank v2.0 – Upper Extremity (screen-to-CAT) measures use the following stopping rules:

- If the responses to the first two items are both the "healthiest" responses then stop.
- If the responses to the first two items are NOT the "healthiest" responses, proceed with the "recommended" CAT stopping rules.

SHORT FORM DIFFERENCES

Adult Profile Short Forms
There are 6 Physical Function short forms for adults; three (4a, 6b, and 8b) are included in the PROMIS Profiles. Items in the 4a, 6b, and 8b short forms were selected based on rankings using two psychometric criteria: 1) maximum interval information; and 2) CAT simulations. Item rankings were similar for both criteria. For the maximum interval criterion, each item information function was integrated (without weighting) for the interval from the mean to 2 SDs worse than the mean. For the CAT simulations, responses to all items in each bank were generated using a random sample of 1,000 simulees drawn separately for each bank (centered on 1.0 SD worse than the general population mean). Items were rank ordered based on their average administration rank over the simulees. Content experts reviewed the items and rankings and made cuts of 4, 6, and 8 items. For each domain, 8-items, 6-items, and 4-items have been selected so that the items are nested/overlap (e.g., the 8-item form is the 6-item form plus two additional items). The 4a, 6b, and 8b short forms can be administered with short forms of similar length from other domains (Depression, Pain Interference, Fatigue, Sleep Disturbance and Ability to Participate in Social Roles and Activities) as part of a PROMIS Profile (see PROMIS-29, 43, or 57 Profile v2.0), though they can also be administered individually.

Other Adult Short Forms
The original adult short forms (10a and 20a) were constructed by the domain team with a focus on representing the range of physical function and also representing the content of the item bank. Domain experts reviewed short forms to give input on the relevance of each item. Psychometric properties and clinical input were both considered.

The PROMIS Short Form v2.0 – Physical Function 24a (PROMIS HAQ) includes PROMIS items that are analogous to the Health Assessment Questionnaire Disability Index (HAQ-DI). The original developer of the HAQ-DI, James F. Fries, MD, was also the principal investigator who led the development of the PROMIS physical function measures for adults. Original HAQ items, in improved form, were incorporated into the PROMIS Physical Function v1.0 item bank. Scores from the original HAQ cannot be compared to PROMIS scores – they use different metrics. However, there is a table that can convert scores from the HAQ-DI to the PROMIS metric on the PROsetta Stone website (www.prosettastone.org).

The **PROMIS Short Form v2.0 – Physical Function 8c** includes 8 items from the v2.0 item bank. These items were selected in collaboration between PROMIS investigators at Northwestern University and the United States Food and Drug Administration (FDA) in October 2018. Candidate items from the Physical Function item bank were identified based upon input on relevance from people with cancer across five countries and literature-informed expert review. Items with good discrimination (slope) parameters were selected to cover the full range of the Physical Function continuum. The final 8 items were selected to maximize reliability across as full a range of physical function as possible.

The FDA prefers to provide a fixed recall period as a reference for respondents. Therefore, the **PROMIS Short Form v2.0 – Physical Function 8c 7-day** includes the same items, but adds a prompt to think about the past 7 days. PROMIS investigators at Northwestern University studied the impact of adding a 7-day recall to the standard PROMIS physical function items and found no evidence of impact, at the scale and item level. As of Feb

2022, the FDA has not made a final determination regarding which version will receive final FDA qualification as a Drug Development Tool. When possible, it is recommended that you use both measures. For those who are not able to use both measures in a study, the PROMIS Short Form v2.0 – Physical Function 8c is recommended.

The PROMIS Short Form v2.0 – Physical Function 8c and PROMIS Short Form v2.0 – Physical Function 8c 7-day are intended for use in clinical trials of people with advanced cancer. They may work well in other settings as well.

Starting with the PROMIS v1.0 pool of universal physical function items, members of the PROMIS-Cancer team created a PROMIS Physical Function Cancer item bank using cancer-specific focus groups, expert reviewers and large-scale field-testing. The PROMIS-Cancer team subsequently selected 10 candidate items from this bank for a short form that were reviewed by multidisciplinary panels of clinical experts working in oncology (consisting of psychologists, nurses, physicians, and pharmacists). All of these items were confirmed to be clinically relevant for use in assessment of patient concerns regarding physical function, particularly in terms of content coverage and identifying cases in need of intervention. Items in PROMIS Physical Function Short Form 10b assess universal physical function. That is, no item refers to specifically to cancer. The measure is appropriate for use across chronic conditions.

Pediatric and Parent Proxy Short Forms
There are two 8-item Pediatric and two 8-item Parent Proxy short forms, one pediatric and one parent proxy each for Mobility and Upper Extremity. Items were selected based on content and psychometric characteristics.

Selecting a Short Form
In selecting between short forms, the difference is instrument length. The reliability and precision of the short forms within a domain is highly similar. If you are working with a sample in which you want the most precise measure, select the longest short form. If you have little room for additional measures but really wanted to capture something as a secondary outcome, select one of the shorter instruments (e.g., 4-item short form).

SELECTING THE PHYSICAL FUNCTION FOR SAMPLES WITH MOBILITY AID USERS INSTRUMENTS

The Physical Function for Samples with Mobility Aid Users instruments are intended for samples in which some participants utilize mobility aids such as wheelchairs. There are two screening questions that ask about one's ability stand and to walk. Based on responses to the screening items, the following questions are tailored. Specifically, if an individual is not able to walk or stand, items in the item bank asking about being able to walk specific distances or jog are not eligible for administration in the CAT. Likewise, short forms will have respondents skip not applicable items. Note that this is not a bank intended only for those who use mobility aids. No items ask specifically about mobility aid use.

PROMIS ADULT CANCER MEASURES

PROMIS-Cancer (PROMIS-Ca) measures (Physical Function, Fatigue, Pain Interference, Depression and Anxiety) were developed under the PROMIS Cancer Supplement (CaPS) grant from NCI. The measures are highly similar to PROMIS measures. Some banks include unique items. In rare instances, a shared item uses different item-level calibrations in each bank.

- PROMIS-Ca Bank v1.1 - Physical Function contains 45 items, 33 of which are also in PROMIS Bank v2.0 - Physical Function.
- PROMIS-Ca Bank v1.0 - Fatigue contains 54 items, all of which are from PROMIS Bank v1.0 - Fatigue.
- PROMIS-Ca Bank v1.0 - Anxiety contains 22 items; 20 items from PROMIS Bank v1.0 - Anxiety, and 2 items unique to CaPS in which cancer specific calibrations were used: EDANX09 & EDANX39.
- PROMIS-Ca Bank v1.0 - Depression item bank contains 30 items; 23 items are from PROMIS Bank v1.0 - Depression and 7 items unique to CaPS in which cancer specific calibrations were used: EDANG09, EDANG29, EDDEP02, EDDEP12, EDDEP16, EDDEP38 & EDDEP55.
- PROMIS-Ca Bank v1.1 - Pain Interference contains 35 items; 32 items from PROMIS Bank v1.1 - Pain Interference v1.1 and 3 items unique to CaPS in which cancer specific calibrations were used: PAININ4, PAININ15 & PAININ30.

PROMIS-Cancer (PROMIS-Ca) measures were developed by having content experts review the adult PROMIS item banks for anxiety, depression, fatigue, pain interference, and physical function. Items were selected through expert consensus and informed by focus groups and cognitive interviews with cancer patients. Multidisciplinary clinical input was obtained to ensure content coverage and the relevance of PROMIS items to patients' cancer and/or cancer treatment experiences. Items' psychometric properties were reviewed when applicable. Next, calibration testing was conducted with cancer patients with different diagnoses and treatments. Data were analyzed to identify if items performed differently in people with cancer than people with other chronic conditions or in the general population. In most cases, PROMIS calibrations ("PROMIS Wave 1") were retained. In rare cases where differential item functioning was identified, calibrations for that item were revised for when that item is used in the PROMIS-Ca item bank. For items that exist only in a PROMIS-Ca item bank, new calibrations were created by using a fixed parameter linking strategy. This set of calibrations is named "Cancer" in the HealthMeasures Scoring Service.

A fixed parameter linking approach was taken because of the additional analyses that were conducted to evaluate the differences between the PROMIS item bank and the PROMIS-Ca item bank. The measures produce slightly different scores. This difference was determined to be so small that comparing scores from a PROMIS measure and PROMIS-Ca measure is acceptable. Because the PROMIS measures have demonstrated validity across diverse patient populations, are linked with other PRO measures (i.e., PROsetta Stone), and have continued to be improved through item bank expansion (e.g., PROMIS Physical Function item bank v2.0), it is recommended to use the general population PROMIS calibrations when assessing individuals with cancer.

SELECTING A PEDIATRIC OR PARENT PROXY INSTRUMENT

In selecting whether to use the pediatric versus parent proxy instrument for this domain, it is important to consider both the population and the domain which you are studying. Pediatric self-report should be considered the standard for measuring patient-reported outcomes among children. However, circumstances exist when the child is too young, cognitively impaired, or too ill to complete a patient-reported outcome instrument. While information derived from self-report and proxy-report is not equivalent, it is optimal to assess both the child and the parent since their perspectives may be independently related to healthcare utilization, risk factors, and quality of care.

WHICH CALIBRATION SAMPLE SHOULD I USE?

Some PROMIS Parent Proxy instruments (Anxiety, Depressive Symptoms, Fatigue, Mobility, Pain Interference, Peer Relationships) have two calibration samples – "Parent Proxy" and "Parent Proxy Without Local

Dependence." The former (Parent Proxy) includes calibrations for all items. This is the default calibration sample. If you aren't sure which calibration sample to use, utilize this one. The Parent Proxy Without Local Dependence does not include calibrations for some items. The items without calibrations are enemy items. That is, a dyad or triad of items was identified in which there are psychometric reasons to only administer one of those items to a given respondent. For example, item Pf1mobil1r and Pf1mobil3r are enemy items. A participant should only see one of these items in a CAT.

The PROMIS Pediatric v2.0 Upper Extremity instrument also includes two calibration samples – "Pediatric" and "Pediatric Without Local Dependence." The Pediatric Without Local Dependence calibration sample is selected as the default. It does not include calibrations for enemy items.

SCORES

For most PROMIS instruments, a score of 50 is the average for the United States general population with a standard deviation of 10 because calibration testing was performed on a large sample of the general population. You can read more about the calibration and centering samples on HealthMeasures.net (http://www.healthmeasures.net/score-and-interpret/interpret-scores/promis). The T-score is provided with an error term (Standard Error or SE). The Standard Error is a statistical measure of variance and represents the "margin of error" for the T-score.

Important: *A higher PROMIS T-score represents more of the concept being measured.* For positively-worded concepts like Physical Function, Mobility, and Upper Extremity function, a T-score of 60 is one SD better than average. By comparison, a Physical Function T-score of 40 is one SD worse than average.

STATISTICAL CHARACTERISTICS

There are four key features of the score for Physical Function:

- **Reliability**: The degree to which a measure is free of error. It can be estimated by the internal consistency of the responses to the measure, or by correlating total scores on the measure from two time points when there has been no true change in what is being measured (for z-scores, reliability = $1 - SE^2$).
- **Precision**: The consistency of the estimated score (reciprocal of error variance).
- **Information:** The precision of an item or multiple items at different levels of the underlying continuum (for z-scores, information = $1/SE^2$).
- **Standard Error (SE):** The possible range of the actual final score based upon the scaled T-score. For example, with a T-score of 52 and a SE of 2, the 95% confidence interval around the actual final score ranges from 48.1 to 55.9 (T-score $\pm$ (1.96*SE) = 52 $\pm$ 3.9 = 48.1 to 55.9).

The final score is represented by the T-score, a standardized score with a mean of 50 and a standard deviation (SD) of 10.

In Figure 2 (v1.0 adult 10a and v1.0 adult 20a short form), the two dotted horizontal lines each represent a degree of internal consistency reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 20-item form. Figure 2 also tells us where on the scale the forms are most informative

based upon the T-score: the 20-item form is more informative than the 10-item form, and the 20-item form offers sufficient reliability over a wider range of T-scores than the 10-item form.

Figure 3 (v1.0 Adult 4a, 6a & 8a short forms) also tells us where on the scale the form is most informative based upon the T-score: the 8-item form is more informative than the 6-item form, which is more informative than the 4-item form. See additional test information figures for Pediatric instruments in Appendix 1.

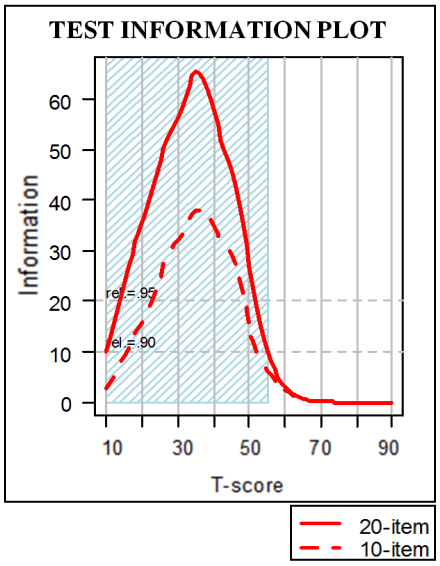

Figure 3

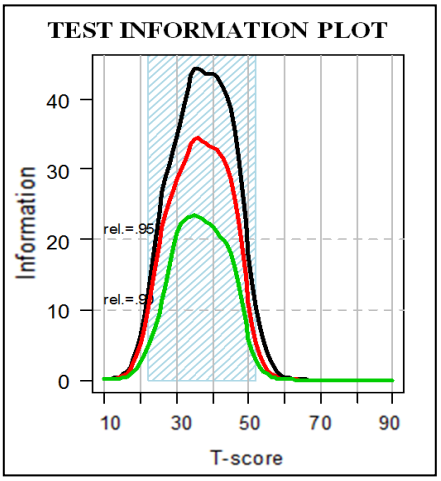

Figure 2

PREVIEW OF SAMPLE ITEM

Figure 4 is an excerpt from the paper version of the v2.0 adult item bank. This is the paper version format used for all Physical Function instruments. It is important to note, CAT is not available for paper administration.

		Without any difficulty	With a little difficulty	With some difficulty	With much difficulty	Unable to do
PFA19r1	Are you able to run or jog for two miles (3 km)?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFA20	Are you able to cut your food using eating utensils?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1

Figure 4

FREQUENTLY ASKED QUESTIONS (FAQ)

Q: I am interested in learning more. Where can I do that?
Review the HealthMeasures website at www.healthmeasures.net.

Q: Are these instruments available in other languages?
Yes! Look at the HealthMeasures website (http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/available-translations) for current information on PROMIS translations.

Q: Can I make my own short form?
Yes, custom short forms can be made by selecting any items from the item bank. This can be scored using the Scoring Service (https://www.assessmentcenter.net/ac_scoringservice).

APPENDIX 1– ADDITIONAL FIGURES

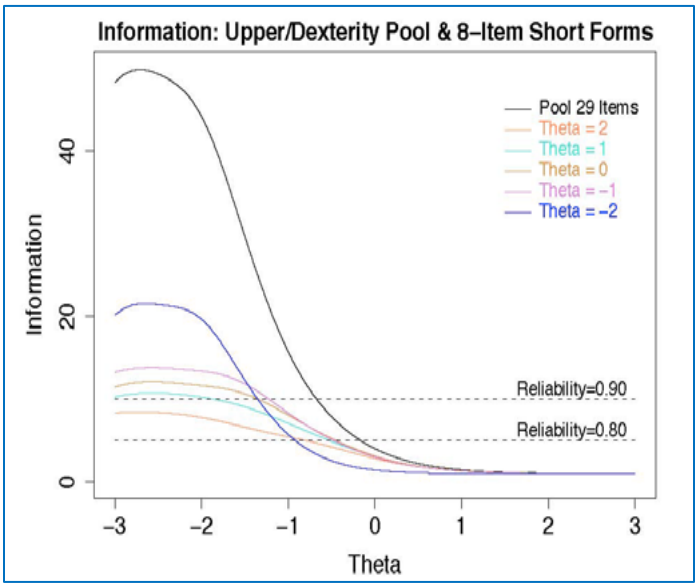

Figure 5 Pediatric Test Information Upper Extremity

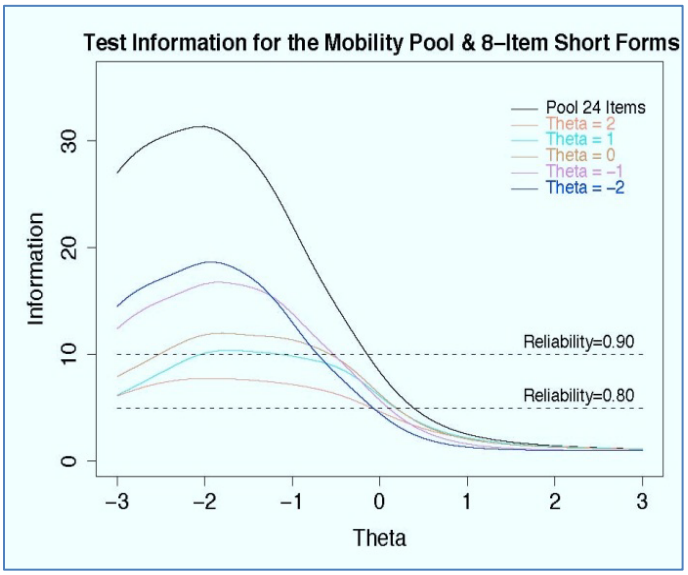

Figure 6 Pediatric Test Information Mobility