

ILLNESS BURDEN SCORING MANUAL

A brief guide to scoring the PROMIS® Illness Burden instruments:

ADULT
PROMIS Bank v1.0 - Illness Burden PROMIS Short Form v1.0 - Illness Burden 6a

COMPARING SCORES ACROSS VERSIONS

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

There is only one version (v1.0) of the Illness Burden instruments.

SCORING THE INSTRUMENT

Short Forms: PROMIS instruments are scored using item-level calibrations. This means that the most accurate way to score a PROMIS instrument is to use the HealthMeasures Scoring Service (https://www.assessmentcenter.net/ac_scoringervice) or a data collection tool that automatically calculates scores (e.g., REDCap auto-score). This method of scoring uses responses to each item for each participant. We refer to this as “response pattern scoring.” Because response pattern scoring is more accurate than the use of raw score/scale score look up table included in this manual, it is preferred. Response pattern scoring is especially useful when there is missing data (i.e., a respondent skipped an item), different groups of participants responded to different items, or you have created a new short form using a subset of questions from a PROMIS item bank.

To use the scoring table in this manual, calculate a summed score. Each question usually has five response options ranging in value from one to five. To find the total raw summed score for a short form with all questions answered, sum the values of the response to each question. For example, for the v1.0 6-item form, the lowest possible raw score is 6; the highest possible raw score is 30 (see short form scoring table in Appendix 1). **All questions must be answered in order to produce a valid score using the scoring tables.** If a participant has skipped a question, use the HealthMeasures Scoring Service (https://www.assessmentcenter.net/ac_scoringervice) to generate a final score.

With the total raw summed score for a measure, locate the score conversion table in Appendix 1 and use this table to translate the total raw score into a T-score for each participant. The T-score rescales the raw score into a standardized score with a mean of 50 and a standard deviation (SD) of 10. Therefore, a person with a T-score of 40 is one SD below the mean. For the Illness Burden v1.0 short form 6a, a raw score of 10 converts to a T-score of 43.4 with a standard error (SE) of 3.0 (see scoring table for the short form in Appendix 1).

CAT: A minimum number of items (i.e., 4) must be answered in order to receive a score for the Illness Burden 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. The CAT will continue until either the standard error drops below a specified level (3.0 on the T-score metric), or the participant has answered the maximum number of questions (8), whichever occurs first.

The Illness Burden item bank was developed with a small exploratory sample. Consequently, the CAT administration algorithm is a little different from other PROMIS measures. This means that the PROMIS Illness Burden CAT tends to be a little longer than other PROMIS CATs.

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. However, for Illness Burden a score of 50 is the average for a sample of individuals with type 2 diabetes mellitus. 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 negatively-worded concepts like Illness Burden, a T-score of 60 is one SD worse than average. By comparison, an Illness Burden T-score of 40 is one SD better than average.

Standard Error (SE): A PROMIS score includes a T-score and a standard error (SE). The standard error is a measure of the variability for a given T-score across hypothetical repeated measurements. The standard error can be used to construct confidence intervals around a T-score. A 95% confidence interval is common. A 95% confidence interval means there is a 95% probability that the true T-score is within this range. The formula for a 95% confidence interval is $(T\text{-score} \pm (1.96 * SE))$. For example, if $T=52$ and $SE=2$, the lower boundary of the confidence interval is $(52 - (1.96 * 2)) = 48$ and the upper boundary is $(52 + (1.96 * 2)) = 56$.

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: How do I handle multiple responses when administering a short form on paper?

Guidelines on how to deal with multiple responses have been established. Resolution depends on the responses noted by the research participant.

- If two or more responses are marked by the respondent, and they are next to one another, then a data entry specialist will be responsible for randomly selecting one of them to be entered and will write down on the form which answer was selected. Note: To randomly select one of two responses, the data entry specialist will flip a coin (heads - higher number will be entered; tails – lower number will be entered).



To randomly select one of three (or more) responses, a table of random numbers should be used with a statistician's assistance.

- If two or more responses are marked, and they are NOT all next to one another, the response will be considered missing.

Q: What is the minimum change on a PROMIS instrument that represents a clinically meaningful difference?

To learn more about research on the meaning of a change in scores, we suggest conducting a literature review to identify the most current information. The HealthMeasures website (<http://www.healthmeasures.net/score-and-interpret/interpret-scores/promis>) has additional information on interpreting scores.

APPENDIX 1-SCORING TABLE

Short Form v1.0 Illness Burden 6a		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
6	32.6	5.1
7	37.4	3.7
8	39.7	3.5
9	41.7	3.2
10	43.4	3.0
11	44.9	2.9
12	46.4	2.9
13	47.7	2.9
14	49.1	2.9
15	50.4	2.9
16	51.8	2.9
17	53.1	2.9
18	54.4	2.9
19	55.7	2.9
20	57.0	2.9
21	58.3	2.9
22	59.5	2.9
23	60.8	2.9
24	62.1	2.9
25	63.5	2.9
26	64.9	3.0
27	66.4	3.1
28	68.3	3.4
29	70.1	3.5
30	73.9	4.5
*SE = Standard Error on T-score metric		