

4-29-2019

Variation and Change Over Time in PROMIS-29 Survey Results Among Primary Care Patients With Type 2 Diabetes

Juell Homco
Kristin Rodriguez
David R. Bardach
Elizabeth A. Hahn
Suzanne Morton
Daren Anderson
David Kendrick
Sarah Hudson Scholle

Follow this and additional works at: <https://aurora.org/jpcrr>



Part of the [Community Health and Preventive Medicine Commons](#), [Endocrine System Diseases Commons](#), [Endocrinology, Diabetes, and Metabolism Commons](#), [Health Services Research Commons](#), and the [Primary Care Commons](#)

Recommended Citation

Homco J, Rodriguez K, Bardach DR, Hahn EA, Morton S, Anderson D, Kendrick D, Scholle SH. Variation and change over time in PROMIS-29 survey results among primary care patients with type 2 diabetes. *J Patient Cent Res Rev*. 2019;6:135-47. doi: [10.17294/2330-0698.1694](https://doi.org/10.17294/2330-0698.1694)

Published quarterly by Midwest-based health system Advocate Aurora Health and indexed in PubMed Central, the Journal of Patient-Centered Research and Reviews (JPCRR) is an open access, peer-reviewed medical journal focused on disseminating scholarly works devoted to improving patient-centered care practices, health outcomes, and the patient experience.

Variation and Change Over Time in PROMIS-29 Survey Results Among Primary Care Patients With Type 2 Diabetes

Juell Homco, MPH,¹ Kristin Rodriguez, MPH,¹ David R. Bardach, PhD,^{2*} Elizabeth A. Hahn, MA,³ Suzanne Morton, MPH, MBA,^{2†} Daren Anderson, MD,⁴ David Kendrick, MD, MPH,¹ Sarah Hudson Scholle, DrPH²

¹University of Oklahoma-University of Tulsa School of Community Medicine, Tulsa, OK; ²National Committee for Quality Assurance, Washington, DC; ³Northwestern University Feinberg School of Medicine, Chicago, IL;

⁴Community Health Center, Inc., Middletown, CT

Purpose	We sought to describe results of patient-reported outcome measures implemented among primary care patients with diabetes and explore factors associated with changes in scores over time.
Methods	Two organizations serving diverse patient populations collected the PROMIS-29 survey at baseline and 3-month follow-up for patients with type 2 diabetes. Bayesian regression analysis was used to examine the relationship between patient characteristics and changes in PROMIS-29 scores. Exploratory analyses assessed relationships between goal-setting and changes in scores.
Results	The study population reported substantially more problems with physical functioning (mean: 42.5 at Site 1 and 38.9 at Site 2) and pain interference (mean: 58.0 at Site 1 and 61.1 at Site 2) compared to the general population (mean: 50; standard deviation: 10). At least 33% of patients had a clinically meaningful change (ie, at least half the standard deviation, or 5 points) in each PROMIS domain. For pain interference, 55% had no change, 22% improved by 5 or more points, and 23% worsened by 5 or more points. Bayesian regression analyses suggest that chronic conditions, insurance status, and Hispanic ethnicity are likely associated with decreased functioning over time. Exploratory analyses found that setting a mental health goal did not appear to be associated with improvement for anxiety or depression.
Conclusions	Use of patient-reported outcome measures in routine clinical care identified areas of functional limitations among people with diabetes. However, changes in participants' PROMIS-29 scores over time were minimal. Research is needed to understand patterns of change in global and domain-specific functioning, particularly among racial/ethnic minorities. (<i>J Patient Cent Res Rev.</i> 2019;6:135-147.)
Keywords	patient-reported outcomes; health-related quality of life; outcome measurement; health care planning

Performance measures are commonly used in value-based payment models as the basis of payments, incentives, and public quality reporting.^{1,2} Many of these performance measures

use clinical data routinely collected at the point of care.¹ However, there is increasing demand to add assessment of patient-reported outcomes in routine clinical care and corresponding performance metrics based on patient-reported outcome measures (PROMs) to health care quality assessment.^{1,2}

*Now with Cedar Gate Technologies (Greenwich, CT).

†Now with American Society for Nutrition (Rockville, MD).

Correspondence: Sarah Hudson Scholle, DrPH, National Committee for Quality Assurance, 1100 13th Street NW, Washington, DC 20005 (Scholle@ncqa.org)

Currently, only a few existing performance metrics use a PROM. For example, the Centers for Medicare & Medicaid Services (CMS) offers incentives for use of performance measures assessing functional outcomes in the Comprehensive Care for Joint Replacement payment model.³ Also, measures assessing the

monitoring and remission of depression use the Patient Health Questionnaire, a 9-item tool that lets patients report their symptoms of depression.⁴ While not widely adopted as quality measures in the U.S. health care system, other countries have used PROM-based performance metrics to evaluate outcomes.^{5,6}

Incorporating PROM data collection into routine clinical care has been shown to improve patient-clinician relationships, advance shared decision-making, and improve patient health outcomes.⁷⁻¹⁰ In addition, the use of PROMs has been shown to empower patients to be more involved in their care.^{10,11} Research trials have found that PROMs can enhance goal-setting for diabetes, depression, and other conditions.^{12,13}

Although patient-reported outcome tools are increasingly being used in routine clinical care, there are challenges to implementation and interpretation for patients with complex chronic conditions.^{10,14-16} Information regarding how to interpret results, the amount of change to expect over time, and how to maintain or improve functioning or forestall slow decline is inadequate.¹⁷ There is growing evidence of the validity of Patient-Reported Outcomes Measurement Information System (PROMIS[®]) tools over time with some clinical populations,^{18,19} yet data for patients with diabetes is lacking. The PROMIS tools were developed with National Institutes of Health funding and include self-reported measures of global, physical, mental, and social health. For example, the PROMIS-29 v2.0 profile is a collection of 4-item short forms assessing anxiety, depression, fatigue, pain interference (the impact of pain on one's ability to perform daily activities), physical function, sleep disturbance, and ability to participate in social roles and activities, along with a single item on pain intensity.^{20,21} PROMIS tools have high reliability, have been demonstrated to be valid and responsive to change in research and clinical populations, are in the public domain, and have existing infrastructure to support their use.²²⁻²⁵ Despite this evidence, more information is needed on expected responses and how to interpret changes in score over time in specific PROMIS domains for specific patient groups.

This project sought to fill a gap in existing literature by describing the results of PROMs, specifically a PROMIS-29 tool, implemented among primary

care patients with type 2 diabetes as part of a larger study assessing the feasibility and usefulness of PROMs in routine clinical care. This project also aimed to understand changes in PROM scores over time by describing their associations with patient demographics, clinical characteristics, and patient goals for this patient population.

METHODS

This study was conducted in two primary care organizations with diverse structures and patient populations (a Federally Qualified Health Center and an academic health center), previously described in detail.¹⁶ This project was reviewed and approved by Chesapeake Research Review, LLC (Columbia, MD) and the institutional review boards at the participating sites. A waiver of consent was requested and approved for the provision of a limited data set that included necessary clinical data elements, PROM data, and goal-setting information collected during routine care.

Study Population

The study population included patients with type 2 diabetes who had a hemoglobin A1c of $\geq 6.5\%$ and a primary care visit with a participating provider within the study period (July 1, 2015–May 31, 2016). Each site had a goal of 200 participants. The target recruitment number was based on a previous study that showed this sample was large enough to detect a meaningful effect size for change in PROMIS-29 score.^{20,21} Both sites modified their existing workflows to include PROM data collection and use of the results for goal-setting conversations with patients in routine clinical care. Details on the study population and implementation strategy have been previously described.¹⁶

Data and Measures

We used the PROMIS-29 v2.0 profile.^{20,21} Norms for PROMIS-29 were derived for the U.S. general population in previous work.²² Specifically, PROMIS questionnaires were administered using a sampling plan to ensure that each item was administered to at least 900 respondents from the U.S. general population and 500 respondents with known chronic medical conditions.²² A subsample was derived to represent the U.S. general population in terms of gender, age, race/ethnicity, and education.²⁶ A T-score of 50 is the mean of this normative subsample.

Medical assistants and care managers gave patients a paper version of the PROMIS-29 at baseline and at the 3-month follow-up visit. Patients generally completed the form themselves, but staff read the form to patients if requested and when the form was completed by telephone. The English version was available and offered to patients at both sites; Site 1 also offered the Spanish version.

Patient age, gender, race/ethnicity, preferred language, and insurance characteristics were obtained from electronic health records (EHRs) at the participating sites. Chronic conditions were identified from patient diagnoses in the EHR and coded using 27 categories of chronic conditions defined by the CMS Chronic Conditions Data Warehouse.²⁷ Results for blood pressure, blood glucose levels (hemoglobin A1c), and body mass index were obtained from EHR data and coded based on existing quality measures or definitions.^{28,29}

Goals

At Site 1, a primary care clinician (or nurse), nutritionist, or behavioral health provider conducted goal-setting conversations at a patient visit or over the telephone after the patient visit. At Site 2, care managers conducted goal-setting conversations and participated in the primary care visit. Patient goals were documented in patient's words (including multiple goals when applicable), along with the action steps (also open-ended text) and the domains within which the patient wished to focus, using specific fields to identify the PROMIS-29 domains.

We categorized each patient's goals and other frequently mentioned topics into 1 or more PROMIS-29 domains. Two National Committee for Quality Assurance staff members independently categorized patient goals until agreement was reached, grouped the goal domains into 3 larger categories (health care goals, health goals, and life goals), and obtained consensus on the categorization from the project's executive committee.

Analysis

PROMIS-29 scoring guidelines were followed — namely, use of T-scores, which are standardized with a mean of 50 and standard deviation (SD) of 10.²² The

means and distributions on each scale at baseline and follow-up assessments were calculated. We assumed follow-up scores were missing at random^{30,31} for this calculation, meaning the missing values were more likely a product of the other observed patient-level variables and not the missing values themselves. To examine mean scores for baseline and follow-up, we estimated linear mixed models with random subject effects to account for the correlation among repeated observations^{32,33} using a hierarchical approach that uses all available data. We chose not to explicitly impute data because the conclusions would then be sensitive to assumptions made in the imputing process and because these models do not require complete cases.

Least squares means, standard errors, and 95% confidence intervals were estimated from the models. We also present the magnitude of change from baseline to follow-up in 5 categories using 0.5 SD (ie, 5 points) as a clinically meaningful difference.^{34,35} In the absence of specific studies evaluating responsiveness to change in a particular population, an effect size in the range of 0.33 to 0.50 generally corresponds to a clinically important difference in PROM outcomes.^{35,36}

We analyzed the association of patient characteristics with changes in PROMIS-29 scores using Bayesian regression.³⁶ Bayesian methods allowed us to answer two questions about the relationship of a patient characteristic with PROMIS-29 scores: 1) How strong is the probable association of a patient factor with outcomes? and 2) How likely is that association? Bayesian methods also allowed us to take advantage of all nonmissing data; thus, we were able to use data for patients with missing follow-up assessments without having to impute data.

Sampling of posterior distributions was done with RStanArm, as accessed through R 3.3.3; all other analyses were done in R 3.3.3 directly.^{37,38} A weakly informed prior (Cauchy distribution centered at 0 and a constant term of 2.5) was used as a conservative but practical estimate for all regression coefficients.³⁹ The sampler ran 8 chains; each chain collected 5000 samples, although the first 2500 samples in each chain were discarded to allow for burn-in. Convergence and possible autocorrelation were assessed using RStanArm's Rhat and n_eff estimates.

The regression models all took the following form:

$$\text{Change in T-score} \sim N(\beta_0 + \beta_1(\text{Initial T-score}) + \beta_2(\text{Days from Baseline}) + \beta_3(\text{Obese}) + \beta_4(\text{High Blood Pressure}) + \beta_5(\text{Hemoglobin A1c} > 9.0) + \beta_6(2\text{--}4 \text{ Chronic Conditions}) + \beta_7(+\text{Chronic Conditions}) + \beta_8(\text{Medicare/Medicaid/Dual Eligible}) + \beta_9(\text{Uninsured}) + \beta_{10}(\text{Age } 65+) + \beta_{11}(\text{Female}) + \beta_{12}(-\text{Hispanic Black}) + \beta_{13}(\text{Hispanic, English-Preferred}) + \beta_{14}(\text{Hispanic, Spanish-Preferred}) + \beta_{15}(\text{Site}), \sigma^2)$$

For modeling change in T-scores, we excluded observations when the baseline score on a PROMIS-29 domain was either the highest or the lowest score possible (therefore allowing either no room for improvement or for worsening). Because of this approach, sample sizes ranged from 178 to 246 patients per domain.

In practice, this exclusion method could reduce the models' ability to specify a narrow range of plausible effect sizes; however, we did not observe that it introduced noteworthy selection bias. We evaluated the impact of patient demographic and clinical characteristics across PROMIS-29 domains by examining the distribution of their effect estimates with site included as a covariate. In addition, Markov chain Monte Carlo (MCMC) simulations were used to inspect each independent variable's parameter estimate for changes in individual PROMIS-29 T-scores. Our conclusions were consistent across the 6 PROMIS-29 domains, and therefore, we combined results into a single plot to make a global statement across domains (after reverse-coding, where appropriate, so that higher scores were always better). The parameter estimates from the MCMC simulations were merged together, resulting in a single vector for each of the common independent variables. Each vector then represented an independent variable's influence over changes in the PROMIS-29 domains based on the global posterior distribution.

Descriptive statistics on goal content were calculated, and exploratory analyses to examine how the content related to PROMIS-29 anxiety and depression scores were conducted. We chose these domains because it was most clear when goals were related to these domains. We did not test statistical significance due to small sample sizes for goal types. We compared the

content of goals among people whose PROMIS-29 scores indicated problems with depression or anxiety (T-score of ≥ 55 on either domain, excluding people without results for the relevant domains) versus those who did not.

Finally, we plotted the change in PROMIS-29 scores among all patients (highlighting those who set goals for mental health and excluding patients missing relevant PROMIS-29 results at baseline or follow-up, thus assuming data were missing completely at random).

RESULTS

PROMIS-29 Scores and Change Over Time

The demographic and clinical characteristics of patients who completed baseline PROMIS-29 data collection are provided, by site, in Table 1. At baseline, our study population ($n=490$) reported more problems with physical functioning (average of 43 at Site 1 and 39 at Site 2, with higher scores indicating better functioning) and pain interference (average of 58 at Site 1 and 61 at Site 2, with lower scores indicating better functioning) than the general population; both indicated functioning about 1 SD poorer than the general population average of 50 (Table 2). For each domain, at least 35% of patients had a clinically meaningful change between the baseline and follow-up assessment (at least 5 points/0.5 SD) in either direction (Table 3). For example, for pain interference, 55% had no change, 22% improved by 5 or more points, and 23% worsened by 5 or more points.

Patient Characteristics Affecting Change in PROMIS-29 Scores

Figure 1 shows the results of Bayesian analyses to examine the relationship of patient demographic and clinical characteristics with change scores on PROMIS-29 domains combined, considering all other covariates in the model. Results suggest that having more chronic conditions, public insurance, and Hispanic ethnicity is likely to be associated with decreased functioning over time. Patients with 2–4 or 5 or more chronic conditions were more than 80% likely to have decreased functioning at follow-up assessment; this is shown by the percentage of the distribution to the left of 0. The most likely decrease is about 5 points, identified in the chart by the vertical bar at the median of the distribution. Having public insurance (Medicare,

Table 1. Demographic and Clinical Characteristics of Patients Participating in Baseline PROM Data Collection, by Site

Characteristic	Site 1		Site 2	
	n	%	n	%
Total patient population	205	100	285	100
Age category				
18–54 years	92	44.9	102	35.8
55–64 years	74	36.1	101	35.4
65+ years	39	19.0	82	28.8
Female	114	55.6	179	62.8
Race/ethnicity				
Non-Hispanic white	51	24.9	152	53.3
Non-Hispanic black	26	12.7	80	28.1
Hispanic	113	55.1	13	4.6
Other/Multiple races	10	4.8	13	4.6
Primary insurance				
Commercial	45	22.0	64	22.5
Medicare/Medicaid/dual eligible	147	71.7	206	72.3
Other	0	0	1	0
Uninsured	13	6.3	14	4.9
Spanish preferred language	72	35.1	3	1.1
Chronic conditions				
Diabetes only	16	7.8	11	3.9
2–4 conditions	126	61.5	145	50.9
5+ conditions	63	30.7	129	45.3
Hemoglobin A1c > 9.0	43	21.0	77	27.0
High BP reading (systolic ≥ 140 mmHg or diastolic ≥ 90 mmHg)	66	32.2	96	33.7
BMI ≥ 30.0 (obese)	144	70.2	197	69.1

BMI, body mass index; BP, blood pressure; PROM, patient-reported outcome measure.

Medicaid, or both) is also probably associated with decreased functioning (most of the distribution is to the left of 0), but the effect is smaller and the range of probable effects narrower. Hispanic patients (with either Spanish- or English-language preference) are also at least 80% likely to have decreased functioning.

Relationship of Goals to Change in PROMIS-29 Scores

The proportion of patients who set a goal varied by site, with 40% at Site 1 versus 90% at Site 2; at Site 2, care managers more consistently worked with patients to identify a goal (Table 4). Nearly all patients who set a goal focused on a health-related topic; however,

mental health (28%) and exercise goals (25%) were most common at Site 1, whereas diet (34%), weight loss (26%), and exercise goals (25%) were most common at Site 2. In many cases, goals addressed multiple topics.

To explore the association of goals with PROMIS-29 scores, we focused on mental health goals and their potential association with scores on the PROMIS-29 anxiety and depression domains. Patients who scored high on anxiety or depression tended to be more likely to set a goal (eg, 42% vs 37% at Site 1) and to identify a mental health goal (32% vs 23% at Site 1) (Table 4).

Table 2. Mean Scores for Baseline and Follow-Up PROMIS-29 Domain, by Site*

Domain	Data Collection	Site 1			Site 2		
		Mean	95% Lower CI	95% Upper CI	Mean	95% Lower CI	95% Upper CI
HIGHER IS BETTER							
Physical function	Baseline	42.5	41.1	43.9	38.9	37.7	40.0
	Follow-up	41.5	40.0	43.0	38.7	37.4	39.9
Ability to participate in social roles and activities	Baseline	49.9	48.4	51.5	48.6	47.3	49.9
	Follow-up	49.8	48.1	51.5	49.2	47.7	50.6
LOWER IS BETTER							
Anxiety	Baseline	53.2	51.7	54.6	52.7	51.5	53.9
	Follow-up	55.9	54.3	57.5	54.3	53.0	55.7
Depression	Baseline	54.4	52.8	55.9	53.1	51.9	54.4
	Follow-up	54.7	53.0	56.3	52.2	50.8	53.5
Fatigue	Baseline	51.2	49.6	52.7	56.8	55.5	58.1
	Follow-up	52.9	51.3	54.6	55.6	54.1	57.1
Sleep disturbance	Baseline	54.5	53.2	55.8	54.5	53.2	55.8
	Follow-up	53.5	52.1	55.0	53.1	51.7	54.5
Pain interference	Baseline	58.0	56.4	59.5	61.1	59.9	62.4
	Follow-up	58.5	56.9	60.2	60.2	58.9	61.6

*Accounting for missing follow-up data.

Table 3. Change in PROMIS-29 Scores Between Baseline and Follow-Up

	Total Non-missing	Improved at Least 1 SD (Change ≥ -10)		Improved 0.5 SD (Change of -5 to Less Than -10)		Minimal Change (Between -5 and 5)		Worse by 0.5 SD (Change of 5 to Less Than 10)		Worse by 1 SD (Change ≥ 10)	
		n	%	n	%	n	%	n	%	n	%
HIGHER IS BETTER											
Physical function	356	30	8.4	42	11.8	230	64.6	32	9.0	22	6.2
Ability to participate in social roles and activities	332	42	12.7	50	15.1	147	44.3	45	13.6	48	14.5
LOWER IS BETTER											
Anxiety	352	41	11.6	32	9.1	162	46.0	46	13.1	71	20.2
Depression	349	41	11.7	47	13.5	190	54.4	38	10.9	33	9.5
Fatigue	348	55	15.8	49	14.1	142	40.8	42	12.1	60	17.2
Sleep disturbance	350	61	17.4	40	11.4	166	47.4	50	14.3	33	9.4
Pain interference	344	39	11.3	36	10.5	190	55.2	42	12.2	37	10.8

*n's vary because not all patients could be scored on all domains at both time periods.

SD, standard deviation.

Setting a mental health goal did not appear to be associated with reductions in anxiety or depression (Figures 2 and 3). For example, as shown for anxiety in Figure 2, there is not a clear pattern of patients who set a mental health goal (shown in black triangles)

compared to other patients (gray dots). Most patients had minimal change (<5 points) in their anxiety score — these are the dots and triangles in between the dashed lines. More patients reported their anxiety worsened (shown by dots above the upper solid line) than

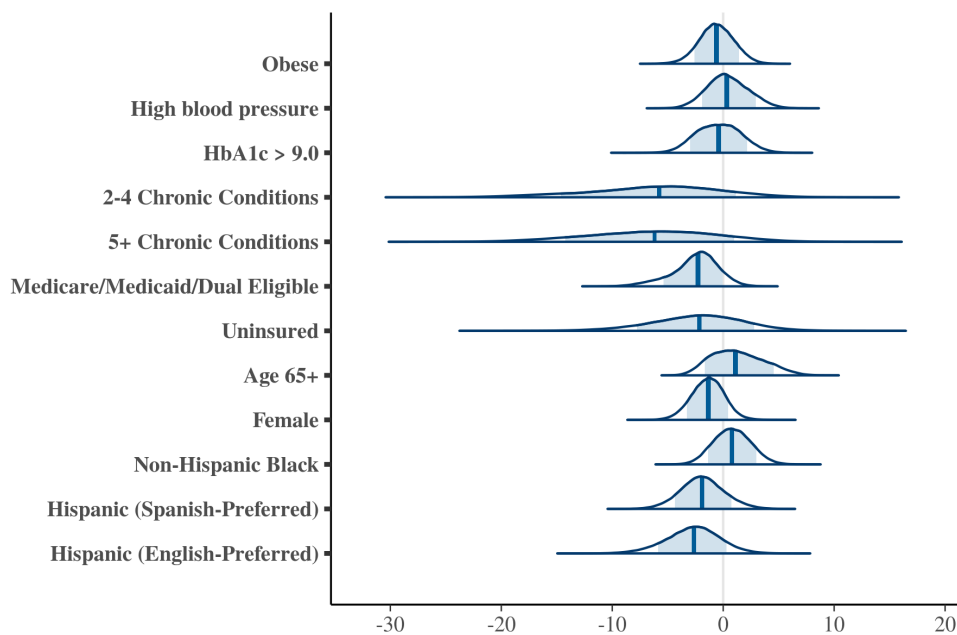


Figure 1. Relationship of patient factors to changes in PROMIS-29 scores. Bayesian regression models were used for this analysis. The chart summarizes distribution of effects of the independent variable on the change in the PROMIS score across all domains after reverse-coding, where appropriate, so higher is always better. The table shows the probability that each patient characteristic is associated with worse-than-average changes in PROMIS results over time, and the central 80% region of estimated effect sizes.

Variable	Probability of Harmful Effect	Central 80% Interval	
Obese	65.0%	-2.57	1.43
High blood pressure	42.0%	-1.92	2.92
HbA1c > 9.0	56.1%	-2.98	2.17
2–4 chronic conditions	86.2%	-14.63	1.14
5+ chronic conditions	87.2%	-14.24	0.99
Medicare/Medicaid/dual eligible	90.0%	-5.40	-0.05
Uninsured	69.7%	-7.77	2.81
Age 65+ years	32.0%	-1.64	4.53
Female gender	81.6%	-3.30	0.46
Non-Hispanic black race	32.1%	-1.34	2.98
Hispanic (Spanish-preferred)	84.2%	-4.37	0.72
Hispanic (English-preferred)	87.0%	-5.88	0.29

improved (shown by dots below the lower solid line). Of note, some patients who reported no problems at baseline had scores above the population average score of 50 at follow-up, and patients with scores above the mean at baseline often reported the lowest score at follow-up (dots/triangles on x and y axes).

For depression (Figure 3), most patients had changes less than 10 points (in either direction), but patterns related to lowest scores also are evident.

DISCUSSION

This work is among the first to document PROMIS-29 results over time in a population with type 2 diabetes identified in a routine clinical care setting.^{40,41} Our study

population reported more problems with physical functioning and pain interference than the general population. Overall, our results are consistent with a previous study that found little change in patient-reported outcomes over time for most primary care patients.¹⁷ The finding that most patients had minimal change is not a surprise since the goal-setting intervention was not intensive. Our exploratory analyses did not suggest a link between setting mental health goals and improvement in PROMIS-29 scores at follow-up. The clustering of responses at the lowest point of the scores (indicating no problems) warrants greater investigation. Also, while we were not surprised to find that chronic conditions and public insurance were associated with a greater probability

Table 4. Content of Goals for Patients With Functional Limitations Based on PROMIS-29 Anxiety and Depression Scores, by Site

Type of Goal	Site 1				Site 2			
	Anxiety or Depression		No Anxiety or Depression		Anxiety or Depression		No Anxiety or Depression	
	n	%	n	%	n	%	n	%
Number of patients	118	100	82	100	163	100	121	100
At least one goal	50	42.4	30	36.6	150	92.0	106	87.6
Health	46	92.0	29	96.7	132	88.0	92	86.8
Diet	9	18.0	4	13.3	52	34.7	35	33.0
Lose weight	6	12.0	2	6.7	33	22.0	33	31.1
Exercise	10	20.0	10	33.3	29	19.3	33	31.1
Reduce pain	9	18.0	2	6.7	34	22.7	17	16.0
Physical function	7	14.0	1	3.3	12	8.0	7	6.6
Mental health*	16	32.0	7	23.3	15	10.0	1	0.9
Stop smoking	2	4.0	0	0.0	9	6.0	6	5.7
Sleep	6	12.0	6	20.0	7	4.7	2	1.9
Fatigue	2	4.0	0	0.0	3	2.0	1	0.9
Maintain weight	0	0.0	0	0.0	1	0.7	0	0.0
Meditation	0	0.0	2	6.7	1	0.7	0	0.0
Health care	8	16.0	2	6.7	67	44.7	31	29.2
Reduce glucose level	0	0.0	1	3.3	27	18.0	19	17.9
Other clinical issue	2	4.0	0	0.0	29	19.3	13	12.3
Medication-related	5	10.0	2	6.7	23	15.3	8	7.5
Diabetes supplies	0	0.0	0	0.0	5	3.3	2	1.9
Other goal	1	2.0	0	0.0	3	2.0	0	0.0
Transportation	2	4.0	0	0.0	2	1.3	0	0.0
Life	7	14.0	1	3.3	20	13.3	5	4.7
Social	5	10.0	1	3.3	9	6.0	4	3.8
Home-related	1	2.0	0	0.0	8	5.3	3	2.8
Financial help	0	0.0	0	0.0	5	3.3	0	0.0
Hobby	1	2.0	0	0.0	2	1.3	0	0.0
Spiritual health	0	0.0	0	0.0	1	0.7	0	0.0

*Functional limitation was defined as T-score of 55 or greater on the PROMIS anxiety or depression scale (where mean is 50 and standard deviation is 10). Excludes patients who did not have a score on either of the anxiety and depression scales at baseline.

of decreased functioning over time, the probable relationship with Hispanic ethnicity is of interest.

Our study offers support and direction for future research to understand the patterns of change and improvement in functional status among patients in primary care and the extent of their sensitivity to goal-setting and care-planning interventions. In particular, research should

explore the variability of PROMIS-29 scores and evaluate whether responses clustered at the lowest point of the score (indicating no problems) occurs in other settings and with varying implementation strategies. Also, research should explore language and ethnicity to consider whether the finding of a relationship between Hispanic ethnicity and decreased functioning over time is borne out in larger, more robust studies.

Anxiety T-score

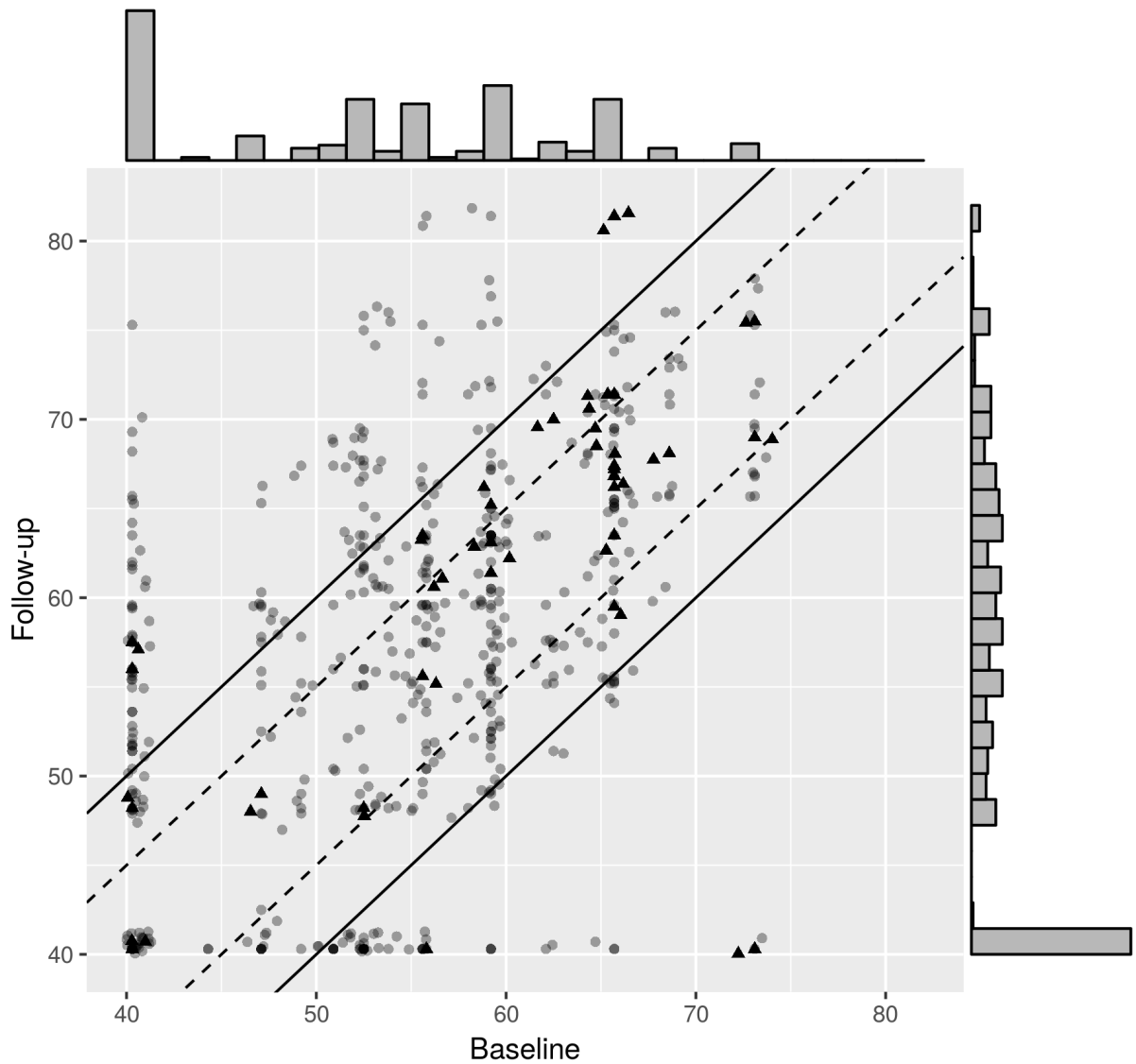


Figure 2. Change in PROMIS-29 anxiety score among patients who set/did not set a mental health goal at baseline (both sites combined). Baseline anxiety score is on the horizontal axis (higher scores indicate worse anxiety). Follow-up score is on the vertical axis. Each patient is a dot. Black triangle dots represent patients who set a mental health goal, and gray circular dots represent patients who did not set a mental health goal. Only patients with scores at both baseline and follow-up are represented. Patients who fall between the dashed lines had minimal change between baseline and follow-up (<5 points in either direction). Patients between the solid lines had a change that was <1 SD (<10 points). Patients above the top solid line had an anxiety score worse by >1 SD at follow-up. Patients below the bottom solid line had an anxiety score that improved by >1 SD at follow-up. Columns at the right and top of the chart show the histograms (number of patients with that score) for each point on the axis.

Results show that most patients are in the middle, with minimal change, but more patients reported that their anxiety was worse. The dots clustered at the origin represent patients who had the lowest score possible at both baseline and follow-up. A number of patients who reported no problems at baseline had scores above the population average (ie, 50) at follow-up; patients with scores above the population average at baseline often reported the lowest score at follow-up.

Depression T-score

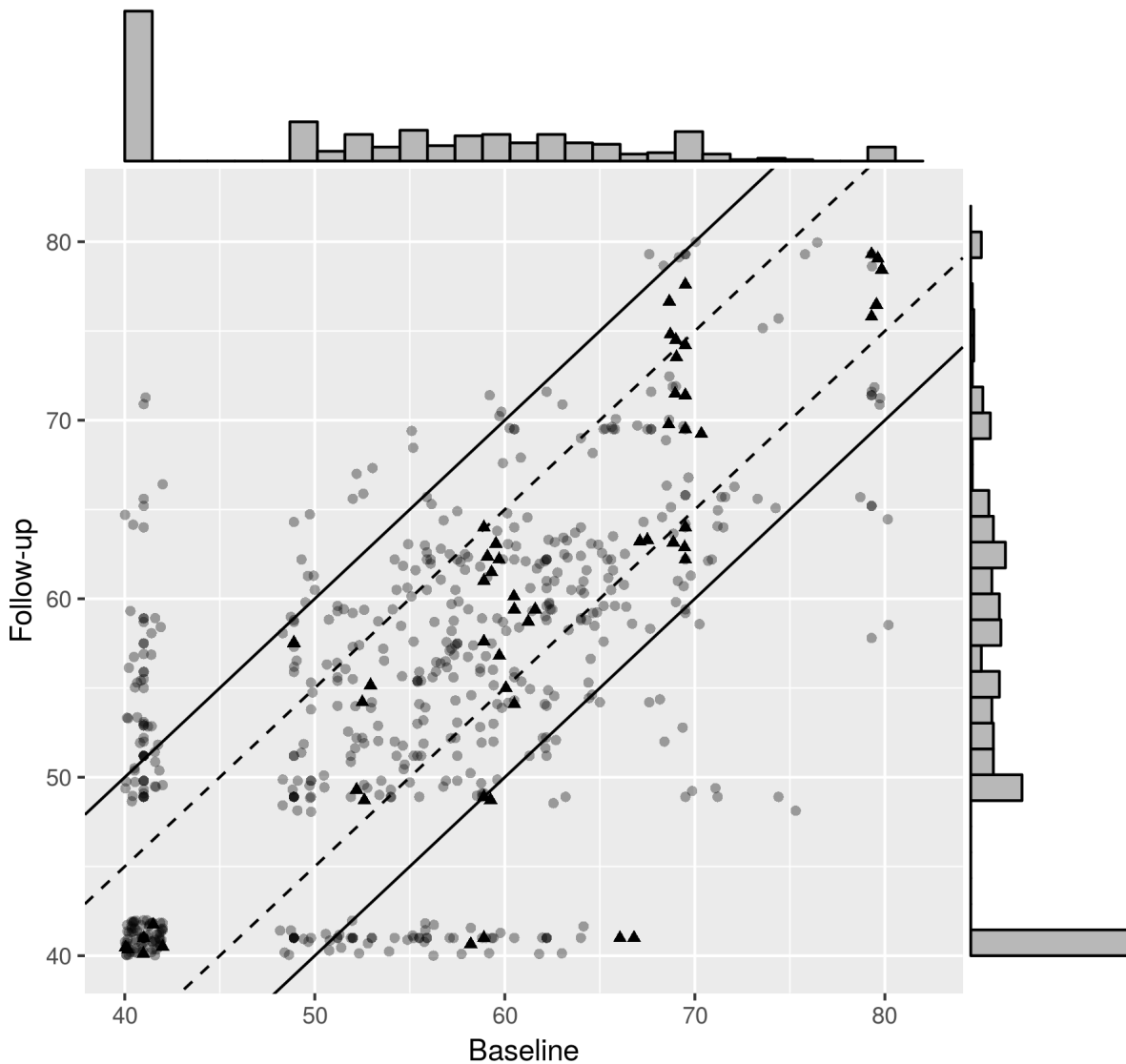


Figure 3. Change in PROMIS-29 depression score among patients who set/did not set a mental health goal at baseline (both sites combined). Baseline depression score is on the horizontal axis. Higher scores indicate worse depression. Follow-up score is on the vertical axis. Each patient is a dot. Black triangle dots represent patients who set a mental health goal, and gray circular dots represent patients who did not set a mental health goal. Only patients with scores at both baseline and follow-up are represented. Patients who fall between the dashed lines had minimal change between baseline and follow-up (<5 points in either direction). Patients between the solid lines had a change that was <1 SD (<10 points). Patients above the top solid line had a depression score worse by >1 SD at follow-up. Patients below the bottom solid line had a depression score that improved by >1 SD at follow-up. Columns at the right and top of the chart show the histograms (number of patients with that score) for each point on the axis.

Results show that most patients are in the middle, with no change. The dots clustered at the origin represent patients who had the lowest score possible at both baseline and follow-up. A number of patients who reported no problems at baseline had scores above the average 50 at follow-up; patients with scores above the mean at baseline often reported the lowest score at follow-up.

Given the site-specific differences in goal-setting implementation and the anticipated challenges with incorporating PROMs and patient goals into practice EHRs and overall workflow, future research should explore effective methods of PROM implementation. For example, human-factor research methods⁴² could play an important role in improving use of patient-reported outcome data in clinical care by helping identify mental models that guide patient engagement and clinical decision-making, patient cultural and language needs, and workflows for data collection and follow-up. To support this work, we believe that it will be important first to gather sufficient data to provide patients, clinicians, and other members of the care team information on PROMIS-29 scores and how they can be used to improve clinical management and treatment as well as monitor outcomes.

Limitations

The generalizability of findings from this pilot feasibility study is limited because it was conducted at only two sites, included only patients with type 2 diabetes, and used a single PROM tool (PROMIS-29). Although changes in PROMIS-29 scores were assessed over a 3-month period, longer longitudinal studies are needed to truly determine clinical relevance. Sites did not use electronic data collection for PROM results or for tracking patients over time due to the significant costs and resources needed to modify practice EHRs. This may have impacted provider engagement in the goal-setting process.

The two sites varied in both their patient population and workflow/care team, so our ability to tease out these factors is limited. The timing of follow-up and documentation of missing data varied due to differences in staffing, workflows, and other competing demands in these busy primary care settings. Despite these limitations, the study is important for capturing data on PROMs in settings serving diverse populations, and the participating sites — a Federally Qualified Health Center and an academic clinic — are likely generalizable to comparable settings.

CONCLUSIONS

The use of patient-reported outcome measures in routine clinical care identified areas of functional limitations among people with diabetes, particularly in pain interference and physical functioning. However,

changes in PROMIS-29 scores over time without targeted intervention were limited. As increased emphasis is placed on using PROMs in routine clinical care and performance measurement, additional research is needed to understand patterns of change in global and domain-specific functioning, particularly among racial/ethnic minorities, as well as the best ways to use these data in care planning. Future studies should consider the cost of implementation along with potential changes in PROM scores to conduct cost-benefit analyses.

Patient-Friendly Recap

- Patient-reported outcomes include any report of health status that comes directly from the patient, for example, an answer to a survey asking how much pain they feel.
- PROMIS-29 is a survey designed to obtain and measure patient-reported quality of life.
- The authors tested which factors contributed to changes in outcomes (ie, better or worse quality of life) for patients with type 2 diabetes visiting a primary care clinic.
- Use of PROMIS-29 identified that, over 3 months, decreased physical function was associated with number of chronic conditions, insurance status, and Hispanic ethnicity.

Acknowledgments

Mary Wilkins and Susan McBrien of University of Oklahoma—University of Tulsa School of Community Medicine; Daisy Burgos and Thomas McCalla of Community Health Center, Inc.; and Brad Roberts of the National Committee for Quality Assurance.

Author Contributions

Study design: Homco, Rodriguez, Hahn, Anderson, Kendrick, Scholle. Data acquisition or analysis: Homco, Rodriguez, Morton, Bardach, Hahn, Morton, Scholle. Manuscript drafting: Homco, Rodriguez, Bardach, Hahn, Morton, Scholle. Critical revision: all authors.

Conflicts of Interest

None.

Funding Sources

This research was funded through a Patient-Centered Outcomes Research Institute (PCORI) Award (ME-1402-10788). The statements in this report are solely the responsibility of the authors and do not necessarily represent the views of PCORI, its Board of Governors, or Methodology Committee.

References

1. Center for Clinical Standards and Quality, Centers for Medicare & Medicaid Services. CMS quality measure development plan: supporting the transition to the quality payment program. 2017 annual report. Issued 2017 Jun 2. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2017-CMS-MDP-Annual-Report.pdf>. Accessed September 10, 2017.
2. Hopkins D, Huff JE. PBGH action brief on patient-reported outcomes, April 2015. http://www.pbgh.org/storage/documents/AB-Patient_Reported_Outcomes.pdf. Accessed September 10, 2017.
3. Centers for Medicare & Medicaid Services. Overview of CJR quality measures, composite quality score, and pay-for-performance methodology. <https://innovation.cms.gov/Files/x/cjr-qualsup.pdf>. Accessed September 10, 2017.
4. MN Community Measurement. Data collection guide: direct data submission. Depression care measures 2013 (02/01/2012 to 01/31/2013 dates of service). Published 2013. http://mncm.org/wp-content/uploads/2013/04/Depression_Care_Measures_DDS_2013_Final_11_29_2012.pdf. Accessed September 10, 2017.
5. Secondary Care Analysis, NHS Digital. Finalised patient reported outcome measures (PROMs) in England: April 2014 to March 2015. Published 2016 Aug 11. <http://content.digital.nhs.uk/catalogue/PUB21189/final-proms-eng-apr14-mar15-fin-report.pdf>. Accessed September 10, 2017.
6. Mäkelä KT, Matilainen M, Pulkkinen P, et al. Countrywise results of total hip replacement. An analysis of 438,733 hips based on the Nordic Arthroplasty Register Association database. *Acta Orthop*. 2014;85:107-16. [CrossRef](#)
7. Basch E. Patient-reported outcomes – harnessing patients’ voices to improve clinical care. *N Engl J Med*. 2017;376:105-8. [CrossRef](#)
8. Griggs CL, Schneider JC, Kazis LE, Ryan CM. Patient-reported outcome measures: a stethoscope for the patient history. *Ann Surg*. 2017;265:1066-7. [CrossRef](#)
9. Cella D, Hahn EA, Jensen SE, et al. *Patient-Reported Outcomes in Performance Measurement*. Research Triangle Park, NC: RTI Press, 2015.
10. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res*. 2012;21:1305-14. [CrossRef](#)
11. Greenhalgh J, Pawson R, Wright J, et al. Functionality and feedback: a protocol for a realist synthesis of the collation, interpretation and utilisation of PROMs data to improve patient care. *BMJ Open*. 2014;4:e005601. [CrossRef](#)
12. Rodriguez HP, Anastario MP, Frankel RM, et al. Can teaching agenda-setting skills to physicians improve clinical interaction quality? A controlled intervention. *BMC Med Educ*. 2008;8(1):3. [CrossRef](#)
13. Rao JK, Anderson LA, Inui TS, Frankel RM. Communication interventions make a difference in conversations between physicians and patients: a systematic review of the evidence. *Med Care*. 2007;45:340-9. [CrossRef](#)
14. Basch E, Torda P, Adams K. Standards for patient-reported outcome-based performance measures. *JAMA*. 2013;310:139-40. [CrossRef](#)
15. Kwan BM, Sills MR, Graham D, et al. Stakeholder engagement in a patient-reported outcomes (PRO) measure implementation: a report from the SAFTINet Practice-based Research Network (PBRN). *J Am Board Fam Med*. 2016;29:102-15. [CrossRef](#)
16. Scholle SH, Morton S, Homco J, et al. Implementation of the PROMIS-29 in routine care for people with diabetes: challenges and opportunities. *J Ambul Care Manage*. 2018;41:274-87. [CrossRef](#)
17. Peters M, Crocker H, Jenkinson C, Doll H, Fitzpatrick R. The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey. *BMJ Open*. 2014;4(2):e003968. [CrossRef](#)
18. Cook KF, Jensen SE, Schalet BD, et al. PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. *J Clin Epidemiol*. 2016;73:89-102. [CrossRef](#)
19. Jensen RE, Moynour CM, Potosky AL, et al. Responsiveness of 8 Patient-Reported Outcomes Measurement Information System (PROMIS) measures in a large, community-based cancer study cohort. *Cancer*. 2017;123:327-35. [CrossRef](#)
20. Rojas Smith L, Ashok M, Morss Dy S, Wines RC, Teixeira-Poit S. *Contextual Frameworks for Research on the Implementation of Complex System Interventions*. Report No.: 14-EHC014-EF. Rockville, MD: Agency for Healthcare Research and Quality, 2014.
21. HealthMeasures. PROMIS® (Patient-Reported Outcomes Measurement Information System). <http://www.healthmeasures.net/explore-measurement-systems/promis>. Accessed April 14, 2018.
22. 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. *J Clin Epidemiol*. 2010;63:1179-94. [CrossRef](#)
23. DeWalt DA, Rothrock N, Yount S, Stone AA; PROMIS Cooperative Group. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care*. 2007;45(5 Suppl 1):S12-S21. [CrossRef](#)
24. Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;45(5 Suppl 1):S22-S31. [CrossRef](#)
25. Carle AC, Cella D, Cai L, et al. Advancing PROMIS’s methodology: results of the Third Patient-Reported Outcomes Measurement Information System (PROMIS®) Psychometric Summit. *Expert Rev Pharmacoecon Outcomes Res*. 2011;11:677-84. [CrossRef](#)
26. Liu H, Cella D, Gershon R, et al. Representativeness of the Patient-Reported Outcomes Measurement Information System internet panel. *J Clin Epidemiol*. 2010;63:1169-78. [CrossRef](#)
27. Chronic Conditions Data Warehouse, Centers for Medicare & Medicaid Services. Condition categories. <https://www.cdwdata.org/web/guest/condition-categories>. Accessed March 27, 2017.
28. National Committee for Quality Assurance. HEDIS measures and technical resources. <http://www.ncqa.org/hedis-quality-measurement/hedis-measures/hedis-2015>. Accessed September 10, 2017.

29. Centers for Disease Control and Prevention. Defining adult overweight and obesity. Last reviewed 2017 Apr 11. <https://www.cdc.gov/obesity/adult/defining.html>. Accessed March 15, 2019.
30. Troxel AB, Fairclough DL, Curran D, Hahn EA. Statistical analysis of quality of life with missing data in cancer clinical trials. *Stat Med*. 1998;17:653-66. [CrossRef](#)
31. Little RJ, Rubin DB. *Statistical Analysis With Missing Data, Second Edition*. Hoboken, NJ: John Wiley & Sons, Inc., 2002. [CrossRef](#)
32. Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data*. New York, NY: Springer, 2000. [CrossRef](#)
33. Hedeker D, Gibbons RD. *Longitudinal Data Analysis*. Hoboken, NJ: John Wiley & Sons, Inc., 2006.
34. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes*. 2006;4:70. [CrossRef](#)
35. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582-92. [CrossRef](#)
36. Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Methods in health service research. An introduction to bayesian methods in health technology assessment. *BMJ*. 1999;319:508-12. [CrossRef](#)
37. RStanArm. Bayesian applied regression modeling via Stan. <https://mc-stan.org/users/interfaces/rstanarm>. Accessed September 10, 2017.
38. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2018. Available online at <https://www.R-project.org/>.
39. Gelman A, Jakulin A, Pittau MG, Su YS. A weakly informative default prior distribution for logistic and other regression models. *Ann Appl Stat*. 2008;2:1360-83. [CrossRef](#)
40. Bermúdez-Millán A, Pérez-Escamilla R, Segura-Pérez S, et al. Psychological distress mediates the association between food insecurity and suboptimal sleep quality in Latinos with type 2 diabetes mellitus. *J Nutr*. 2016;146:2051-7. [CrossRef](#)
41. Hahn EA, Burns JL, Jacobs EA, et al. Health literacy and patient-reported outcomes: a cross-sectional study of underserved English- and Spanish-speaking patients with type 2 diabetes. *J Health Commun*. 2015;20 Suppl 2:4-15. [CrossRef](#)
42. Gibbons MC, Lowry SZ, Patterson ES. Applying human factors principles to mitigate usability issues related to embedded assumptions in health information technology design. *JMIR Hum Factors*. 2014;1(1):e3. [CrossRef](#)

© 2019 Aurora Health Care, Inc.