Applied Research Quality Life (2018) 13:89–103
DOI 10.1007/s11482-017-9507-x

Clinical Validation of PROMIS Global Short Form in Pregnancy

Lisbet S. Lundsberg[1] · **Eleanor B. Schwarz**[2] · **Nicole A. Vilardo**[1] · **Kimberly A. Yonkers**[3] · **Aileen M. Gariepy**[1]

Received: 17 March 2016 / Accepted: 17 January 2017 / Published online: 2 February 2017

Abstract Health-related quality of life (HRQoL) is an important patient-reported outcome, yet research regarding HRQoL during pregnancy is limited. We examined HRQoL during pregnancy using the Patient Reported Outcomes Measurement Information System (PROMIS) Global Short Form (GSF) and validated the GSF compared to legacy HRQoL measures. We evaluated HRQoL among 161 women seeking pregnancy care in urban clinic settings. Participants completed measures of HRQoL, social support, antenatal depression, and utility. Descriptive statistics and correlation coefficients were calculated. Participants averaged 27(±6.6) years and were culturally diverse: 42% self-identified as Hispanic, 37% Black, non-Hispanic, 14% White, non-Hispanic and 7% multiracial or other. Mean estimated gestational age was 9 (±4.6) weeks. PROMIS GSF Physical T-scores were significantly correlated with SF-12 Physical Component Score (PCS) and Mental Component Score (MCS) HRQoL measures (correlation coefficient=0.40 and 0.49, p-value<0.0001, respectively), the Modified Kendler Social Support Index (MKSSI) (correlation coefficient=0.42, p-value<0.0001), and the Visual Analog Scale (VAS) measure of utility (correlation coefficient=0.19, p-value =0.04). GSF Mental T-scores were associated with SF-12 MCS and PCS (correlation coefficient=0.66, p-value<0.0001, and 0.26, p-value<0.01, respectively), MKSSI (correlation coefficient=0.50, p-value<0.0001), and VAS (correlation coefficient=0.29, p-value<0.01). GSF Physical and Mental scores were inversely associated with the Edinburgh Postpartum Depression Scale (EPDS), correlation

✉ Lisbet S. Lundsberg
lisbet.lundsberg@yale.edu

[1] Department of Obstetrics, Gynecology and Reproductive Sciences, Yale School of Medicine, 310 Cedar Street, New Haven, CT 06520, USA

[2] University of California, Davis, 4150 V Street, Sacramento, CA 95817, USA

[3] Department of Psychiatry, Yale School of Medicine, New Haven, CT 06520, USA

coefficient= −0.62 and − 0.71, respectively (p-value<0.0001). GSF-derived utility measures demonstrate significant correlation with SF-12 PCS and MCS, MKSSI, EPDS, and VAS. Overall, PROMIS GSF domains demonstrate correlation with legacy HRQoL measures as well as validated measures of social support, depression, and utility among a diverse cohort of pregnant women.

Keywords Health related quality of life (HRQoL) · Utility · Pregnancy · PROMIS

Introduction

Health related quality of life (HRQoL) is a standardized patient-reported outcome that quantifies the impact of a disease or health state on an individual (Torrance 1987; Ubel et al. 2003). However, HRQoL of reproductive health outcomes, including pregnancy, is understudied (Goldhaber-Fiebert and Brandeau 2015; Myers 2015). HRQoL during pregnancy is difficult to assess because of its transient nature, variability during the course of gestation, and the complexity of the maternal/fetal dyad. Regardless, millions of women around the world spend considerable amounts of time being pregnant, including six million pregnancies among women of child-bearing age (15–44 years) annually in the United States (Curtin et al. 2013). Pregnancy is a unique health state that may significantly impact women's quality of life, in both positive and negative ways, for years to come. Legacy HRQoL measures have often been developed to measure the impact of chronic illness or disease on quality of life, which are distinct from pregnancy as a health state. Capturing women's HRQoL is essential to appropriately measure overall experiences related to reproductive health (Mogos et al. 2013). There is an increasing interest in evaluating the HRQoL of specific reproductive health states, including decision and cost-effective interventions affecting reproduction (Goldhaber-Fiebert and Brandeau 2015; Myers 2015; Washington et al. 2015). Recent literature evaluating the cost-effectiveness of interventions affecting fertility and child-bearing reveal a lack of transparency and inconsistency in reporting and measuring health effects and quality adjusted life years (Goldhaber-Fiebert and Brandeau 2015). The lack of a systematic framework for evaluating pregnancy-related outcomes in cost-effectiveness analyses reinforces the need for standardized and validated instruments measuring HRQoL in this population.

HRQoL is traditionally measured with carefully designed questionnaires or semi-structured interviews such as the SF-12 Health Form (Ware et al. 1996). These instruments are general and may not adequately capture the unique experiences of pregnant women (Mogos et al. 2013). For example, traditional quality of life (QoL) measures that assess patients' ability to perform physical daily activities, such as playing golf, may not be applicable to pregnancy (Ware et al. 1996). Yet legacy HRQoL instruments may not incorporate specific parameters of health, including emotional states that are relevant to pregnancy, including depression. The Edinburgh Postpartum Depression Scale (EPDS) (Cox et al. 1987) has been used to identify depression among both antepartum and postpartum women (Biratu and Haile 2015; Cox et al. 1996; Gibson et al. 2009; Senturk et al. 2011; Heron et al. 2004; Rochat et al. 2006). Similarly, the Kendler Social Support Index has been modified and validated among a pregnant population (Spoozak et al. 2009).

Robust measurements of HRQoL during pregnancy are also important because they can be converted into a composite index (Revicki et al. 2009), the European Quality of Life Index Score Five Dimensions, or EurolQoL (EQ-5D), to derive utility measurements of the value or desirability of a health state (Weinstein et al. 2009) integral to cost effectiveness analyses.

Given the lack of data regarding quality of life during pregnancy, robust measurement tools for determining HRQoL during pregnancy are needed. The Patient Reported Outcomes Measurement Information System (PROMIS®) has developed a multi-dimensional instrument that may better capture the physical and emotional aspects of HRQoL during pregnancy. Developed by the National Institutes of Health to reliably assess patients' self-reported health (Cella et al. 2007, 2010; DeWalt et al. 2007; Hays et al. 2009; Klem et al. 2009; Teresi et al. 2009) the PROMIS database includes global and health state specific HRQoL questionnaires and computerized adaptive tests that identify physical and mental health domains affected by a certain disease or health state (Cella et al. 2007, 2010; DeWalt et al. 2007). Furthermore, PROMIS measures have standardized norms so that HRQoL can be compared among different health states. The EQ-5D measure of QoL (EuroQol 1990) can also be calculated from PROMIS measures to provide a summary health utility index for the health state (Revicki et al. 2009).

While evidence from PROMIS' validation phase suggests it is appropriate to use for many health outcomes (Cella et al. 2010), PROMIS instruments have not been well tested in pregnancy. Therefore, we sought to clinically validate the 10-item PROMIS Global Short Form (GSF), which includes measurement of both physical and mental health domains, among a cohort of pregnant women. We hypothesized that PROMIS-GSF outcomes would be significantly correlated with scores from the SF-12 and other common measures of HRQoL. Specifically, our objective was to demonstrate that PROMIS-GSF Physical and Mental Health domains would demonstrate convergent validity with traditional measures of HRQoL, in addition to other selected measures of social support, depression and utility.

Methods

Study Setting and Design

We evaluated antenatal HRQol among a cohort of women presenting for pregnancy testing or abortion care at clinics in New Haven, CT. The primary study was designed to evaluate measures of pregnancy context, quality of life, social support, and depression during pregnancy and association with maternal and neonatal outcomes. Women either testing positive for pregnancy or scheduled for abortion services during the period of June 2014 to June 2015 were approached regarding participation in the study. Eligibility criteria included: English or Spanish speaking, gestational age <24 completed weeks, age 15–44 years, and study enrollment within 1 week of their pregnancy test or pre-operative clinic visit. Among 361 women presenting at clinics for abortion services or having a positive pregnancy test, 269 women were determined to be eligible

and approached regarding study participation; of those, 196 were interested in participating and 32 were unable to stay for enrollment or lost to follow up within 1 week of contact. Thus, 164 were enrolled. One individual provided consent but did not complete the baseline assessment and two individuals initially testing positive for pregnancy were subsequently determined not to be pregnant. Therefore, a total of 161 participants are included in this analysis.

Participants completed a self-administered questionnaire at time of enrollment. Demographic and maternal characteristics were assessed, including age, race, ethnicity, education, employment, relationship status, as well as smoking, alcohol and drug use within the last 3 months. Information regarding chronic medical conditions (e.g. asthma, diabetes, thyroid disease) and previous diagnosis of depression or anxiety was collected. Reproductive history was ascertained, including number of previous births, miscarriage, abortion, and age at first pregnancy. Gestational age was based on reported last menstrual period or clinician's estimated gestational age at time of enrollment.

Measures of HRQoL

The PROMIS Global Short Form (GSF) consists of ten questions (Hays et al. 2009). Nine questions utilize a 5-point response scale and the pain intensity assessment is administered using an 11-point scale. Separate domain scores for Global Short Form Physical (GSF Physical) and Mental (GSF Mental) Health components are calculated based on eight questions, using four questions for each domain. Converting response values and summing scores based on preset algorithms yields PROMIS raw scores for each domain, from which T-score metrics are calculated (PROMIS 2010). Individual T-scores for both PROMIS GSF Physical and GSF Mental domains are reported in this analysis. A T-score value of 50 represents the mean value within the U.S. population, with a standard deviation (SD) of 10. Higher PROMIS domain T-scores represent higher quality of life.

The 12-item Short Form Health Survey (SF-12) (Ware et al. 1996) was also administered and scored independently by Optum®/QualityMetric (Lincoln, RI) for standardization and to ensure validity. The SF-12 generates 8 health domain scales which are then used to compute the two summary component measures: the Mental Component Score (MCS) and the Physical Component Score (PCS). Selected participants with missing information on paired questions could not be scored according to preset algorithms. In other cases a standardized missing data estimation approach was applied to response values, yielding 96% complete data overall. SF-12v1 and SF-12v2 forms were utilized (94% and 6%, respectively), both generating summary physical and mental component scores (PCS and MCS, respectively). Higher SF-12 component score values represent higher quality of life.

Depression Score

Edinburgh Postpartum Depression Scale (EPDS) is based on 10 questions with four potential response categories each, ranging in scored values from 0 to 3, developed originally for use among postpartum women and validated for use among antepartum

women (Cox et al. 1987; Matthey et al. 2006). Two of the ten EPDS questions were reversed scored. Total EPDS was calculated by summing scored values for each participant across all questions. To identify women with a positive depression screen, we used a cutoff of greater or equal to 13, consistent with previous studies among antenatal populations (Biratu and Haile 2015; Senturk et al. 2011; Heron et al. 2004; Rochat et al. 2006).

Measures of Social Support

The Modified Kendler Social Support Index (MKSSI) has been used as a validated measure of social support during pregnancy (Spoozak et al. 2009). Questions assessed social support from the respondent's spouse or partner, mother, father, siblings, other relatives, and friends, forming the subscale components for the MKSSI and encompassing a total of 25 questions. Individual questions measured social support based on how much the individual "listens to your problems or worries", "understands the way you think and feel", and "goes out of their way to help you", using a 5-point Likert scale ranging from 1 to 5. For the subscales of social support from mother, father, frequency of contact was also included, ranging in value from 0 (never or if deceased) to 5 (at least once a day). For friends, a measure of how many close confidants was assessed, ranging from 0 to a maximum of 5. Individual component values for the MKSSI subscales were averaged to result in a subscale score. Overall MKSSI scores were calculated by averaging subscale component values for social support.

Measures of Utility

The Visual Analogue Scale (VAS) has been used for the valuation of health states (Robinson et al. 2001) including pregnancy (Creinin 2000; Schwarz et al. 2008). Study participants were instructed to place an 'X' on a 10-cm line to indicate how they felt after learning they were pregnant, with the line ranging from 0 ('As if I was dying') to 10 ('In perfect health') (Schwarz et al. 2008). Measurements were recorded and recalibrated to a 0–1.0 scale.

An algorithm utilizing questions of the PROMIS GSF provided measurement of utility metric approximating the European Quality of Life (EuroQoL) 5 Dimensions (EQ-5D) (Revicki et al. 2009). Eight of the ten PROMIS GSF question responses were incorporated to calculate the composite index score as a measure of utility.

Statistical Analysis

Descriptive and bivariate analyses were performed including mean values and standard deviations, medians and interquartile range, and overall range values for measure of HRQoL, social support, depression and utility. Spearman's rank correlation coefficients were calculated for each PROMIS GSF domain and SF-12 legacy measure of HRQoL, as well as MKSSI, EPDS, PROMIS GSF-derived utility and VAS. Cronbach alpha was calculated for components of PROMIS GSF and GSF domains, GSF-derived utility, MKSSI, and EDPS to assess internal consistency of these measures within the cohort. Statistical analysis was performed using SAS 9.4 (SAS Institute, Cary, NC).

The study protocol was reviewed and approved by the Yale University Human Research Protection Program as well as participating clinical sites. Study participants provided written informed consent prior to enrollment.

Results

Sample Characteristics

Average age of participants among our sample was 27(±6.6) years and gestational age at study enrollment was 9(±4.6) weeks (Table 1). Twenty-six percent completed the study in Spanish and the remainder in English (74%). Forty-two percent self-identified as Hispanic, 37% Black, non-Hispanic, 14% White, non-Hispanic and 7% multiracial. Most participants had at least a high school degree (43%) or some college or college degree (39%), were single or living with partner but not married (73%), and had a previous birth (75%). Approximately half (48%) were employed either part-time or full-time, 20% identified as homemaker, and 32% were unemployed. Over one-third (36%) reported having a previous miscarriage and 44% reported having a previous abortion. Current chronic medical conditions (e.g. asthma, hypertension, diabetes) were reported by 21%, with a similar proportion reporting previous diagnosis of depression or anxiety (21% and 21%, respectively). In the 3 months prior to enrollment, 32% reported tobacco use and 24% reported smoking marijuana, while 52% reported alcohol consumption during that period. When asked about the period just before becoming pregnant, 73% of the cohort indicated they either did not intend to get pregnant or their intentions kept changing; 27% intended to get pregnant. At enrollment, 61% were planning to parent, 26% were planning to have an abortion, 1% planned for adoption, and 12% did not know what their plans were for their pregnancy.

Descriptive Analysis of HRQoL MKSSI, EPDS, and Measures of Utility

Mean PROMIS GSF Physical T-scores among the study sample was 48.5(±8.4), ranging from 29.6 to 67.7 and mean GSF Mental T-score was 50.1(±9.9), ranging from 28.4 to 67.6 (Table 2). SF-12 Physical Composite Score mean value was 49.4(±8.1) and ranged from 24.6 to 65.3, while SF-12 Mental Composite Score mean value was 48.2(±12.0), ranging from 5.9 to 66.3. Total MKSSI score averaged 3.2(±0.8), ranging from 0.9 to 4.9; MKSSI subscale average values ranged from 2.48(±1.63) for father support to 3.76(±1.40) for mother support. Overall EPDS scores ranged from 0 to 27, with average value of 7.5(±6.4); 36 participants (22.4% of the sample) scored ≥13 on the EPDS, indicating a positive screening for depression. PROMIS-derived utility demonstrated an average index score of 0.71(±0.10), ranging from 0.45 to 0.88. VAS measurements were reported for 80% of the sample; mean VAS value was 0.76(±0.27), ranging from 0.03 to 1.00. Cronbach alpha correlation coefficients demonstrated good overall reliability for the PROMIS GSF HRQoL measures including the 10-item PROMIS GSF, eight PROMIS physical and mental domain components (correlation coefficient = 0.89 and 0.86, respectively); the 4-item GSF Physical and Mental Health domains demonstrated reliability correlation coefficients of 0.63 and 0.85 respectively. The PROMIS GSF-derived utility components also

Table 1 Participant demographics and characteristics at enrollment ($N = 161$)[a]

Characteristic	N (%)
Age in years	
Mean (SD)	27.2 (6.6)
Median (range)	26 (16–44)
Gestational age at enrollment	
Mean (SD)	9.1 (4.6)
Median (range)	7.4 (3.7–23)
Language study conducted in	
English	119 (73.9)
Spanish	42 (26.1)
Race/ethnicity	
Hispanic	67 (42.4)
Black, non-Hispanic	58 (36.7)
White, non-Hispanic	22 (13.9)
Multiracial, other	11 (7.0)
Education (years)	
Less than 12	29 (18.1)
12 or high school equivalent	68 (42.5)
Some college/college degree	63 (39.4)
Employment	
Full-time	39 (24.4)
Part-time	38 (23.7)
Unemployed	51 (31.9)
Homemaker	32 (20.0)
Relationship status	
Single or living with partner, not married	117 (73.1)
Married	24 (15.0)
Separated, divorced, widowed, other	19 (11.9)
Current chronic medical problems (e.g. including asthma, hypertension, diabetes, multiple sclerosis)	34 (21.1)
Ever diagnosed with depression	34 (21.1)
Ever diagnosed with anxiety	33 (20.5)
Smoking/use of tobacco products past 3 months	
None	109 (67.7)
Once/twice or monthly	15 (9.3)
Weekly or daily	37 (23.0)
Drinking alcohol past 3 months	
None	77 (47.8)
Once/twice or monthly	67 (41.6)
Weekly or daily	17 (10.6)
Marijuana use in past 3 months	
None	123 (76.4)
Once/twice or monthly	24 (14.9)

Table 1 (continued)

Characteristic	N (%)
Weekly or daily	14 (8.7)
Parity	
0	40 (25.2)
1	56 (35.2)
2+	63 (39.6)
Previous miscarriage	54 (36.0)
Previous abortion	67 (43.8)
Intention before becoming pregnant	
Intended	44 (27.3)
Not intended, intentions changing	117 (72.7)
What are you planning to do	
Planning to parent	97 (60.6)
Planning for adoption	2 (1.3)
Planning for abortion	41 (25.6)
Don't know	20 (12.5)

[a] Totals may not add to 161 due to missing information

demonstrated internal consistency with correlation coefficient = 0.85, as well as internal consistency for the 10-item EPDS, correlation coefficient = 0.90. Reliability correlation coefficient of the six subscale components of the MKSSI was 0.62.

Table 2 Characteristics of study cohort using measures of quality of life, $N = 161$[a]

Measure	N	Mean (SD)	Median (IQR)	Range
PROMIS Global Short Form				
PROMIS Global Physical Health Score	153	48.5 (8.4)	47.7 (42.3–54.1)	29.6–67.7
PROMIS Global Mental Health Score	155	50.1 (9.9)	50.8 (43.5–59.0)	28.4–67.6
12-item Short Form Health Survey (SF12)				
Physical Composite Score (PCS)	154	49.4 (8.1)	51.1 (43.5–55.4)	24.6–65.3
Mental Composite Score (MCS)	154	48.2 (12.1)	51.3 (41.2–57.1)	5.9–66.3
Modified Kendler Social Support Index (MKSSI, total)	145	3.17 (0.83)	3.19 (2.57–3.81)	0.95–4.95
MKSSI Subscales				
Social support - Partner	161	2.81 (1.73)	3.00 (1.00–4.67)	1.0–5.0
Social support - Father	159	2.48 (1.63)	2.75 (0.75–4.00)	0.75–5.0
Social support - Mother	159	3.76 (1.40)	4.25 (3.25–4.75)	0.75–5.0
Social support - Siblings	160	3.74 (1.30)	4.00 (3.00–5.00)	1.0–5.0
Social support - Others	159	3.22 (1.40)	3.33 (2.00–4.33)	1.0–5.0
Social support - Friends	149	3.25 (1.20)	3.25 (2.25–4.25)	0.75–5.0
Edinburgh Postpartum Depression Scale (EPDS)	161	7.5 (6.4)	6.0 (2.0–11.0)	0–27.0
PROMIS GSF-derived utility index	150	0.71 (0.10)	0.71 (0.63–0.79)	0.45–0.88
Visual Analog Scale (VAS)	128	0.76 (0.27)	0.88 (0.53–1.00)	0.03–1.00

[a] Totals may not add to 161 due to missing information

Correlation of HRQoL, MKSSI, EPDS, and Measures of Utility

PROMIS GSF Physical and Mental domain scores were highly correlated with corresponding HRQoL SF-12 physical and mental component scores, as well as measures of social support (MKSSI), depression (EPDS) and utility (GSF-derived and VAS) (Table 3). PROMIS-GSF Physical scores were significantly correlated with the SF-12 PCS (correlation coefficient = 0.40, $p < 0.0001$). Similarly, PROMIS-GSF Mental scores were significantly correlated with SF-12 MCS (correlation coefficient = 0.66, $p < 0.0001$). Positive correlation was also shown between PROMIS-GSF Physical and Mental domains with total MKSSI (correlation coefficient = 0.42 and 0.50, $p < 0.0001$, respectively), and measures of GSF-derived utility (correlation coefficients = 0.92 and 0.84, $p < 0.0001$, respectively). PROMIS-GSF Physical and Mental domains were also significantly correlated with VAS (correlation coefficient = 0.19 and 0.29, respectively), however the magnitude of the correlation was reduced compared to GSF-derived utility. Both GSF Physical and Mental scores were significantly correlated with all MKSSI subscales, with correlation coefficient ranging from 0.19 ($p = 0.0455$) to 0.44 ($p < 0.0001$) (data not shown). Significant inverse correlations were demonstrated between EPDS and measures of HRQoL including GSF Physical and Mental domains, SF-12 PCS and MCS, as well as MKSSI, GSF-derived utility and VAS, with correlation coefficients ranging from −0.26 ($p = 0.0012$) for EPDS and SF-12 PCS, to −0.76 ($p < 0.0001$) for EPDS and SF-12 MCS. PROMIS GSF-derived utility measures were significantly correlated with measures of HRQoL, MKSSI, EPDS, and VAS with correlation coefficients ranging from 0.22 ($p = 0.0177$) for GSF-derived utility and VAS, to 0.92 ($p < 0.0001$) for GSF-derived utility and GSF Physical T scores. Correlation between SF-12 PCS with both SF-12 MCS and VAS demonstrated lack of statistical significance (correlation coefficient = −0.01, $p = 0.8864$, and −0.01, $p = 0.8862$, respectively).

Discussion

Among a diverse cohort of 161women seeking pregnancy services at clinics in New Haven, CT, we demonstrated that PROMIS GSF Physical and GSF Mental domains are significantly correlated with legacy measures of HRQoL, validated measures for social support (MKSSI), screening for antenatal depression (EPDS), and utility measured by the PROMIS GSF-derived utility and the visual analog scale (VAS). Additionally, PROMIS GSF-derived utility estimating EQ-5D demonstrated a significant correlation with these measures. Our study represents a comprehensive evaluation of selected measures of HRQoL, social support, and depression among a pregnant population, capturing a broad range of HRQoL and utility measures representative of diverse experiences among pregnant women. To our knowledge this is the first study to clinically validate and utilize such PROMIS GSF measures among a pregnant population.

Pregnancy represents a unique health state and has a significant impact on women's lives. To date, studies measuring HRQoL among pregnant women are limited. Previous studies utilized the SF-36 to examine quality of life among pregnant women with a history of domestic violence (Tavoli et al. 2016), assessed SF-36v2 factor structure in early pregnancy (Jomeen and Martin 2005), and evaluated sleep and quality of life

Table 3 Correlation matrix for measures of HRQoL, utility, social support and depression during pregnancy: correlation coefficient (p-value)

	PROMIS GSF Physical T-Score	PROMIS GSF Mental T-Score	SF-12 PCS	SF-12 MCS	MKSSI	EPDS	PROMIS GSF-derived utility	VAS
PROMIS GSF Physical T-Score	1.0000							
PROMIS GSF Mental T-Score	0.7024 (<0.0001)	1.0000						
SF-12 PCS	0.4015 (<0.0001)	0.2568 (0.0016)	1.0000					
SF-12 MCS	0.4881 (<0.0001)	0.6561 (<0.0001)	*−0.0116 (0.8864)*	1.0000				
MKSSI	0.4176 (<0.0001)	0.5014 (<0.0001)	0.2459 (0.0036)	0.3748 (<0.0001)	1.0000			
EPDS	−0.6199 (<0.0001)	−0.7064 (<0.0001)	−0.2587 (0.0012)	−0.7649 (<0.0001)	−0.4644 (<0.0001)	1.0000		
PROMIS GSF-derived utility	0.9201 (<0.0001)	0.8416 (<0.0001)	0.3804 (<0.0001)	0.6080 (<0.0001)	0.5238 (<0.0001)	−0.7222 (<0.0001)	1.0000	
VAS	0.1877 (0.0385)	0.2862 (0.0014)	*−0.0131 (0.8862)*	0.2596 (0.0039)	0.1999 (0.0322)	−0.3275 (0.0002)	0.2172 (0.0177)	1.0000

Correlation coefficients in italics represent those not meeting statistical significance

MKSSI (Modified Kendler Social Support Index)

EPDS (Edinburgh Postpartum Depression Scale)

VAS (Visual Analog Scale)

during pregnancy using SF-12v2 (Tsai et al. 2016). The PROMIS Fatigue Short Form has been evaluated among a cohort of pregnant women (Lyon et al. 2014) and a small case series study evaluated quality of life among pregnant chiropractic patients using the PROMIS-29 (Alcantara et al. 2015). Additionally, the World Health Organization Quality of Life Questionnaire and General Health Questionnaire have been used to measure Iranian women's quality of life during the antenatal and postpartum periods (Mortazavi et al. 2014) and the SF-36v1 has also been used to measure HRQoL during pregnancy among Taiwanese women (Chang et al. 2014). In Canada, a recent study evaluated HRQoL during gestation and postpartum among women with assisted reproduction (Vinturache et al. 2015). Another study examined HRQoL among women with ectopic pregnancies using SF-36 and other measures (van Mello et al. 2015). While these studies focused on subgroups of pregnant women using a specific QoL tool, our study evaluated PROMIS GSF and validated scales of HRQoL among a diverse population of pregnant women in the United States.

Findings from our study are important for several reasons. PROMIS GSF domains were found to be highly correlated with HRQoL legacy measures, including SF-12 Physical and Mental Component scores, demonstrating PROMIS GSF is clinically useful in evaluating HRQoL among a population of pregnant women. PROMIS GSF may also have advantages over the SF-12 and legacy HRQoL measures in that they are publically available at no cost, brief in scope and administration time, translated into multiple languages (Spanish, Dutch, French, German and Italian), compartmentalized into physical and mental health scores, and can be used to calculate a utility score for both clinical and health economic analyses. Given the specific health state pregnancy represents, these findings may lead to development of pregnancy-specific HRQoL instruments and domain measures, further use of specific PROMIS measures (e.g. depression, pain, fatigue) during pregnancy, as well as implementation of HRQoL instruments utilizing computerized adaptive testing (CAT).

Within our cohort, quality of life measures exhibit wide variation in pregnancy, indicated by PROMIS GSF T-score and SF-12 distributions, demonstrating pregnancy is a health state experienced differently by individual women. PROMIS T-score measures among our cohort included both the upper and lower bounds of both Physical and Mental Health domains, illustrating the breadth of individual HRQoL measures. Establishing a useful tool to identify significantly lower levels of HRQoL during pregnancy may identify women at risk for poor pregnancy or maternal outcomes, individuals requiring additional antepartum and postpartum care, and those in need of social support services or mental health referrals. The strong inverse correlation of PROMIS domains with EPDS measures represents lower quality of life scores associated with high EPDS scores (which represent a positive screening for depression), and may serve as a useful tool to identify women at risk of depression or indicative of additional mental health screening. Our findings also support the recent recommendation by the United States Preventative Services Task Force to screen all pregnant women for depression (Siu et al. 2016). Additionally, PROMIS GSF domains correlate significantly with MKSSI measures of overall social support and social support subscales. Among our cohort, PROMIS GSF-derived utility scores estimating the

EQ-5D demonstrate a strong correlation with HRQoL legacy measures, MKSSI, and EPDS, and VAS, similar to PROMIS-GSF Physical and GSF Mental T-scores, illustrating the utility of this measure among a diverse pregnant population. Limited studies have evaluated EQ-5D among reproductive populations (Lubinga et al. 2013; Petrou et al. 2009; Shaheen and Lindholm 2006), and our findings further support the use of this GSF-derived utility measure within reproductive health research.

There are several limitations of this study. Our cohort included women seeking care from urban, inner city clinics, therefore participants may not be representative of the general population of pregnant women. Women who declined participation may also differ from those who enrolled. However, our sample is diverse in socio-demographic characteristics, such as race and ethnicity, age, education, and employment status, with approximately one-quarter of study participants completing the study in Spanish. Additionally, 73% of women reported they did not intend to get pregnant or their intentions were changing (57% unintended and 16% intentions changing, respectively); the proportion of unintended pregnancies in this cohort is similar to reported national rates of unintended pregnancy (51%) (Finer and Zolna 2014), further demonstrating generalizability of this cohort. While we have demonstrated overall internal consistency of the PROMIS GSF and related domains, GSF-derived utility, and EPDS (coefficients greater than 0.8), coefficients for the GSF Physical domain (0.63) and MKSSI (0.62) may warrant further evaluation.

With respect to HRQoL measures, questions contained in the PROMIS and legacy HRQoL instruments are not specifically tailored to pregnancy, and may not fully assess HRQoL during this health state. Development of more refined, pregnancy-specific measures may optimize quality of life assessment among these groups. Nevertheless, PROMIS measures among this cohort represent a broad range of T-score values within each domain and were highly correlated with other measures of HRQoL, social support and depression scales, as well as measures of utility. Another limitation of the study is that our measures reflect HRQoL in early pregnancy only; however, this is important as prospective, antenatal HRQoL assessment has not been comprehensively evaluated. Our analysis was also restricted to descriptive evaluation and bivariate measures of correlation with validated instruments. This approach is an initial step to better understand the usefulness of HRQoL measures among a pregnant population and association with existing traditional measures, as well as hypothesis-generating information for future research. PROMIS-GSF may provide a robust tool for measuring HRQoL during pregnancy, to allow comparison of HRQoL within different pregnancy states, such as HRQoL scores for women with intended compared to unintended pregnancy, and comparison of pregnancy HRQoL scores to other health states, such as asthma, obesity and diabetes. Due to the limited research evaluating HRQoL among pregnant populations, this study adds to the literature by identifying tools to adequately measure quality of life and other health parameters, which is important for timely and appropriate pregnancy care, services, and support.

Among our cohort of 161 pregnant women, PROMIS GSF Physical and Mental Health domains significantly correlated with legacy measures of HRQoL, including SF-12 component scores, and measures of social support, depression, and utility. The GSF-derived utility measure also demonstrated high correlation with HRQoL measures and traditional utility (VAS) among this cohort. As a brief 10-item tool that is publicly available, the PROMIS GSF instrument demonstrates clinical validity among a diverse

population of newly diagnosed pregnant women seeking care in urban, clinical settings. With HRQoL increasingly recognized as a critical measure of patient-reported outcomes research, the use of PROMIS GSF during pregnancy and continued development of validated QoL instruments specific for pregnancy and other reproductive health outcomes is essential in furthering efforts to optimize women's reproductive health and overall HRQoL.

Compliance with Ethical Standards

Conflict of Interest Dr. Gariepy received a Clinical Translational Science Award (CTSA Grant Number UL1 TR000142) from the National Center for Advancing Translational Science (NCATS). Dr. Lundsberg was supported by CTSA Grant Number UL1 TR000142. Dr. Yonkers has no conflicts of interest related to the current publication; outside the submitted work; she has worked as a consultant for Pontifax regarding premenstrual dysphoric disorders, received NIH funding, and royalties from Up-to-Date. Dr. Schwarz and Dr. Vilardo declare no conflict of interest.

Ethical Approval All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Funding This study was funded by CTSA Grant Number UL1 TR000142 from the National Center for Advancing Translational Science (NCATS), a component of the National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NIH.

References

Alcantara, J., Alcantara, J. D., & Alcantara, J. (2015). The use of validated outcome measures in the chiropractic care of pregnant patients: a systematic review of the literature. *Complementary Therapies in Clinical Practice, 21*(2), 131–136. doi:10.1016/j.ctcp.2015.01.003.

Biratu, A., & Haile, D. (2015). Prevalence of antenatal depression and associated factors among pregnant women in Addis Ababa, Ethiopia: a cross-sectional study. *Reproductive Health, 12*, 99. doi:10.1186/s12978-015-0092-x.

Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., et al. (2010). 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, 63*(11), 1179–1194. doi:10.1016/j.jclinepi.2010.04.011.

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., et al. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Medical Care, 45*(5 Suppl 1), S3–S11. doi:10.1097/01.mlr.0000258615.42478.55.

Chang, S. R., Chen, K. H., Lin, M. I., Lin, H. H., Huang, L. H., & Lin, W. A. (2014). A repeated measures study of changes in health-related quality of life during pregnancy and the relationship with obstetric factors. *Journal of Advanced Nursing, 70*(10), 2245–2256. doi:10.1111/jan.12374.

Cox, J. L., Chapman, G., Murray, D., & Jones, P. (1996). Validation of the Edinburgh Postnatal Depression Scale (EPDS) in non-postnatal women. *Journal of Affective Disorders, 39*(3), 185–189.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh postnatal depression scale. *British Journal of Psychiatry, 150*, 782–786.

Creinin, M. D. (2000). Randomized comparison of efficacy, acceptability and cost of medical versus surgical abortion. *Contraception, 62*(3), 117–124.

Curtin, S.C., Abma, J.C., & Ventura, S.J. (2013). Pregnancy rates for U.S. women continue to drop. In NCHS Data Brief Number 136, U.S. Department of Health and Human Services, CDC.

DeWalt, D. A., Rothrock, N., Yount, S., Stone, A. A., & Group, P. C. (2007). Evaluation of item candidates: the PROMIS qualitative item review. *Medical Care, 45*(5 Suppl 1), S12–S21. doi:10.1097/01.mlr.0000254567.79743.e2.

EuroQol, G. (1990). EuroQol–a new facility for the measurement of health-related quality of life. *Health Policy, 16*(3), 199–208.

Finer, L. B., & Zolna, M. R. (2014). Shifts in intended and unintended pregnancies in the United States, 2001–2008. *American Journal of Public Health, 104*(Suppl 1), S43–S48. doi:10.2105/AJPH.2013.301416.

Gibson, J., McKenzie-McHarg, K., Shakespeare, J., Price, J., & Gray, R. (2009). A systematic review of studies validating the Edinburgh postnatal depression scale in antepartum and postpartum women. *Acta Psychiatrica Scandinavica, 119*(5), 350–364. doi:10.1111/j.1600-0447.2009.01363.x.

Goldhaber-Fiebert, J. D., & Brandeau, M. L. (2015). Evaluating cost-effectiveness of interventions that affect fertility and childbearing: how health effects are measured matters. *Medical Decision Making, 35*(7), 818–846. doi:10.1177/0272989X15583845.

Hays, R. D., Bjorner, J. B., Revicki, D. A., Spritzer, K. L., & Cella, D. (2009). Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Quality of Life Research, 18*(7), 873–880. doi:10.1007/s11136-009-9496-9.

Heron, J., O'Connor, T. G., Evans, J., Golding, J., Glover, V., & Team, A. S. (2004). The course of anxiety and depression through pregnancy and the postpartum in a community sample. *Journal of Affective Disorders, 80*(1), 65–73. doi:10.1016/j.jad.2003.08.004.

Jomeen, J., & Martin, C. R. (2005). The factor structure of the SF-36 in early pregnancy. *Journal of Psychosomatic Research, 59*(3), 131–138. doi:10.1016/j.jpsychores.2005.02.018.

Klem, M., Saghafi, E., Abromitis, R., Stover, A., Dew, M. A., & Pilkonis, P. (2009). Building PROMIS item banks: librarians as co-investigators. *Quality of Life Research, 18*(7), 881–888. doi:10.1007/s11136-009-9498-7.

Lubinga, S. J., Levine, G. A., Jenny, A. M., Ngonzi, J., Mukasa-Kivunike, P., Stergachis, A., et al. (2013). Health-related quality of life and social support among women treated for abortion complications in western Uganda. *Health and Quality of Life Outcomes, 11*, 118. doi:10.1186/1477-7525-11-118.

Lyon, D., McCain, N., Elswick, R. K., Sturgill, J., Ameringer, S., Jallo, N., et al. (2014). Biobehavioral examination of fatigue across populations: report from a P30 Center of Excellence. *Nursing Outlook, 62*(5), 322–331. doi:10.1016/j.outlook.2014.06.008.

Matthey, S., Henshaw, C., Elliott, S., & Barnett, B. (2006). Variability in use of cut-off scores and formats on the Edinburgh postnatal depression scale: implications for clinical and research practice. *Archives of Women's Mental Health, 9*(6), 309–315. doi:10.1007/s00737-006-0152-x.

Mogos, M. F., August, E. M., Salinas-Miranda, A. A., Sultan, D. H., & Salihu, H. M. (2013). A systematic review of quality of life measures in pregnant and postpartum mothers. *Applied Research in Quality of Life, 8*(2), 219–250.

Mortazavi, F., Mousavi, S. A., Chaman, R., & Khosravi, A. (2014). Maternal quality of life during the transition to motherhood. *Iranian Red Crescent Medical Journal, 16*(5), e8443. doi:10.5812/ircmj.8443.

Myers, E. R. (2015). How should we estimate the cost-effectiveness of interventions that affect reproduction? *Medical Decision Making, 35*(7), 812–814. doi:10.1177/0272989X15602227.

Petrou, S., Morrell, J., & Spiby, H. (2009). Assessing the empirical validity of alternative multi-attribute utility measures in the maternity context. *Health and Quality of Life Outcomes, 7*, 40. doi:10.1186/1477-7525-7-40.

PROMIS (2010). Scoring PROMIS Global Short Form. https://www.assessmentcenter.net/documents/Scoring%20PROMIS%20Global%20short%20form.pdf. Accessed 27 Jan 2016.

Revicki, D. A., Kawata, A. K., Harnam, N., Chen, W. H., Hays, R. D., & Cella, D. (2009). Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample. *Quality of Life Research, 18*(6), 783–791. doi:10.1007/s11136-009-9489-8.

Robinson, A., Loomes, G., & Jones-Lee, M. (2001). Visual analog scales, standard gambles, and relative risk aversion. *Medical Decision Making, 21*(1), 17–27.

Rochat, T. J., Richter, L. M., Doll, H. A., Buthelezi, N. P., Tomkins, A., & Stein, A. (2006). Depression among pregnant rural South African women undergoing HIV testing. *JAMA, 295*(12), 1376–1378. doi:10.1001/jama.295.12.1376.

Schwarz, E. B., Smith, R., Steinauer, J., Reeves, M. F., & Caughey, A. B. (2008). Measuring the effects of unintended pregnancy on women's quality of life. *Contraception, 78*(3), 204–210. doi:10.1016/j.contraception.2008.04.120.

Senturk, V., Abas, M., Berksun, O., & Stewart, R. (2011). Social support and antenatal depression in extended and nuclear family environments in Turkey: a cross-sectional survey. *BMC Psychiatry, 11*, 48. doi:10.1186/1471-244X-11-48.

Shaheen, R., & Lindholm, L. (2006). Quality of life among pregnant women with chronic energy deficiency in rural Bangladesh. *Health Policy, 78*(2–3), 128–134. doi:10.1016/j.healthpol.2005.11.008.

Siu, A. L., U. S. P. S. T. Force, Bibbins-Domingo, K., Grossman, D. C., Baumann, L. C., Davidson, K. W., et al. (2016). Screening for depression in adults: US preventive services task force recommendation statement. *JAMA, 315*(4), 380–387. doi:10.1001/jama.2015.18392.

Spoozak, L., Gotman, N., Smith, M. V., Belanger, K., & Yonkers, K. A. (2009). Evaluation of a social support measure that may indicate risk of depression during pregnancy. *Journal of Affective Disorders, 114*(1–3), 216–223. doi:10.1016/j.jad.2008.07.015.

Tavoli, Z., Tavoli, A., Amirpour, R., Hosseini, R., & Montazeri, A. (2016). Quality of life in women who were exposed to domestic violence during pregnancy. *BMC Pregnancy and Childbirth, 16*(1), 19. doi:10.1186/s12884-016-0810-6.

Teresi, J. A., Ocepek-Welikson, K., Kleinman, M., Eimicke, J. P., Crane, P. K., Jones, R. N., et al. (2009). Analysis of differential item functioning in the depression item bank from the Patient Reported Outcome Measurement Information System (PROMIS): an item response theory approach. *Psychology Science Quarterly, 51*(2), 148–180.

Torrance, G. W. (1987). Utility approach to measuring health-related quality of life. *Journal of Chronic Diseases, 40*(6), 593–603.

Tsai, S. Y., Lee, P. L., Lin, J. W., & Lee, C. N. (2016). Cross-sectional and longitudinal associations between sleep and health-related quality of life in pregnant women: a prospective observational study. *International Journal of Nursing Studies*. doi:10.1016/j.ijnurstu.2016.01.001.

Ubel, P. A., Loewenstein, G., & Jepson, C. (2003). Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. *Quality of Life Research, 12*(6), 599–607.

van Mello, N. M., Mol, F., Hajenius, P. J., Ankum, W. M., Mol, B. W., van der Veen, F., et al. (2015). Randomized comparison of health-related quality of life in women with ectopic pregnancy or pregnancy of unknown location treated with systemic methotrexate or expectant management. *European Journal of Obstetrics, Gynecology and Reproductive Biology, 192*, 1–5. doi:10.1016/j.ejogrb.2015.06.007.

Vinturache, A., Stephenson, N., McDonald, S., Wu, M., Bayrampour, H., & Tough, S. (2015). Health-related quality of life in pregnancy and postpartum among women with assisted conception in Canada. *Fertility and Sterility, 104*(1), 188 e181–195 e181. doi:10.1016/j.fertnstert.2015.04.012.

Ware, J., Jr., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220–233.

Washington, C. I., Jamshidi, R., Thung, S. F., Nayeri, U. A., Caughey, A. B., & Werner, E. F. (2015). Timing of postpartum intrauterine device placement: a cost-effectiveness analysis. *Fertility and Sterility, 103*(1), 131–137. doi:10.1016/j.fertnstert.2014.09.032.

Weinstein, M. C., Torrance, G., & McGuire, A. (2009). QALYs: the basics. *Value in Health, 12*(Suppl 1), S5–S9. doi:10.1111/j.1524-4733.2009.00515.x.