Received: 21 September 2017 | First decision: 10 October 2017 | Accepted: 2 January 2018

DOI: 10.1111/apt.14531

WILEY AP&T Alimentary Pharmacology & Therapeutics

Psychometric properties of the PROMIS short form measures in a U.S. cohort of 961 patients with chronic hepatitis C prescribed direct acting antiviral therapy

D. M. Evon[1] | J. Amador[2] | P. Stewart[2] | B. B. Reeve[3] | A. S. Lok[4] | R. K. Sterling[5] | A. M. Di Bisceglie[6] | N. Reau[7] | M. Serper[8] | S. Sarkar[9] | J. K. Lim[10] | C. E. Golin[11,12] | M. W. Fried[1]

[1]Division of Gastroenterology and Hepatology, University of North Carolina, Chapel Hill, USA

[2]Department of Biostatistics, University of North Carolina, Chapel Hill, USA

[3]Department of Population Health Sciences, Duke University, Durham, USA

[4]Division of Gastroenterology and Hepatology, University of Michigan, Ann Arbor, USA

[5]Division of Gastroenterology, Hepatology, and Nutrition Virginia Commonwealth University, Richmond, USA

[6]Department of Internal Medicine, Saint Louis University, St Louis, USA

[7]Department of Internal Medicine, Rush Medical Center, Chicago, USA

[8]Department of Medicine, University of Pennsylvania, Philadelphia, USA

[9]Department of Internal Medicine, University of California, Davis, USA

[10]Department of Internal Medicine, Yale University, New Haven, USA

[11]Division of General Medicine and Clinical Epidemiology, University of North Carolina, Chapel Hill, USA

[12]Department of Health Behavior, University of North Carolina, Chapel Hill, USA

Correspondence
Dr D M Evon, Division of Gastroenterology and Hepatology, University of North Carolina, Chapel Hill, NC, USA.
Email: donna_evon@med.unc.edu

Funding information
University of North Carolina Center for AIDS Research, Grant/Award Number: P30 AI-50410; Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: K24-HD06920; Patient-Centered Outcomes Research Institute, Grant/Award Number: CER-1408-20660; Center for Gastrointestinal Biology and Disease, Grant/Award Number: P30-DK34987

Summary

Background: To better understand symptoms experienced by patients infected with chronic hepatitis C virus (HCV), valid and reliable patient-reported outcome (PRO) measures are needed.

Aim: To assess the reliability and validity of 10 patient-reported outcomes measurement information system (PROMIS) measures and the Headache Impact Test-6 (HIT-6) in a large national sample of patients with HCV.

Methods: Pre-treatment data from 961 patients with HCV starting direct acting antiviral therapy at 11 U.S. liver centers were analyzed. Internal reliability was evaluated using Cronbach's alpha coefficient; frequency distributions were examined for floor and ceiling effects; structural validity was investigated via item-response-theory models; convergent validity was evaluated using correlations with theoretically-similar items from the HCV-PRO and memorial symptom assessment scale (MSAS); and known-groups validity was investigated by observing PRO differences by liver disease status and number of comorbidities.

Results: The HIT-6 and the majority of the PROMIS measures yielded excellent reliability (alphas ≥ 0.87). Ceiling effects were infrequent ($< 4\%$), while 30%-59% of patients reported no symptoms (floor effects). The data supported structural validity of the HIT-6 and most PROMIS measures. The PROMIS measures showed moderate to strong correlations with theoretically-similar items from the HCV-PRO and MSAS (0.39-0.77). Trends were observed between worse PRO scores and advanced cirrhosis and greater number of comorbidities, lending support for known-groups validity.

Conclusions: The psychometric properties of the HIT-6 and PROMIS measures performed satisfactorily in this large cohort of patients with HCV starting direct acting antiviral therapy. Opportunities exist for further refinement of these PROs. Evaluation of performance over time and in under-represented subgroups is needed.

The Handling Editor for this article was Professor Grace Wong, and it was accepted for publication after full peer-review.

1 | INTRODUCTION

People living with chronic hepatitis C virus (HCV) infection often report a broad array of physical and mental symptoms[1-4] including but not limited to, fatigue, musculoskeletal pain, poor appetite, nausea, abdominal pain, cognitive impairment, depression, anxiety, irritability and sleep distubance.[1,5] Numerous studies have investigated health related quality of life (HRQOL) in patients living with chronic HCV.[6,7] However, a comprehensive quantitative analysis of patients' experiences of *specific symptoms* that may be associated with HCV has not been conducted.[5]

The new direct acting antiviral (DAA) therapies to treat HCV are very well tolerated compared to previous treatments; nonetheless, they can still cause side effects, such as fatigue, nausea, and headache.[8-11] Thus far, our understanding of patients' experiences with DAA therapy has been limited to data derived from industry-sponsored DAA trials.[12,13] Those studies provided very useful information regarding the effects of DAAs on patients' HRQOL, work productivity, and fatigue.[14] However, no other symptom experiences have been evaluated from the patients' perspective during DAA therapy.

The impact of disease, treatment and viral cure on peoples' lives is nearly impossible to understand without direct input from patients themselves. Patients' experiences are evaluated using patient-reported outcome (PRO) measures, defined as "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else."[15] PRO measures are the best way to capture patients' experiences about how a disease, its treatment, or cure affects their lives in a meaningful way.

A recent literature review identified 22 key PRO concepts that were found to be important to people living with HCV in qualitative studies.[5] The majority of these concepts are *specific symptoms* such as depression, fatigue, anxiety, cognitive dysfunction, musculoskeletal pain, irritability, sleep problems, lack of appetite, and gastrointestinal symptoms. Many of these *specific symptoms* that concern patients have received inadequate attention in HCV clinical studies. Of the 18 PRO measures utilized in HCV clinical studies, only four have actually been developed or validated in the HCV population, three of which are broad HRQOL measures which are not designed to thoroughly measure *specific symptoms or side effects*.[6,16,17]

In response to increasing recognition of the importance of evaluating patients' experiences of illness and treatment, the National Institutes of Health Patient-Reported Outcomes Measurement Information System (PROMIS) initiative has developed a comprehensive well-evaluated set of self-report tools, some of which measure *specific symptoms*.[18,19] PROMIS measures have been used in a few hepatology studies,[20,21] one of which validated several instruments delivered via a computerized adaptive testing system in patients with cirrhosis.[20] PROMIS measures have several advantages over other PRO measures: items were developed using qualitative and quantitative methods; surveys are designed to measure a single symptom and not overlap with others; surveys are brief (4-8 items); and scores can be compared across other health conditions and populations.

In the current study, we evaluated the reliability and validity of several PROMIS short form measures and the Headache Impact Test (HIT-6), a symptom not assessed by PROMIS.[22] The specific aims of the study were to evaluate the following psychometric properties: (1) internal reliability; (2) floor and ceiling effects; and (3) construct validity (structural validity, convergent validity, and known-groups validity).

2 | MATERIALS AND METHODS

2.1 | Parent study

The psychometric properties of PROMIS measures were evaluated using baseline data from the Patient-Reported Outcomes Project of HCV-TARGET ("PROP UP") (Clinical trial.gov: NCT02601820). A complete description of the PROP UP study is reported elsewhere.[23] Briefly, PROP UP is a multi-site, longitudinal, observational, study designed to enroll 1,600 patients with chronic HCV infection at 11 U.S. liver centers to characterize patients' experiences before, during and after DAA therapy. At five time points during the study, patients completed multiple PRO measures in the clinic, via email link to a data capture system, or through phone interview with trained staff from a centralized call center. Completion of surveys at each time point averaged 25 minutes (range: 15-45 minutes). Recruitment began in January 2016. All sites received approval from their local Institutional Review Board prior to recruitment.

2.2 | Study design

Data from baseline assessments were used to conduct a cross-sectional study of the reliability and validity of the HIT-6 and 10 PROMIS measures. Patients who consented and completed baseline surveys were included in this analysis.

2.3 | Research participants

Inclusion/exclusion criteria have been previously reported.[23] Briefly, patients were invited to participate in PROP UP if they were diagnosed with chronic HCV infection of any genotype; English-speaking; age 21 years or older; and prescribed one of five DAA regimens. For this analysis, we included data from a cohort of 961 patients who had completed baseline surveys at the time of data retrieval (11/28/2016).

2.4 | Measures

2.4.1 | PROMIS measures

The 10 PROMIS short forms included Fatigue-7a, Depression-8a, Anger-5a, Anxiety-4a, Pain Interference-8a, Sleep Disturbance-8a,

Applied Cognition-General Concerns-8a, Belly Pain-6, Diarrhea-6, and Nausea/Vomiting-4 (available via www.HealthMeasures.net). The short forms were used to measure symptoms most common or salient to HCV or DAA therapy. Each PROMIS short form includes a subset of items from a larger item bank that were the best performing items in terms of content validity and reliability.[18,19] PROMIS scores were scaled to standardized T-scores, with a mean of 50 and standard deviation of 10 in the U.S. general population. An exception is Sleep Disturbance, which was normed on a clinical and general population. Higher scores indicate worse symptoms. Studies in other medical populations suggest that the minimally important difference within or between groups generally ranges from 2-5 points.[24-27]

2.4.2 | Headache impact test-6 (HIT-6)

Headaches have been reported as an adverse event during industry-sponsored DAA trials,[8-11] but no PROMIS instrument specifically measures this symptom. Therefore, we used the 6-item Headache Impact Test to measure headaches that could occur during treatment.[22] The six items are responded to on a 5-point Likert scale ranging from "Never" to "Always." The final score is summed and can range from 36 to 78, with higher scores reflecting worse headaches.

2.4.3 | Memorial symptom assessment scale (MSAS)

The MSAS is a reliable and validated instrument used to evaluate 32 symptoms associated with common medical conditions treatment.[28,29] Many of these symptoms may be experienced by patients with HCV or undergoing treatment. Participants first indicate the presence or absence of the 32 symptoms, and if present, then rate the symptom on three attributes (severity, frequency, interference with functioning) on a scale of 1 to 4. Each total item score ranges from 0 (if symptom is absent) to 4 (if present, the average of three attributes). Higher scores indicate worse symptoms. In this study, we used the MSAS as a comparison legacy measure (ie, a previously evaluated measure), using several symptom item scores to evaluate convergent validity with specific PROMIS surveys (eg, MSAS total pain score correlated with PROMIS Pain Interference score). Patient data were excluded from the analysis if more than 13% of items were missing, as recommended by the MSAS developers.[28]

2.4.4 | HCV-PRO

The HCV-PRO is a disease-specific survey that assesses well-being and HRQOL in people with chronic HCV.[30,31] The measure was developed in accordance with the Food and Drug Administration guidelines and demonstrates good reliability and validity and was used as a second legacy measure. The survey includes 16 items with responses ranging from "1 = all of the time" to "5 = none of the time" with the sum transformed to a 0-100 scale. In contrast to the other PRO measures, a *higher* HCV-PRO score indicates *higher* HRQOL. In this study, we used several individual items from the HCV-PRO to evaluate convergent validity with specific PROMIS measures (eg, HCV-PRO item "feeling tired" correlated with PROMIS Fatigue).

2.4.5 | Sociodemographics

Patients self-reported their age, sex, race, ethnicity, marital status, education, and employment status.

2.4.6 | Number of health comorbidities

Patients completed a 35-item medical history form to indicate the presence or absence of various common health conditions.

2.4.7 | Presence or absence of cirrhosis

Classification of cirrhosis status (Yes/No) was based on one or more of the following source documents reviewed by trained study coordinators at each site: Transient Elastography (eg, Fibroscan, Echosens, Waltham, Massachusetts, USA), blood serum markers (eg, FibroSure, Laboratory Corporation of America, Burlington, North Carolina, USA)), and treating clinician's impressions (based on cumulative medical history, liver biopsy, ultrasound, procedures, physical exam, laboratory tests). An expert hepatologist (M.W.F.) assisted with cross-checking cirrhosis status with other variables in the dataset to confirm classification, such as treatment duration and use of ribavirin, laboratory values, and the following blood serum markers: AST to platelet ratio index (APRI) > 2.0,[32] FIB-4 score > 3.25,[33] or Model for End-Stage Liver Disease (MELD) score > 6.[34] APRI > 2.0 and MELD > 6 are suggestive of cirrhosis, FIB-4 > 3.25 is suggestive of advanced fibrosis, and MELD > 12 is suggestive of advanced cirrhosis.[32,33]

2.5 | Statistical analysis strategy

The internal consistency reliability of each measure was evaluated using point- and interval-estimates of Cronbach's alpha. We considered alpha > 0.70 to be an acceptable minimum criterion for establishing reliability in the cohort of patients studied.[35,36] For generalization from our sample to a target population, we infer reliability in that population if the observed lower 95% confidence limit for alpha exceeds 0.70.[37] Due to the substantial number of participants who did not experience the symptom, the estimates of Cronbach's alpha for the PROMIS Belly Pain, Diarrhea, and Nausea/Vomiting were computed using only the data from participants who reported the presence of that symptom because many patients did not experience this symptom at all. In sensitivity analyses used to evaluate the impact of this approach, the alpha estimates were also computed based on all participants.

Ceiling and floor effects were evaluated in terms of the proportion of participants who had the maximum score (or minimum

score).[38] Skewness and other properties of the frequency distributions of the scores were examined via graphical descriptive methods.

Factor-analytic assessment of structural validity (ie, underlying unidimensionality) of each measure relied on unidimensional and multidimensional graded-response item-response-theory models.[39] Overall model fit was evaluated based on the root mean square error of approximation (RMSEA).[40] An RMSEA value less than or equal to 0.06 was considered to indicate good fit, a value within 0.06-0.08 is a fairly good fit, and a value above 0.10 is a poor fit.[41] The *S* chi-square test procedure was used to test the null hypothesis which states that the item-level fit is adequate.[42] Relative to longer questionnaires, shorter questionnaires provide less power and precision to detect and quantify item-level misfit. We assumed that item-level fit for the short-form measures is similar to the item-level fit for long-form measures. The standardized local dependence (LD) chi-square test procedure was used to detect items that are excessively related after controlling for the underlying domain: LD test statistic values larger than 10.0 identified items with substantial LD.[43] In sensitivity analyses, to evaluate the robustness of the main results to reasonable perturbation of the methods used, the LD tests were adjusted for multiple comparisons. Sensitivity analyses also included an additional check on the underlying dimensionality of the measure performed by fitting a bi-factor graded-response item-response-theory modelwith each identified LD pair or set of items as a second order factor.[44] Any LD violations were deemed negligible if they explained common variance (ECV) was at least 0.90.[45,46] ECV represents the variance explained by the general factor in the bi-factor model.

To evaluate convergent validity, we used the absolute value of point- and interval-estimates of the Spearman correlations (ρ) between measures of theoretically-similar domains. For this purpose, for each of the PROMIS and HIT-6 measures, we identified item pairs which, logically, should be positively correlated. We anticipated that there should exist substantial positive correlations between those item pairs. We used Dancey and Reidy's classification which specifies that 0.40-0.70 is the range of a "moderate" correlation.[47]

Known-groups validity helps to support construct validity if it can be demonstrated that an instrument discriminates between groups that are known to differ on a given construct. The extant literature in HCV is extremely limited on definitive groups that would differ according to *precise symptoms.* However, a few studies, though not all suggest that HRQOL and one symptom, fatigue, may differ among patients with and without cirrhosis.[16,48-50] Based on this literature, the a priori analysis plan was to compare patients categorized as cirrhotic or noncirrhotic based on medical chart review (as described above). Secondarily, we used a three-category approach to better observe PRO scores in patients with and without cirrhosis, stratified by their baseline MELD scores: non-cirrhosis according to medical chart review; MELD score 6-12; and MELD > 12. Thirdly, based on findings from Rothrock et al, we evaluated potential known-groups validity by exploring the association between the number of health comorbidities as a continuous variable with symptom scores, speculating that patients with more comorbid conditions should experience greater symptoms.[51]

Unless otherwise specified, all statistical computations were performed using SAS System software version 9.4 (SAS Institute, Cary, NC). Analysis of the graded-response item-response-theory models were performed using the SAS IRT procedure and using IRT-PRO software, version 4.1 (Scientific Software International, Lincolnwood, IL). PROMIS T-scores were computed using RSTUDIO software, version 1.0.136 (RStudio Inc.).

3 | RESULTS

3.1 | Patient characteristics

The characteristics of the 961 study participants at baseline are presented in Table 1. Mean age was 57 (SD = 10.7) years old. In addition to having chronic HCV and liver disease, participants reported an average of 4 other health comorbidities (range 0-15).

3.2 | Missing values

The frequency of missing data among the HIT-6 and PROMIS short form scores were less than 2.2%. The frequency of missing data for the HCV-PRO score was 7.7%. We judged the reasons for the incomplete data to satisfy the "missing completely at random" criterion.[52]

TABLE 1 Patient characteristics (N = 961)

	N[a]	Mean (SD), Median	Range
Age (y)	955	57 (10.7), 58	23-82
Number of comorbidities	961	4.1 (2.9), 4	0-15
	N[a]	**Categories**	**N (%)**
Sex	951	Male	523 (55.0)
		Female	428 (45.0)
Race	948	White	581 (61.3)
		Black or African American	311 (32.8)
		Other	56 (5.9)
Ethnicity	907	Not Hispanic or Latino	760 (83.8)
		Hispanic or Latino	37 (4.1)
		Other	110 (12.1)
Marital status	953	Married/in committed relationship	357 (37.5)
		Single	345 (36.2)
		Separated/Divorced/Widowed	251 (26.3)
Education	953	High school or equivalent diploma	527 (55.3)
		Vocational school or higher	426 (44.7)
Employment	937	Unemployed/Disabled/Applying	460 (49.1)
		Working full/part time	369 (39.4)
		Retired/Homemaker/Student	108 (11.5)
Cirrhosis status	912	Present	456 (50.0)
		Absent	456 (50.0)

SD, Standard deviation.
[a]N is the number of non-missing values. Missing data ranged from 0% to 5.6%; collectively only 1.5% of data values were missing.

3.3 | Internal consistency reliability

Reliability of each of the symptom domains from the HIT-6, PROMIS short forms, and HCV-PRO are presented in Table2. The HIT-6 and HCV-PRO measures had high reliability (all Cronbach's alphas > 0.90). All PROMIS short forms, except the PROMIS Nausea/Vomiting, yielded very good reliability estimates between 0.87 and 0.98.

3.4 | Floor and ceiling effects

Ceiling effects were consistently low among all PRO measures (< 4% of participants had scores at the upper limit), suggesting that all measures performed well in capturing severe symptoms at the upper limits of the scales (Table 2). In contrast, many patients did not report symptoms and therefore notable floor effects (range 30% to 59%) were observed for the HIT-6 and a majority of the PROMIS measures. Floor effects were small for PROMIS Fatigue (1.6%) and Sleep Disturbance (6.8%) indicating that most patients reported at least some mild fatigue and sleep disturbance problems. The HCV-PRO had a negligible floor (0.10%) effect and a small ceiling effect (5.9%).

3.5 | Structural validity

Findings supported the unidimensionality of the HIT-6, PROMIS measures, and HCV-PRO (Table 3). Eight PRO measures had a very good fit (RMSEA ≤ 0.06). None of the measures had a poor fit (RMSEA > 0.10). The majority of the confirmatory factor analysis models fitted to the measures produced high factor loadings (> 0.70). Exceptions were noted for the PROMIS Fatigue, Nausea/Vomiting and HCV-PRO, which showed multidimensionality. The PROMIS Fatigue measure had a small factor loading (0.30) for one item ("How often did you have enough energy to exercise strenuously"). This item is the only item on the survey that is reverse-scored and may explain why it did not hang together with the other items. Nonetheless, a good fit was suggested by the small RMSEA value of 0.04 and the absence of local dependence (LD). The PROMIS Nausea/Vomiting survey and HCV-PRO had items with factor loadings below 0.70 likely because these measures tap into multiple constructs (eg, nausea and vomiting). Despite this, the overall model fit was still very good with RMSEA values < 0.06.

3.6 | Convergent validity

As shown in Table 4, both the HIT-6 score and the PROMIS Belly Pain score were moderately correlated with the HCV-PRO item "I felt bothered by pain or physical discomfort" and with the MSAS pain score (range 0.37 to 0.48).

The PROMIS Pain Interference score was strongly correlated with the HCV-PRO item "I felt bothered by pain or physical discomfort" (0.74) and MSAS pain score (0.76). PROMIS Fatigue, Sleep Disturbance, and Cognitive Concerns were strongly correlated with HCV-PRO items and MSAS scores of similar constructs (range 0.67 to 0.78). PROMIS Depression, Anger, Anxiety, and Diarrhea were moderately correlated with specific HCV-PRO items and MSAS scores of similar constructs (range 0.48 to 0.67). PROMIS Nausea/

TABLE 2 Reliability of the HIT-6, PROMIS short forms, and HCV-PRO (N = 961)

Measure	N	Items	Alpha[a]	95% CI[b]	N	Floor (%)[c]	Ceiling (%)[c]
HIT-6 score[d]	944	6	0.93	[0.92, 0.94]	944	30.9	0.2
PROMIS T-scores[d]							
Fatigue	941	7	0.87	[0.86, 0.88]	960	1.6	0.2
Pain interference	944	8	0.98	[0.98, 0.98]	961	36.7	3.9
Sleep disturbance	945	8	0.94	[0.93, 0.95]	960	6.8	2.5
Depression	942	8	0.96	[0.95, 0.96]	960	33.9	0.7
Cognitive concerns	940	8	0.96	[0.96, 0.97]	961	30.1	1.4
Anger	958	5	0.91	[0.91, 0.92]	961	18.3	1.0
Anxiety	954	4	0.91	[0.90, 0.92]	954	40.0	0.6
Belly pain[e]	470	6	0.87	[0.85, 0.89]	960	49.7	0.1
Diarrhea[e]	388	6	0.88	[0.87, 0.90]	948	58.5	0.3
Nausea/Vomiting[e]	452	4	0.63	[0.58, 0.69]	949	52.4	0.1
HCV-PRO[f]	887	16	0.94	[0.94, 0.95]	887	0.1	5.9

[a]Cronbach's alpha coefficient.
[b]95% confidence interval (CI) for Cronbach's alpha.
[c]Proportion of responses at the minimum (floor) or maximum (ceiling) of the scale.
[d]Higher values of HIT-6 and PROMIS T-scores indicate more severe symptoms.
[e]Only patients who reported presence of the symptom were analyzed.
[f]Shown for reference in examining the properties of the HIT-6 and PROMIS scores. Higher values of HCV-PRO score indicate higher functional well-being.

TABLE 3 Structural validity of the HIT-6, PROMIS short forms, and HCV-PRO (N = 961)

Measure	Items	Factor loadings[a]	RMSEA[b]	ECV[c]
HIT-6 score[d]	6	0.86-0.97	0.06	—[e]
PROMIS[d]				
Fatigue	7	0.30[f], 0.76-0.91	0.04	—[e]
Pain interference	8	0.95-0.98	0.09	0.96
Sleep disturbance	8	0.84-0.92	0.08	0.86
Depression	8	0.90-0.96	0.05	0.96
Cognitive Concerns	8	0.90-0.96	0.07	0.98
Anger	5	0.88-0.92	0.06	0.96
Anxiety	4	0.89-0.93	0.05	0.96
Belly pain	6	0.93-0.98	0.05	0.99
Diarrhea	6	0.83-0.98	0.07	0.87
Nausea/Vomiting	4	0.66[g], 0.84-0.96	0.05	—[e]
HCV-PRO[h]	16	0.55[i], 0.67[j], 0.71-0.86	0.04	0.82

[a]Factor loadings obtained from the unidimensional confirmatory factor analysis. Factor loadings > 0.70 support unidimensionality.
[b]RMSEA, root mean square error approximation. RMSEA values ≤ 0.06 indicate good fit, values ≤ 0.08 are fair, and values above 0.10 generally reflect poor fit. Statistics were based on full marginal tables. Model-based weight matrix was used.
[c]ECV, Expected Common Variance. ECV represents the variance explained by the general factor in the bifactor model. ECV > 0.90 support unidimensionality.
[d]Higher values of HIT-6 and PROMIS T-scores indicate more severe symptoms.
[e]Unidimensionality was assessed using a one-factor confirmatory factor analysis.
[f]Factor loading of the item: "In the past 7 days, how often did you have enough energy to exercise strenuously?".
[g]Factor loading of the item: "In the past 7 days, how often did you have a poor appetite?".
[h]Shown for reference in examining the properties of the HIT-6 and PROMIS scores. Higher values of HCV-PRO score indicate higher functional well-being..
[i]Factor loading of the item: "Having Hepatitis C affected my sex life".
[j]Factor loading of the item: "Having Hepatitis C was very stressful to me".

Vomiting was moderately correlated with the MSAS vomiting and nausea items (0.35 and 0.64, respectively).

3.7 | Known-groups validity

Differences in PRO scores between patients defined as noncirrhotic and cirrhotic via medical chart review were negligible (data not shown). In a subsequent analysis (Tables 5 and 6) of three liver disease categories (ie, noncirrhotic, MELD 6-12, MELD > 12), a severity-response trend was observed such that worse symptoms were associated with advancing liver disease. In particular, patients with MELD > 12 had higher (worse) scores on all symptoms compared to the other two groups, with mean differences ranging from 1.9 to 4.1. While the sample size of those with MELD > 12 was small (n = 37) and provided limited precision, the mean differences were comparable to established PROMIS minimally important differences reported in other medical populations.[24-27] In the final known-groups validity analysis, we found positive correlations between a greater number of comorbid health conditions and worse symptom scores, consistent with prior studies.[51] Patients with more comorbid conditions had worse scores on PROMIS Fatigue, Sleep Disturbance, Pain Interference, and HRQOL on the HCV-PRO (moderate correlations ranging from 0.39 to 0.53). Smaller positive correlations were observed between the number of comorbid conditions and worse scores on PROMIS Depression, Cognitive Concerns, Anger, Anxiety, Belly Pain, Nausea, and HIT-6 (ranging from 0.29 to 0.34).

4 | DISCUSSION

The objectives of the current study were to establish the reliability and validity of several PRO measures, specifically 10 PROMIS measures and the HIT-6, to evaluate *specific symptoms* in patients with chronic HCV. Given the large national sample of patients recruited for the PROP UP study, it represented a reasonable platform in which to evaluate the psychometric properties of several new PRO measures that could be useful in future patient-centered outcomes research in the field of chronic HCV. The overall findings from this study suggest that the psychometric properties of these PROs are sufficiently satisfactory to be used in future HCV studies, with specific caveats noted below.

The internal consistency of the symptom scales was well above the acceptable limit of 0.70 (range: 0.87-0.98) indicating very good to excellent reliability. Only one exception was noted for the PROMIS Nausea/Vomiting measure which appears to be multidimensional. This measure had a lower reliability because its four items tap three constructs (nausea, poor appetite, vomiting), which while associated, are not highly correlated.

None of the PROs had notable ceiling effects, which is critically important to investigations that need to capture very severe symptoms, as might be the case with patients listed for liver transplantation. Thus, these PROs are suitable to evaluate patients' experiences of severe and debilitating symptoms.

In contrast, floor effects were notable for most of the PROs, with the exception of the HCV-PRO and PROMIS Fatigue and Sleep Disturbance measures. Floor effects can be present when a measure does not represent the full range of possible human experiences, but we do not believe this to be the case with the PROMIS and HIT-6 measures. Floor effects can also be evident when the PRO score has both a binary component (presence or absence of the symptom) and a continuous component representing the frequency or severity of a symptom, *when present*. In the current sample of patients starting new DAA therapy, we would anticipate that approximately half of these patients would be asymptomatic for any one of the various symptom scales, thus producing minimal scores and demonstrating sizable floor effects.

TABLE 4 Correlation of the HIT-6 and PROMIS short forms with HCV-PRO and MSAS items (N = 961)

Measure	Corresponding measure	N[a]	Correlation[b]	95% CI[c]
HIT-6 score[d]	HCV-PRO Item: I felt bothered by pain or physical discomfort	935	0.45	[0.39, 0.50]
	MSAS[e] pain score	936	0.37	[0.31, 0.43]
PROMIS T-score[d]				
Fatigue	HCV-PRO Item: I felt too tired during the day to get done what I needed	958	0.74	[0.70, 0.77]
	MSAS lack of energy score	949	0.77	[0.74, 0.80]
Pain interference	HCV-PRO Item: I felt bothered by pain or physical discomfort	951	0.74	[0.70, 0.77]
	MSAS pain score	952	0.76	[0.73, 0.79]
Sleep disturbance	HCV-PRO Item: I had difficulty sleeping or slept too much	954	0.74	[0.71, 0.78]
	MSAS difficulty sleeping score	949	0.77	[0.75, 0.80]
Depression	HCV-PRO Item: I felt downhearted and sad	951	0.67	[0.63, 0.71]
	MSAS feeling sad score	949	0.64	[0.60, 0.68]
Cognitive concerns	HCV-PRO Item: I was unable to think clearly or focus on my thoughts	957	0.78	[0.75, 0.81]
	MSAS difficulty concentrating score	951	0.67	[0.64, 0.71]
Anger	HCV-PRO Item: I felt restless or on edge	952	0.62	[0.58, 0.67]
	MSAS feeling irritable score	943	0.58	[0.54, 0.63]
Anxiety	HCV-PRO Item: Having Hepatitis C was very stressful to me	948	0.48	[0.43, 0.53]
	MSAS feeling nervous score	941	0.56	[0.51, 0.61]
	MSAS worrying score	939	0.61	[0.57, 0.66]
Belly pain	HCV-PRO Item: I felt bothered by pain or physical discomfort	950	0.48	[0.42, 0.53]
	MSAS pain score	951	0.39	[0.34, 0.45]
Diarrhea	MSAS diarrhea score	937	0.51	[0.46, 0.57]
Nausea/	MSAS nausea score	939	0.64	[0.60, 0.69]
Vomiting	MSAS vomiting score	940	0.35	[0.29, 0.40]

[a]The number of participants with a non-missing value.
[b]Absolute value of Spearman's rank correlation coefficient between each measure and a similar construct.
[c]95% confidence interval (CI) for Spearman's rank correlation coefficient.
[d]Higher values on the HIT-6 and PROMIS scores indicate more severe symptoms.
[e]MSAS, memorial symptom assessment scale; MSAS score, average of the three attributes: severity, frequency, and interference.

Floor effects occurred in this study because many patients did not have the symptom: 59% reported no diarrhea, 52% reported no nausea/vomiting, 50% reported no abdominal pain, and 37% reported no pain. These floor effects are not surprising given that half of the sample did not have cirrhosis and specific symptoms (nausea/vomiting, abdominal pain) may not be present in patients in the absence of advanced cirrhosis. Viewed from another perspective, these data indicate that 41-63% of patients actually *do experience* mild to severe symptoms, representing a substantial number of patients. One caveat for investigators who need to capture subtle differences in symptoms in patients with mild disease, these PROs may lack sufficient sensitivity. Future studies could evaluate if additional items from the PROMIS item banks would increase variability in scores of patients with mild disease. Additionally, a variety of statistical analysis methods designed for measures that have both a binary component (presence or absence) and continuous component may be useful; for example, mixture models such as a zero-inflated log-normal model can be applied for improved interpretability and precision.

The structural validity of the PROMIS surveys and the HIT-6 supported unidimensionality of most surveys, although not all. Strong support for the unidimensionality for the HIT-6, and PROMIS Depression, Cognitive Concern, Anger, Anxiety and Belly Pain scales was observed. Some surveys tap multidimensional constructs (eg, PROMIS Nausea/Vomiting), while other surveys may include an item that does not hang together well with the other items (eg, PROMIS Fatigue).

The assessment of convergent validity in this study focused on the strength of correlation between the HIT-6 and PROMIS surveys with theoretically-similar items from two previously validated legacy measures: the HCV-PRO and the MSAS. Almost all surveys demonstrated moderate (0.40-0.70) to strong (> 0.70) correlations with items tapping theoretically-similar constructs from the HCV-PRO and MSAS, thus providing evidence that these instruments measure the purported construct. These findings are consistent with another recent psychometric study conducted in cirrhotic patients that documented satisfactory convergent and discriminant validity for the same PROMIS measures.[20] A few weak to moderate correlations

TABLE 5 Known-groups validity: PROs scores by liver disease status

Measure	Group A		Group B	
	No cirrhosis N = 456[a] Mean (SD)	Cirrhosis (MELD 6-11) N = 334[b] Mean (SD)	Cirrhosis (MELD > 12) N = 37[c] Mean (SD)	Group A-Group B[d] Mean difference (95% CI)
HIT-6 score[e]	46 (11)	45 (10)	47 (11)	−1.5 [−4.9, 2.0]
PROMIS T-scores[e]				
Fatigue	52 (11)	52 (11)	54 (10)	−1.9 [−5.6, 1.7]
Pain interference	53 (12)	53 (11)	56 (10)	−3.0 [−6.7, 0.7]
Sleep disturbance	52 (11)	52 (12)	54 (10)	−2.4 [−6.2, 1.4]
Depression	50 (11)	49 (11)	54 (9)	−4.1 [−7.6, −0.6]
Cognitive concerns	33 (9)	33 (9)	35 (8)	−1.8 [−4.9, 1.2]
Anger	49 (12)	48 (11)	51 (10)	−2.9 [−6.7, 0.8]
Anxiety	51 (11)	49 (10)	54 (10)	−3.5 [−6.9, 0.0]
Belly pain	37 (13)	37 (13)	40 (14)	−3.3 [−7.7, 1.1]
Diarrhea	36 (9)	35 (8)	39 (12)	−3.8 [−7.7, 0.2]
Nausea/Vomiting	42 (9)	42 (9)	46 (9)	−3.9 [−6.8, −1.0]
HCV-PRO[f]	71 (23)	71 (23)	63 (23)	7.7 [0.0, 15.3]

SD, standard deviation; CI, confidence interval.
[a]Sample size ranges from 415 to 456.
[b]Sample size ranges from 312 to 334.
[c]Sample size ranges from 36 to 37.
[d]Group A, No cirrhosis or mild cirrhosis (MELD 6-12); Group B, Advanced Cirrhosis(MELD > 12).
[e]Higher values indicate more severe symptoms.
[f]Higher values indicate higher functional well-being.

TABLE 6 Known groups validity: Correlations between PRO scores and Number of Comorbidities

Measure	Number of comorbidities Spearman's correlation[a] (95% CI)	
HIT-6 score[b]	0.31	[0.25, 0.37]
PROMIS[b]		
Fatigue	0.39	[0.34, 0.44]
Pain interference	0.53	[0.48, 0.57]
Sleep disturbance	0.41	[0.36, 0.46]
Depression	0.30	[0.24, 0.36]
Cognitive concerns	0.33	[0.27, 0.38]
Anger	0.29	[0.23, 0.35]
Anxiety	0.30	[0.24, 0.36]
Belly pain	0.34	[0.28, 0.40]
Diarrhea	0.20	[0.14, 0.27]
Nausea/Vomiting	0.29	[0.23, 0.35]
HCV-PRO[c]	−0.48	[−0.53, −0.43]

CI, confidence interval; Sample size ranges from 887 to 961.
[a]Spearman's rank correlation coefficient (Rho).
[b]Higher values of HIT-6 and PROMIS T-scores indicate more severe symptoms.
[c]Higher values indicate higher functional well-being.

were observed for the HIT-6 and PROMIS Belly Pain, which may be because the general "pain" item we used to evaluate headache and abdominal pain was too nonspecific. The absence of other legacy measures or items impeded our ability to adequately examine convergent validity for a couple of scales, representing one limitation of this study.

Finally, we explored known-groups validity in patients with and without cirrhosis, with varying levels of liver disease, and with a range of comorbid medical conditions. Differences in PRO scores were negligible between patients with no cirrhosis and mild cirrhosis, however, patients with advanced cirrhosis (MELD > 12) had worse symptom scores than patients with no or minimal cirrhosis, by 1.9 to 4.1 point differences. In particular, it is reasonable to expect that gastrointestinal symptoms (nausea, abdominal pain, and diarrhea) and pain would be worse in patients with advanced cirrhosis due to decompensation compared to those with minimal or no cirrhosis (point differences ≥ 3). This finding is consistent with Bajaj et al who found differences between patients with compensated and decompensated cirrhosis on several of the same PROMIS measures.[20] The lack of differentiation in this study between patients with no cirrhosis and those with mild cirrhosis may be because the former group had stage 2-3 fibrosis, while the latter group had early stage 4 fibrosis; the small difference in liver disease stage may explain the similarity in symptoms in these two groups. Unfortunately, liver disease staging data are unavailable to substantiate this hypothesis, although we know that at the time of enrollment into this study, many payers only covered DAA therapy in patients with advanced fibrosis or cirrhosis. Future studies with the PROMIS measures could determine if other items from the larger item banks could provide better differentiation between these fibrosis levels. In a second analysis of know-groups validity, we found evidence of an association between the

PROMIS measures and patients' number of comorbidities, such that patients with more comorbidities reported worse symptoms.[51] These analyses support known-groups validity by differentiating between patients with varying degrees of comorbidities and liver disease.

The PROMIS measures, HIT-6, and HCV-PRO complement other PROs that have been used to study the subjective experiences of patients with HCV. A recent literature review identified 22 concepts, almost all *precise symptoms*, described as important to patients that have been inadequately examined in HCV clinical trials.[5] Of 18 PRO measures utilized in previous studies, only four were satisfactorily developed and/or validated in patients with HCV, and three are quality of life instruments. As for *specific symptoms*, only the Fatigue Severity Scale had previously been validated in patients with HCV.[53] Investigators can now use the current PROMIS and HIT-6 measures to evaluate several additional *precise symptoms* that are important to patients yet completely under-studied.

Our study has limitations that point to future directions for improvements in PRO measurement in the field of HCV. All patients were English speaking. Survey data were collected via various data collection modalities and our limited data preclude evaluation of social desirability or effects of data collection modality. Other legacy instruments were not utilized in the parent study due to concern about patient burden and could be an area of future investigation to explore convergent and discriminant validity. Most of the PRO measures showed floor effects, ostensibly because many patients were not experiencing a specific symptom. There is a chance that these measures may not be sensitive enough to capture very mild side effects of DAA therapy or subtle symptom improvement associated with viral cure. However, investigators could seek to improve the surveys by identifying additional items from the large PROMIS item banks to create greater variability at the healthier end of the spectrum. Likewise, future studies might identify other items from the PROMIS banks that help surveys better distinguish between patients with no cirrhosis and early cirrhosis. There may be other symptoms of HCV that were not evaluated in this study (eg, itching). These PROs were chosen to capture a range of burdensome symptoms; however it would be interesting to explore *positive patient experiences* after being cured of HCV, that are not captured by these surveys (eg, positive growth, psychological freedom, happiness) that are not just the absence of symptoms, but altogether different human experiences. Future studies might evaluate the responsiveness of the PRO measures over time and in other patient subpopulations who might have worse symptomatology (eg, patients with certain sociodemographic characteristics or those with alcohol, drug use and mental health issues).[54-58]

To conclude, the current study found the psychometric properties of 10 PROMIS short forms, the HIT-6, and the HCV-PRO performed satisfactorily in patients with chronic HCV. Slight modifications to the surveys may enhance performance among healthy subjects and better distinguish between patients with and without cirrhosis. Our study indicates that these symptom PROs can be used in clinical research investigations of similar patients and clinical settings.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contributions of the following people: Alan Franciscus of www.HCVAdvocate.org, Anquenette Sloan, Summer Wadsworth, Scott Kixmiller, Larry Huston, Finton Brown (UNC Patient Engagement Group); Shani Alston, Virginia Sharpless, Ken Berquist, Herleesha Anderson, Courtenay Pierce, Maxwell Drew Froman, Jenn Barr, Bryonna Jackson, Jane Giang, Stephen Zacks, Jama Darling, A. Sidney Barrit IV, Vicki Metheny, Scott Elliot, Dawn Harrison, Danielle Cardona (University of North Carolina); K. Rajender Reddy, Kelly Borges (University of Pennsylvania); Jessica Mason, Sophie Dessie (Virginia Commonwealth University); Lelani Fetrow, Vicki Shah (Rush University); LaKita May (Saint Louis University); Marechristi Candido, Sophia Zaragoza, Rebecca Hluahanich, Kathleen Haight (University of California, Davis); Elizabeth Wu (University of Michigan); AnnMarie Liapakis, MD, Kristine Drozd, Carol Eggers, Hong Chau and Claudia Bertuccio (Yale University) David Nelson, Patrick Horne (University of Florida); William King, Jean E. Nichols, Heather Goldstein, Megan Marles (Wilmington Gastroenterology Associates); William R. Harlan III, Roberta Golden (Asheville Gastroenterology Associates).

Declarations of personal interests: Donna M. Evon has received research funding from Gilead. Anna S. Lok has received research support from AbbVie, BMS, Gilead, and Merck; and served as an advisor for Gilead. Adrian M. Di Bisceglie has received research support from AbbVie, BMS and Gilead and has served on advisory boards for AbbVie, BMS and Merck. Richard K. Sterling has received research support from AbbVie, BMS, Gilead, Merck, and Roche and served as a consultant for Merck, Bayer, Salix, AbbVie, Gilead, Janssen, ViiV, Baxter, and Pfizer. Nancy Reau has received research funding (paid to Rush) from AbbVie and Intercept and has served as a consultant for AbbVie, Gilead, Merck, and BMS. Souvik Sarkar served on a Gilead Advisory Board. Joseph K. Lim has received research contracts (paid to Yale) from BMS, Genfit, Gilead, Hologic, Intercept, and Prometheus; and has served as consultant to AbbVie, BMS, and Gilead. Michael Fried has received research funding from and served as a consultant for AbbVie, BMS, Gilead, Merck, and TARGET PharmaSolutions. Stock in TARGET PharmaSolutions is held in an independently managed trust. Marina Serper, Jipcy Amador, Carol E. Golin, Paul Stewart, and Bryce Reeve declare that they have no conflict of interests to disclose.

Declaration of funding interests: This study was funded by a Patient-Centered Outcomes Research Institute (PCORI) Award to Donna M. Evon Award (CER-1408-20660). Additional support for this study (data management) came from the Center for Gastrointestinal Biology and Disease (CGIBD; P30-DK34987). Additional support for Dr. Golin's salary was partially supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (K24-HD06920) and by the University of North Carolina Center for AIDS Research (CFAR) (P30 AI-50410). The statements presented in this article are solely the responsibility of the authors and do not necessarily represent the

views of PCORI, its Board of Governors or Methodology Committee.

AUTHORSHIP

Guarantor of article: Donna M. Evon.

Author contributions: Donna M. Evon, Paul Stewart, Bryce B. Reeve, Carol E. Golin, and Michael W. Fried designed the study. Anna S. Lok, Adrian M. Di Bisceglie, Richard K. Sterling, Nancy Reau, Marina Serper, Souvik Sarkar, Joseph K. Lim contributed to data acquisition. Jipcy Amador, Paul Stewart, and Bryce B. Reeve performed the data analyses. Donna M. Evon, Paul Stewart, Jipcy Amador, Bryce B. Reeve, Anna S. Lok, Adrian M. Di Bisceglie, Richard K. Sterling, Nancy Reau, Marina Serper, Souvik Sarkar, Joseph K. Lim, Carol E. Golin, and Michael W. Fried contributed to the interpretation of the data, drafted the paper or revised it critically. All authors listed approved the final version of the article, including the authorship list.

ORCID

D. M. Evon http://orcid.org/0000-0002-1414-1846
M. Serper http://orcid.org/0000-0003-4899-2160
S. Sarkar http://orcid.org/0000-0002-9358-4257
J. K. Lim http://orcid.org/0000-0003-1126-8128

REFERENCES

1. Lang CA, Conrad S, Garrett L, et al. Symptom prevalence and clustering of symptoms in people living with chronic hepatitis C infection. *J Pain Symptom Manage*. 2006;31:335-344.
2. Younossi Z, Kallman J, Kincaid J. The effects of HCV infection and management on health-related quality of life. *Hepatology*. 2007;45:806-816.
3. Carlson MD, Hilsabeck RC, Barakat F, Perry W. Role of sleep disturbance in chronic hepatitis C infection. *Curr Hepat Rep*. 2010;9:25-29.
4. Lovejoy TI, Dobscha SK, Cavanagh R, Turk DC, Morasco BJ. Chronic pain treatment and health service utilization of veterans with hepatitis C virus infection. *Pain Med*. 2012;13:1407-1416.
5. Kleinman L, Mannix S, Yuan Y, Kummer S, L'Italien G, Revicki D. Review of patient-reported outcome measures in chronic hepatitis C. *Health Qual Life Outcomes*. 2012;10:92. https://doi.org/10.1186/1477-7525-10-92.:92-10
6. Ware JE Jr, Bayliss MS, Mannocchia M, Davis GL. Health-related quality of life in chronic hepatitis C: impact of disease and treatment response. The Interventional Therapy Group. *Hepatology*. 1999;30:550-555.
7. Spiegel BM, Younossi ZM, Hays RD, Revicki D, Robbins S, Kanwal F. Impact of hepatitis C on health related quality of life: a systematic review and quantitative assessment. *Hepatology*. 2005;41:790-800.
8. Afdhal N, Reddy KR, Nelson DR, et al. Ledipasvir and sofosbuvir for previously treated HCV genotype 1 infection. *N Engl J Med*. 2014;370:1483-1493.
9. Zeuzem S, Ghalib R, Reddy KR, et al. Grazoprevir-Elbasvir combination therapy for treatment-naive cirrhotic and noncirrhotic patients with chronic hepatitis C virus genotype 1, 4, or 6 infection: a randomized trial. *Ann Intern Med*. 2015;163:1-13.
10. Ferenci P, Bernstein D, Lalezari J, et al. ABT-450/r-ombitasvir and dasabuvir with or without ribavirin for HCV. *N Engl J Med*. 2014;370:1983-1992.
11. Everson GT, Towner WJ, Davis MN, et al. Sofosbuvir with Velpatasvir in treatment-naive noncirrhotic patients with genotype 1 to 6 hepatitis C virus infection: a randomized trial. *Ann Intern Med*. 2015;163:818-826.
12. Younossi ZM, Stepanova M, Henry L, et al. Effects of sofosbuvir-based treatment, with and without interferon, on outcome and productivity of patients with chronic hepatitis C. *Clin Gastroenterol Hepatol*. 2014 Aug;12:1349-1359.
13. Younossi ZM, Stepanova M, Zeuzem S, et al. Patient-reported outcomes assessment in chronic hepatitis C treated with sofosbuvir and ribavirin: the VALENCE study. *J Hepatol*. 2014;61:228-234.
14. Younossi Z, Henry L. Systematic review: patient-reported outcomes in chronic hepatitis C - the impact of liver disease and new treatment regimens. *Aliment Pharmacol Ther*. 2015;41:497-520.
15. Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER) and the Center for Devices and Radiological Health (CDRH). Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. December 2009. https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf. Accessed March 6, 2017.
16. Younossi ZM, Guyatt G, Kiwi M, Boparai N, King D. Development of a disease specific questionnaire to measure health related quality of life in patients with chronic liver disease. *Gut*. 1999;45:295-300.
17. Gralnek IM, Hays RD, Kilbourne A, et al. Development and evaluation of the Liver Disease Quality of Life instrument in persons with advanced, chronic liver disease–the LDQOL 1.0. *Am J Gastroenterol*. 2000;95:3552-3565.
18. 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;63:1179-1194.
19. 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.
20. Bajaj JS, Thacker LR, Wade JB, et al. PROMIS computerised adaptive tests are dynamic instruments to measure health-related quality of life in patients with cirrhosis. *Aliment Pharmacol Ther*. 2011;34:1123-1132.
21. Elliott C, Frith J, Pairman J, Jones DE, Newton JL. Reduction in functional ability is significant postliver transplantation compared with matched liver disease and community dwelling controls. *Transpl Int*. 2011;24:588-595.
22. Kosinski M, Bayliss MS, Bjorner JB, et al. A six-item short-form survey for measuring headache impact: the HIT-6. *Qual Life Res*. 2003;12:963-974.
23. Evon DM, Golin CE, Stewart P, et al. Patient engagement and study design of PROP UP: a multi-site patient-centered prospective observational study of patients undergoing hepatitis C treatment. *Contemporary clinical trials*. 2017;57:58-68.
24. Lee AC, Driban JB, Price LL, Harvey WF, Rodday AM, Wang C. Responsiveness and minimally important differences for 4 patient-reported outcomes measurement information system short forms: physical function, pain interference, depression, and anxiety in knee osteoarthritis. *J Pain*. 2017 Sep;18:1096-1110.
25. Khanna D, Hays RD, Shreiner AB, et al. Responsiveness to change and minimally important differences of the patient-reported outcomes measurement information system gastrointestinal symptoms scales. *Dig Dis Sci*. 2017;62:1186-1192.
26. Amtmann D, Kim J, Chung H, Askew RL, Park R, Cook KF. Minimally important differences for patient reported outcomes measurement

information system pain interference for individuals with back pain. *Journal of pain research*. 2016;9:251-255.
27. Kroenke K, Yu Z, Wu J, Kean J, Monahan PO. Operating characteristics of PROMIS four-item depression and anxiety scales in primary care patients with chronic pain. *Pain Med*. 2014;15:1892-1901.
28. Portenoy RK, Thaler HT, Kornblith AB, et al. The memorial symptom assessment scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. *Eur J Cancer*. 1994;30:1326-1336.
29. Chang VT, Hwang SS, Thaler HT, Kasimis BS, Portenoy RK. Memorial symptom assessment scale. *Expert Rev Pharmacoecon Outcomes Res*. 2004;4:171-178.
30. Anderson RT, Baran RW, Dietz B, Kallwitz E, Erickson P, Revicki DA. Development and initial psychometric evaluation of the hepatitis C virus-patient-reported outcomes (HCV-PRO) instrument. *Qual Life Res*. 2014;23:561-570.
31. Anderson RT, Baran RW, Erickson P, Revicki DA, Dietz B, Gooch K. Psychometric evaluation of the hepatitis C virus patient-reported outcomes (HCV-PRO) instrument: validity, responsiveness, and identification of the minimally important difference in a phase 2 clinical trial. *Qual Life Res*. 2014;23:877-886.
32. Wai CT, Greenson JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology*. 2003;38:518-526.
33. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology*. 2006;43:1317-1325.
34. United Network for Organ Sharing (UNOS) MELD/PELD Calculator Documentation. January 28, 2009. https://www.unos.org/wp-content/uploads/unos/MELD_PELD_Calculator_Documentation.pdf. Accessed July 5, 2017.
35. Nunnally JC, Bernstein IH. *Pschometric Theory*, 3rd edn. New York: McGraw-Hill; 1994.
36. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*. 2013;22:1889-1905.
37. Iacobucci D, Duhachek A. Advancing Alpha: measuring Reliability With Confidence. *Journal of Consumer Psychology*. 2003;13:478-487.
38. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34-42.
39. Samejima F. Graded response model. In: van der Linden WJ, Hambleton RK, eds. *Handbook of modern item response theory*. New York NY: Springer; 1997:85-100.
40. Steiger JH, Lind J. *Statistically-based tests for the number of common factors 1980*. Located at: Iowa City, IA: Paper presented at the Annual Spring Meeting of the Psychometric Society; 1980.
41. Browne MW, Cudeck R, Bollen KA, Long JS. *Alternative ways of assessing model fit*, vol. 154. Newbury Park, CA: Sage Publications; 1993.
42. Kang T, Chen TT. Performance of the generalized S-X2 item fit indexfor the graded response model. *Asia Pac Educ Rev*. 2011;12:89-96.
43. Chen WH, Thissen D. Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*. 1997;22:265-289.
44. Cai L, Yang JS, Hansen M. Generalized full-information item bifactor analysis. *Psychol Methods*. 2011;16:221-248.
45. Bentler PM. Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*. 2009;74:137-143.
46. Reise SP, Moore TM, Haviland MG. Bifactor models and rotations: exploring the extent to which multidimensional data yield univocal scale scores. *J Pers Assess*. 2010;92:544-559.
47. Dancey C, Reidy J. *Statistics without Maths for Psychology: using SPSS Windows*, 3rd edn. Harow, England: Prentice Hall; 2004.
48. Younossi ZM, Stepanova M, Afdhal N, et al. Improvement of health-related quality of life and work productivity in chronic hepatitis C patients with early and advanced fibrosis treated with ledipasvir and sofosbuvir. *J Hepatol*. 2015;63:337-345.
49. Bonkovsky HL, Snow KK, Malet PF, et al. Health-related quality of life in patients with chronic hepatitis C and advanced fibrosis. *J Hepatol*. 2007;46:420-431.
50. Snow KK, Bonkovsky HL, Fontana RJ, et al. Changes in quality of life and sexual health are associated with low-dose peginterferon therapy and disease progression in patients with chronic hepatitis C. *Aliment Pharmacol Ther*. 2010;31:719-734.
51. Rothrock NE, Hays RD, Spritzer K, Yount SE, Riley W, Cella D. Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol*. 2010;63:1195-1204.
52. Rubin DB. Inference and Missing Data. *Biometrika*. 1976;63:581-592.
53. Rosa K, Fu M, Gilles L, et al. Validation of the Fatigue Severity Scale in chronic hepatitis C. *Health Qual Life Outcomes*. 2014;12:90.
54. Foster GR, Goldin RD, Thomas HC. Chronic hepatitis C virus infection causes a significant reduction in quality of life in the absence of cirrhosis. *Hepatology*. 1998;27:209-212.
55. Hauser W, Zimmer C, Schiedermaier P, Grandt D. Biopsychosocial predictors of health-related quality of life in patients with chronic hepatitis C. *Psychosom Med*. 2004;66:954-958.
56. Rowan PJ, Al-Jurdi R, Tavakoli-Tabasi S, Kunik ME, Satrom SL, El-Serag HB. Physical and psychosocial contributors to quality of life in veterans with hepatitis C not on antiviral therapy. *J Clin Gastroenterol*. 2005;39:731-736.
57. Stepanova M, Younossi Z. Letter: new treatments for hepatitis C have implications for quality of life in people who inject drug–authors' reply. *Aliment Pharmacol Ther*. 2016;43:841-842.
58. Higgs P, Wright C, Hellard M. Letter: new treatments for hepatitis C have implications for quality of life in people who inject drugs. *Aliment Pharmacol Ther*. 2016;43:840-841.

How to cite this article: Evon DM, Amador J, Stewart P, et al. Psychometric properties of the PROMIS short form measures in a U.S. cohort of 961 patients with chronic hepatitis C prescribed direct acting antiviral therapy. *Aliment Pharmacol Ther*. 2018;47:1001–1011. https://doi.org/10.1111/apt.14531