Received: 26 July 2022 | First decision: 18 August 2022 | Accepted: 9 September 2022

DOI: 10.1111/apt.17230

AP&T Alimentary Pharmacology & Therapeutics WILEY

Validation of the Achalasia Patient-Reported Outcomes Questionnaire

John E. Pandolfino[1] | Dustin A. Carlson[1] | Josie McGarva[1] | Peter J. Kahrilas[1] | Michael Vaezi[2] | David Katzka[3] | Tiffany H. Taft[1]

[1]Division of Gastroenterology and Hepatology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

[2]Division of Gastroenterology, Vanderbildt University Medical Center, Nashville, Tennessee, USA

[3]Division of Gastroenterology, Mayo Clinic, Rochester, Minnesota, USA

Correspondence
Tiffany H. Taft, Division of Gastroenterology and Hepatology, Northwestern University Feinberg School of Medicine, 676 N. Saint Clair Street, Suite 1400, Chicago, IL 60611, USA.
Email: ttaft@northwestern.edu

Funding information
National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Number: 1P01DK117824-01

Summary

Background: Achalasia is a debilitating major motor disorder of the oesophagus. Hypervigilance and symptom-specific anxiety substantially impact dysphagia symptom reporting, and quality of life is a critical patient outcome. Earlier achalasia symptom scales did not consider these constructs in their psychometric development.

Aim: To develop a new symptom measure, the Achalasia Patient-Reported Outcomes (APRO) Questionnaire

Methods: Four gastroenterologists with achalasia expertise generated preliminary items. Patients reviewed items via cognitive interviews. Patients undergoing high-resolution manometry completed the APRO with Oesophageal Hypervigilance and Anxiety Scale, Northwestern Oesophageal Quality of Life Scale, and three measures of reflux and dysphagia. Full APRO psychometric assessment (reliability, validity, factor structure) was done. Cluster analysis evaluated APRO + symptom-anxiety/hypervigilance patient phenotypes.

Results: We included 961 patients with normal motility and 296 with achalasia. The APRO yielded three subscales: dysphagia, reflux, chest pain with two items for weight change and diet modifications. Reliability and validity were excellent. Twenty-five percent of achalasia patients may have high levels of anxiety/hypervigilance despite low symptoms, while 8% may report severe symptoms with low anxiety/hypervigilance. The APRO significantly predicted quality of life, but less cognitive-affective processes.

Conclusions: The APRO is a reliable and valid measure of achalasia symptoms that addresses the limitations of existing questionnaires. Symptom anxiety and hypervigilance moderate the relationship between APRO and quality of life; 33% of patients with achalasia exhibit concerning patterns in symptom severity, anxiety and hypervigilance that may contribute to poorer outcomes.

The Handling Editor for this article was Dr Colin Howden, and it was accepted for publication after full peer-review.

1 | INTRODUCTION

Patient subjective experience and self-report of their symptoms is fundamental to managing chronic digestive diseases. This necessitates access to reliable and valid questionnaires easily administered in clinical and research settings. Prior to 2009 and the publication of updated guidelines from the U.S. Food and Drug Administration (FDA), a variety of psychometric methods were used to create symptom scales. In oesophageal disease, there are multiple questionnaires of varying length and complexity, including the more widely used gastro-oesophageal reflux questionnaire[1] and Eckardt score,[2,3] as well as the NIH-PROMIS upper gastrointestinal symptom scales[4] and Eosinophilic Esophagitis Symptom Activity Index.[5]

Our understanding of how patients interpret physiological sensations and report them as symptoms continues to evolve. Subjective experience of oesophageal disease is multifactorial and often only modestly correlated with objective assessment of disease severity.[6–8] Gauging both the frequency and the intensity of each symptom is a newer method incorporated in certain oesophageal measures such as the Brief Oesophageal Dysphagia Questionnaire (BEDQ)[9] and the EEsAI,[5] providing a clearer understanding of symptom experiences. In conditions where reflux of stomach or oesophageal contents is involved (i.e. gastro-oesophageal reflux disease or achalasia), evaluating night-time symptoms when in a supine position is also important not only for sleep disturbance but also risk for aspiration. Weight change in scales such as the Eckardt score has historically only evaluated weight reductions rather than considering some achalasia patients gain weight despite their oesophageal symptoms.[10]

Two emerging and consistently important cognitive-affective processes in oesophageal diseases are hypervigilance to bodily sensations and anxiety about symptoms of dysphagia, reflux or abdominal pain.[11–14] Across oesophageal conditions, these constructs are emerging as more comprehensive predictors of symptom severity reporting than physiological markers of disease activity. As such, it behoves those who develop symptom questionnaires to consider how hypervigilance and anxiety may influence instrument scores and moderate relationships with other constructs like health-related quality of life (HRQoL).[11]

In achalasia, the Eckardt score,[3] developed in 1992, is considered the gold standard assessment for symptom severity, including for clinical trial outcomes; it has an established cut-off score of 3 or higher used for 'active disease'. Our 2018 evaluation of the Eckardt score's reliability and construct validity found the questionnaire met 'fair' psychometric performance standards with weight loss and chest pain potentially decreasing its validity.[2] The score was also heavily influenced by dysphagia severity. Achalasia is a chronic motility disorder of considerable morbidity with the main treatments being surgical interventions.[15] Considering the severity of the disease, the contribution of both sensory and motor abnormalities and invasiveness of intervention, it is imperative to accurately measure outcomes.

Based on these discoveries, we sought to create a new measure of achalasia symptom severity, the Achalasia Patient-Reported Outcomes (APRO) Questionnaire, which leverages the strengths of the Eckardt score while attempting to enhance its utility that results from some of its limitations. We also aimed to create the APRO while simultaneously considering hypervigilance and anxiety in its scoring and relationships with HRQoL.

2 | METHODS

English-speaking patients ages 18–80 presenting to an outpatient gastroenterology laboratory between January 2017 and March 2021 for routine high-resolution manometry testing were recruited. Those with a Chicago Classification v3.0 or v4.0 (dependent on date of testing) indicating any type of achalasia (I, II, III; esophagogastric junction outflow obstruction [EGJOO] patients were not included) were selected along with patients with normal oesophageal motility presenting with symptoms of dysphagia, chest pain, and/or reflux as a comparator group. Patients with treated achalasia (previous pneumatic dilation, laparoscopic Heller's myotomy or PerOral Endoscopic Myotomy) were identified to control for this status in statistical analyses. Pre-post assessment of the APRO was done with a subset of treated patients with achalasia that completed assessment before and after intervention. A sample of 39 healthy volunteers also completed the APRO to further support validation. Participants filled out a series of paper-based questionnaires after being checked in, but prior to their procedure.

2.1 | Questionnaires

2.1.1 | APRO scale

The APRO is an 11-item questionnaire developed using psychometric protocols as outlined by the U.S. FDA guidelines for questionnaire development. Specifically, four gastroenterologists with expertise in managing oesophageal disease (Authors: J.E.P., P.J.K., M.V., D.K.) generated the initial items based on the most reported symptoms of achalasia: dysphagia, chest pain and regurgitation. In addition to the severity of these symptoms, items that measure the frequency were also created and how symptoms interrupt sleep. Two additional items assessing dietary changes in response to oesophageal symptoms (Yes = 1, No = 0) and weight change over the last 30 days were included. Questions were rated on Likert scales as follows: Frequency items: 0 days (Coded 0), 1 day (1), 2–3 days (2), 4–6 days (3), Every Day (4), Constant (5); Severity items: None (0), Mild (1), Moderate (2), Severe (3), Very Severe (4), Cannot Swallow (5). Weight change was coded from −1 (Gained weight) to 4 (Lost more than 15 pounds). The initial questionnaire was given to a group of achalasia patients who underwent cognitive interviews

regarding each item's wording and their understanding of the construct being measured. The study team members (Authors: J.E.P., T.H.T., D.A.C.) incorporated minor changes and reached consensus on the final content of the first version of the APRO for validation in the present study.

2.1.2 | Eckardt score

The Eckardt score (ESS)[2,3] is a 4-item measure of achalasia severity. Items include weight loss, chest pain, regurgitation and dysphagia graded on a score of 0–3. The maximum score is 12 and scoring greater than 3 suggests active achalasia.

2.1.3 | Northwestern Oesophageal Quality of Life scale

The Northwestern Oesophageal Quality of Life (NEQOL)[16] is a 14-item measure of oesophageal-disease-specific HRQoL. Items evaluate social, emotional, financial, eating and sleep impacts of oesophageal symptoms. Scores range from 0 to 60 with higher scores indicating better HRQoL. There is no established cut-point for a 'high' versus 'low' score.

2.1.4 | Oesophageal Hypervigilance and Anxiety Scale

The Oesophageal Hypervigilance and Anxiety Scale (EHAS)[17] is a 15-item measure of hypervigilance to oesophageal sensations and anxiety about the presence or possibility of symptoms over the last 30 days. Scores range from 0 to 60 for the total scale, 0–36 for the anxiety subscale and 0–24 for hypervigilance. A score greater than 23 on the total scale is considered elevated hypervigilance and anxiety. There are no established cut points for each subscale.

2.1.5 | Gastro-Oesophageal Reflux Questionnaire

The Gastro-Oesophageal Reflux Questionnaire (GERDQ)[1] is a 6-item measure of reflux severity. Items include four positively scored items: heartburn, regurgitation, sleep disruption from symptoms, and increases in medication to control GERD, and two negatively scored items: epigastric pain and nausea. Scores greater than 8 are indicative of GERD.

2.1.6 | Brief Oesophageal Dysphagia Questionnaire

The Brief Oesophageal Dysphagia Questionnaire (BEDQ)[9] is a 10-item measure of oesophageal dysphagia with an additional assessment of food impactions. It measures both frequency and difficulty with swallowing solid foods, soft foods and liquids over the past 30 days. Two food impaction items, not included in the total score, assess the number of impactions lasting more than 30 min but clearing on its own, or those requiring an emergency room visit in the past year. Scores range from 0 to 40 with higher scores equating to worse dysphagia. A score of 6 or higher is indicative of significant dysphagia.

2.2 | Demographic, clinical and physiological data

In addition to age and gender of the participant, the following information was collected:

2.2.1 | Primary indication

The primary indication for high-resolution manometry was recorded and included: dysphagia, reflux, chest pain, follow-up, pre-operative, and other.

2.2.2 | Proton pump inhibitor use

Current use of proton pump inhibitor coded Yes = 1, No = 0.

2.2.3 | Chicago Classification v3.0 or v4.0

Patients were placed into four categories per Chicago Classification v3.0 or v4.0[18,19] findings on high-resolution manometry, depending on the date of the testing: Achalasia Type I, Achalasia Type II, Achalasia Type III, Normal Motility. Patients with EGJOO were excluded.

2.3 | Statistical analyses

Data were exported to SPSS v27 for Macintosh for analyses. Tests for normal distribution were performed on all variables (skewness and kurtosis ±2.0) and found no need for non-parametric tests. Descriptive statistics for the sample are presented as mean (standard deviation [SD]) for continuous variables and percentage (frequency) for categorical. Data visualisation was done using R Statistical Software (v4.1.2; R Core Team 2021) package ggplot2.

2.3.1 | Psychometric assessments of APRO

Internal consistency for the APRO was evaluated via Cronbach alpha, with a minimum acceptable criterion of 0.70; split-half reliability was measured using the same criterion. To establish the factor structure of the APRO a principal components factor analysis was performed. Items 1–9 of the scale were entered, and the component matrix was

rotated using a Promax method. Subscales were established using a Scree Plot and Eigenvalues greater than 1. Inter-item correlations identified highly related APRO items with Pearson's coefficient >0.75.

2.3.2 | Validity of the APRO

Total scores and applicable subscales scores were calculated for each of the questionnaires. For the EHAS: Anxiety and EHAS: Hypervigilance subscales, a median split was used to identify 'High' and 'Low' levels of each, and participants were categorised accordingly. Differences between group variables (Gender, COVID pre/post, Chicago Classification Category) were evaluated using independent sample's *t*-tests or one-way analysis of variance. Pearson's correlations were used to measure concurrent validity of the APRO with the Eckardt score, and convergent validity with the BEDQ, GERDQ, EHAS and NEQOL. To account for possible effects of the COVID-19 pandemic, participants were grouped into PRE and POST based on a date of 15 March 2020 and evaluated for any differences on questionnaire scores.

For only Achalasia patients, the mean APRO and Eckardt score of patients with a history of previous foregut surgery was compared to scores of those who were treatment naïve using analysis of covariance to control for the effects of prior surgery. In a subset of patients who underwent foregut surgery during the study period, the change in mean scores pre- and post-treatment was measured using paired samples *t*-tests with Cohen's *d* to evaluate effect size and clinical significance.

To assess how the APRO relates to two important constructs in achalasia management (symptom-specific anxiety/hypervigilance and quality of life) two-step cluster analyses and hierarchical linear regression were employed. First, the cluster analyses assessed APRO scores with categorical EHAS Total, EHAS: Anxiety, and EHAS: Hypervigilance to determine potential patterns in patient responses and phenotyping. Then, two separate hierarchical regression analyses evaluated the strength of the relationship between the APRO and quality of life when controlling for Hypervigilance and Anxiety, and treatment status. Adjusted R square and standardised beta weights are reported for the APRO, each EHAS subscale and surgical history.

3 | RESULTS

In the study period, 4924 patients were seen in the GI clinic. Of these, 3472 completed the questionnaires (70.5%). Of these, a total of 1257 patients met the Chicago Classification v3 or v4 selection criteria and had complete data for all study measures (36.2%). Of the 296 achalasia patients, 95 had a history of foregut surgery and 49 of these patients had pre-post treatment questionnaire data. Sample characteristics are in Table 1. Participants were primarily middle-aged and female, with a 3:1 ratio of normal motility ($N = 961$) to achalasia diagnosis. Forty percent met criteria for Type I achalasia, 36% for Type II and 24% for Type III. Significant differences existed between the two groups across all demographic and clinical variables: achalasia patients were older, more likely to be male, more likely to have had previous foregut surgery, and reported dysphagia as their primary symptom. Nearly three-quarters of achalasia patients modified their diet to manage symptoms, and half reported some degree of weight loss. No COVID-19-related pandemic differences existed for any of the questionnaire data ($p = 0.070$–0.685). The median cut-off for the Hypervigilance subscale was 13 and Symptom Anxiety was 19 in this sample.

3.1 | Psychometric properties of the APRO

The mean (SD) score on the APRO for the entire sample was 12.55 (9.23) and ranged from −1 to 42 (out of a maximum score of 45). Females scored higher on the APRO than males (13.07 [9.05] vs. 11.84 [9.62], $p = 0.027$) as did younger patients. However, the significant finding for difference by age is likely a function of the large sample size as the correlation coefficient is small ($r = -0.121$, $p < 0.001$). Healthy participants scored substantially lower on the APRO (0.342 [0.847], $p < 0.001$) with a range in scores of −1 to 4. The APRO demonstrated good internal consistency (Cronbach $\alpha = 0.84$) and split-half reliability (Guttman statistic = 0.80). Inter-item correlations were small to moderate and did not meet criteria for removal except for items 2 and 8 ($r = 0.778$, $p < 0.001$). However, since these items are measuring frequency and intensity of the same symptom, both were retained.

The principal components factor analysis was well-powered and identified three subscales (Table 2). Factor 1, measuring dysphagia, accounted for 46.7% of the variance, factor 2, measuring reflux, accounted for 16.5%, and factor 3, measuring chest pain, accounted for 11.7%. Items 10 and 11 were removed from the final analysis as these measure behaviour (e.g. diet change) and weight loss. However, when included in the factor analysis, these items loaded onto the dysphagia subscale (0.440, 0.451 respectively). Items 10 and 11 are included in the total APRO score but not subscale scores.

3.2 | Construct validity of the APRO

Pearson's correlation coefficients indicate the APRO has good construct validity based on relationships with the Eckardt score, measures of dysphagia, reflux, quality of life, and symptom-specific anxiety and hypervigilance (Table 3). A larger correlation between the APRO and the measure of dysphagia than with the reflux scale supports convergent validity based on the percentage variance in APRO score for dysphagia versus reflux and chest pain symptoms. The APRO also demonstrated a moderate negative correlation with quality of life and moderate positive correlations with symptom-anxiety and hypervigilance, further supporting its validity.

Construct validity of the APRO is further supported by significant differences in score between achalasia patients with and without a history of foregut surgery (Table 3). Untreated achalasia patients scored twice as high on the APRO as the normal motility

TABLE 1 Demographic and clinical characteristics of study sample.

	Achalasia group (*N* = 296)	Normal motility group (*N* = 961)	*p*
Age (years)	55.56 (16.88)	50.51 (16.19)	<0.001
Gender			
Female	45.7% (162)	63.2% (607)	0.006
Male	45.3% (134)	36.8% (354)	
Achalasia subtype			
I	40.2% (119)	NA	NA
II	36.1% (107)	NA	
III	23.6% (70)	NA	
Previous foregut surgery	32.1% (95)	10.7% (135)	<0.001
Reason for testing			
Chest pain	2.4% (7)	7.2% (69)	<0.001
Dysphagia	74.0% (219)	40.6% (390)	
Follow-up	16.1% (48)	2.8% (27)	
Heartburn/Reflux	2.7% (8)	26.6% (256)	
Pre-operative	0.3% (1)	9.9% (95)	
Other	3.0% (9)	8.4% (81)	
Current PPI use	50.6% (119)	65.8% (495)	<0.001

Abbreviation: NA, not applicable; PPI, proton pump inhibitor.

TABLE 2 Factor structure of APRO.

	Dysphagia	Reflux	Chest pain
Eigenvalue	4.21	1.48	1.05
% Variance	46.74%	16.49%	11.65%
Frequency of trouble swallowing solids	0.876		
Discomfort swallowing solids	0.878		
Discomfort swallowing liquids	0.822		
Frequency of trouble swallowing liquids	0.818		
Frequency of regurgitation at night		0.980	
Sleep disruption due to oesophageal symptoms		0.827	
Frequency of regurgitation after meals		0.672	
Frequency of spasm chest pain			0.883
Frequency of burning chest pain			0.824
Score range	0–20	0–15	0–10

Abbreviation: APRO, Achalasia Patient-Reported Outcomes.

group, on average, and 35% higher than treated achalasia patients. Treated patients continued to score significantly higher on the APRO than those with normal motility ($p = 0.017$). There were no significant differences in mean score on the APRO for patients with normal motility who had a history of foregut surgery (11.82 [9.42]) versus those who did not (10.82 [8.18]; $p = 0.306$).

Lastly, significant differences existed between the achalasia subtypes for APRO score when controlling for surgical history. Those with Type III scored the lowest (14.31, standard error [SE] = 1.12) which was significantly lower than patients with Type II (19.02, SE = 0.92, $p = 001$) and Type I (18.79, SE = 0.89, $p = 0.002$); Type I and Type II did not differ in APRO score when controlling for surgical history ($p = 0.857$).

3.3 | Change in APRO versus Eckardt score with treatment

In 49 patients with APRO and Eckardt score data prior to surgical intervention for achalasia (Figure 1), we found the APRO score reduced significantly (Pre: 19.13 [8.40], Post: 7.95 [7.73], $p < 0.001$; 58% reduction) with a large effect size (Cohen's $d = 1.01$). In these

TABLE 3 Mean between-group differences and correlation coefficients for each measure.

	Treated achalasia group	Untreated achalasia group	Normal motility group	*p*	Pearson's *r* with APRO score
APRO total score	13.08 (9.69)	20.07 (9.47)	10.92 (8.30)	<0.001	-
Eckardt score	4.62 (2.67)	6.47 (2.61)	3.85 (2.42)	<0.001	0.772**
BEDQ score	10.89 (9.38)	18.87 (10.01)	7.93 (8.74)	<0.001	0.781**
GERDQ score	4.03 (3.12)	5.23 (3.02)	5.06 (3.12)	<0.001	0.659**
NEQOL score	35.55 (14.77)	27.05 (14.31)	35.55 (13.62)	<0.001	−0.593**
EHAS: Anxiety	15.99 (9.22)	20.93 (8.19)	17.85 (9.16)	<0.001	0.517**
EHAS: Hypervigilance	11.53 (6.32)	13.39 (5.80)	11.68 (6.09)	0.006	0.437**

Abbreviations: APRO, Achalasia Patient-Reported Outcomes; BEDQ, Brief Oesophageal Dysphagia Questionnaire; EHAS, Oesophageal Hypervigilance and Anxiety Scale; GERDQ, Gastro-oesophageal Reflux Questionnaire; NEQOL, Northwestern Oesophageal Quality of Life.
**$p < 0.001$.

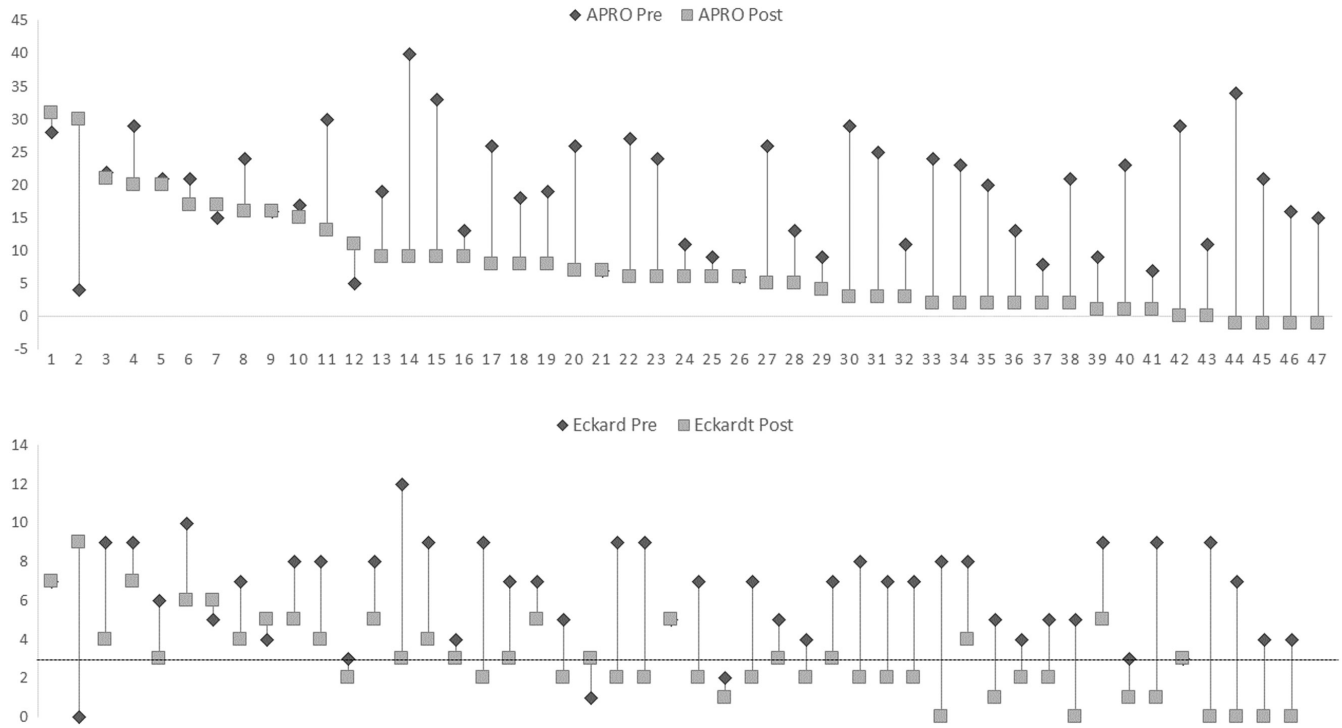

FIGURE 1 Mean change in score for the APRO and Eckardt before and after surgery. APRO, Achalasia Patient-Reported Outcomes Questionnaire

same patients, the Eckardt score reduced less significantly (Pre: 5.25 [2.89], Post: 3.23 [1.99], $p < 0.001$; 38% reduction) with a medium effect size (Cohen's $d = 0.60$). Patients scoring below 3 on Eckardt post-surgery also scored significantly lower on the APRO (4.52 [4.06] vs. 12.29 [7.49], $p < 0.001$).

3.4 | Relationship with symptom anxiety and hypervigilance

Patients with higher APRO scores also endorsed more symptom-specific anxiety and hypervigilance. Further, treatment naïve achalasia patients scored significantly higher for both anxiety and hypervigilance than treated achalasia patients and those with normal motility (Table 3). Patients using diet to manage symptoms reported significantly higher anxiety (20.04 [8.73] vs. 14.39 [9.01], $p < 0.001$) and hypervigilance (13.32 [5.88] vs. 9.86 [5.95], $p < 0.001$).

Each of the cluster analyses had excellent cohesion and separation and yielded four phenotypic groups (Figure 2). For total EHAS score, most patients grouped as would be expected (high APRO, high EHAS or low APRO, low EHAS). However, one-quarter reported elevated EHAS with low symptom severity and 8% had high symptom severity with low EHAS. When examining these

groups by symptom-specific anxiety, 16% reported high symptom severity with low anxiety and 13% had low symptoms with high anxiety. Similarly, 19% reported high symptom severity with low hypervigilance and 17% had low symptoms with high hypervigilance.

3.5 | Relationship with HRQoL

The APRO appears to be a predictor of quality of life when controlling for symptom-specific anxiety and hypervigilance, and previous foregut surgery status, in achalasia patients (Table 4). Standardised beta weights are moderate in size for the APRO, suggesting some direct effect on HRQoL regardless of anxiety and hypervigilance or treatment status. Specifically, the APRO explained between 6% and 13% of the variance in HRQoL when considering these other variables. Simple linear regression finds the APRO predicts 41.4% of the variance in NEQOL score on its own. Patients who endorsed using diet to manage symptoms on the APRO scored significantly lower for HRQoL (30.16 [13.37] vs. 41.51 [12.81], $p<0.001$). The Eckardt score predicted quality of life similarly to the APRO, explaining 5%–10% of the variance in score when controlling for anxiety, hypervigilance and treatment status. Simple linear regression finds the Eckardt score predicts 36.1% of the variance in quality of life on its own.

Violin plots of the APRO + EHAS clusters with the mean (SD) quality of life score further demonstrate the relationship between these three constructs (Figure 3). Patients with high achalasia symptoms and high symptom anxiety/hypervigilance reported the poorest quality of life while those with low symptoms and low anxiety/hypervigilance reported the highest quality of life scores.

4 | DISCUSSION

A 2021 review discusses the widespread use of the Eckardt score as a measure of achalasia outcomes, including endorsement by the International Society of Diseases of the Oesophagus and the American Society of Gastrointestinal Endoscopy.[20] A RAND Delphi process to establish consensus for quality indicators in achalasia, also from 2021, suggests a lack of confidence in current patient-reported outcome measures to accurately detect treatment failure and need for further intervention.[21] As such, we aimed to validate a new measure of achalasia symptom severity (the APRO) in a large sample of patients with treated and untreated achalasia, and those with normal motility as a comparator group.

Overall, the APRO demonstrates good to excellent reliability and construct validity and yields three subscales for dysphagia, reflux and chest pain. In addition to these scales, the APRO also evaluates whether a patient has changed their diet to manage their achalasia symptoms, and weight changes including weight gain. The APRO accurately differentiates between achalasia patients and those with normal motility, and people without gastrointestinal disease score very low on this scale. The APRO can also potentially distinguish patients with Type III from those with achalasia Types I and II when considering surgical history.

An important consideration of a patient outcome measure is its ability to detect change between pre- and post-treatment assessments. In our cohort, achalasia patients who had a history of foregut surgery *prior* to the present study scored lower on the APRO than untreated patients, but their scores remained higher than the normal motility group. In patients who underwent surgery during the study, a 58% reduction in APRO score occurred after surgery. The Eckardt score also declined, but less so (by 38%) and the average score remained above the clinical cutoff of 3. The scoring structure of the APRO may allow for superior precision in detecting meaningful change after achalasia treatments than the Eckardt score. Since these analyses only occurred in a small subset of the sample, replication studies are needed to evaluate how the APRO performs when evaluating achalasia treatments in parallel with the Eckardt score and objective treatment outcomes (e.g. timed barium oesophagram).

Health-related quality of life is emerging as an outcome as important as symptom control and physiological healing.[22] A recent study evaluated the Eckardt score with the Achalasia Quality of Life

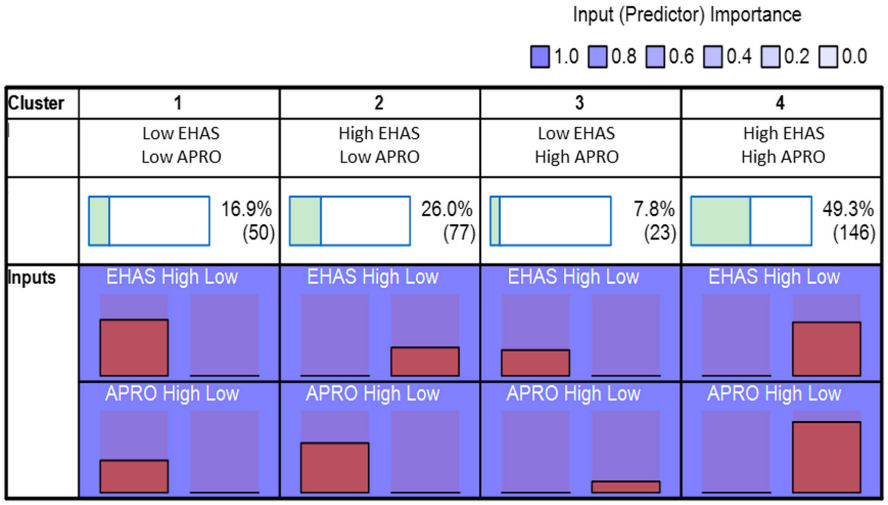

FIGURE 2 Achalasia patient clusters for APRO score, hypervigilance and symptom-anxiety. APRO, Achalasia Patient-Reported Outcomes Questionnaire

TABLE 4 Hierarchical regression models for predictive relationships of APRO and HRQoL in achalasia.

	R^2_{Adj}	*F* change	*p*	Standardised *β*
Model 1	0.594	427.565	<0.001	
EHAS: Anxiety				−0.771
Model 2	0.655	52.316	<0.001	
EHAS: Anxiety				−0.598
APRO total				−0.304
Prior foregut surgery (1 = Yes)			0.749	
Model 1	0.481	272.529	<0.001	
EHAS: Hypervigilance				−0.695
Model 2	0.607	95.117	<0.001	
EHAS: Hypervigilance				−0.500
APRO total				−0.407
Prior foregut surgery (1 = Yes)			0.076	

Abbreviations: APRO, Achalasia Patient-Reported Outcomes Questionnaire; EHAS, Oesophageal Hypervigilance and Anxiety Scale; HRQoL, health-related quality of life.

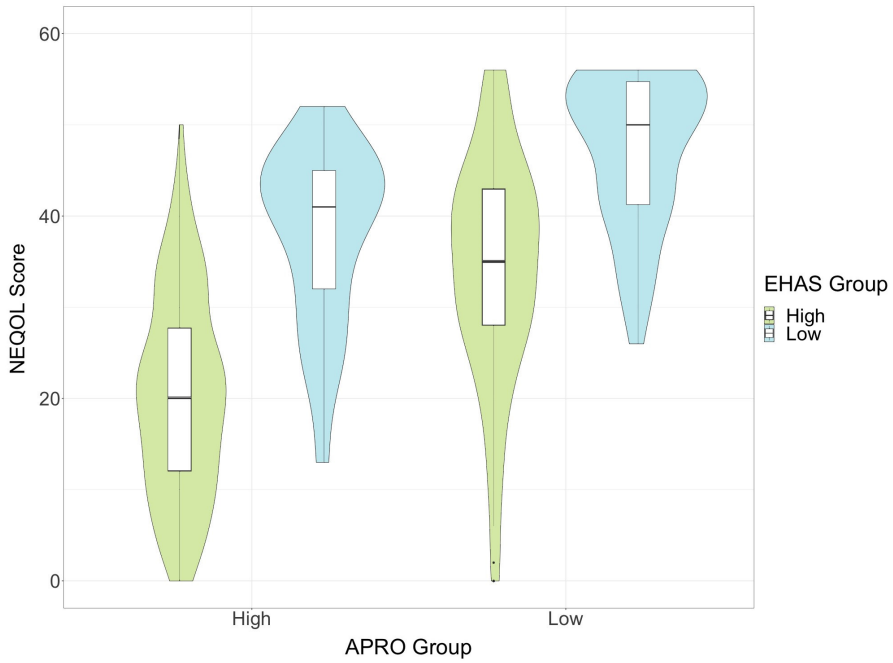

FIGURE 3 Violin plot for mean health-related quality of life by APRO and EHAS grouping. APRO, Achalasia Patient-Reported Outcomes Questionnaire; EHAS, Oesophageal Hypervigilance and Anxiety Scale

Scale and found these two scales to be reliable predictors of treatment success, but also found the two scales to have a large correlation suggesting they are not measuring distinct concepts ($r = 0.85$).[23] This is likely because 70% of the ASQ is symptom-based questions (e.g. how much of a problem was heartburn for you?). Debate exists on whether a quality-of-life measure should include any symptom questions as these can distort a unique construct regarding symptom *impacts and expectations*.[24,25] We used a measure of oesophageal quality of life with no symptom items (NEQOL), and the APRO demonstrated a modest relationship. This suggests using the APRO with the NEQOL (or other non-symptom-based quality of life tool) provides a more accurate assessment of how achalasia symptoms may affect social, emotional and financial domains of patient lives.

Symptom anxiety and hypervigilance have also emerged as critical constructs in understanding oesophageal patient outcomes. Like prior studies, patients with higher APRO scores reported greater anxiety and hypervigilance, and vice versa.[11–14,17,26] Four possible patient phenotypes of APRO symptom severity and anxiety/hypervigilance emerged. Two aligned as expected (high symptoms, high anxiety/hypervigilance; low symptoms, low anxiety/hypervigilance) and two identified patients who may be at higher risk of the poorest outcomes. In clinical practice, patients who report lower achalasia symptoms but have elevated anxiety and hypervigilance are at risk for comorbidities and may continue to seek medical care despite well-managed disease.[27] Alternatively, those with high levels of achalasia symptoms with low anxiety/hypervigilance may

simply be very well adjusted to their illness but may also be at risk of under-reporting disease activity and presenting later to follow-up than would be ideal. As such, understanding a patient's phenotype can guide clinical decision making and utilisation of adjunctive behavioural medicine services.

This study has some limitations that should be considered when interpreting the results. While we have a large sample size, patients were recruited from a tertiary university-based medical centre that specialises in achalasia care. As such, patients may have more severe disease and may be more likely to have had surgical intervention. Patients in the normal motility group are still symptomatic, although more often presenting for workup for reflux than dysphagia. Thus, this group is not a true control group which may obscure some of the comparisons. Further, we did not compare the achalasia cohort to those with abnormal motility patterns but not meeting the diagnostic criteria for achalasia. We do not have measures of other affective conditions like non-illness-related anxiety or depression, which could impact how the APRO performs as an outcomes tool. Lastly, we did not generate a cut-off value to identify 'active achalasia' which exists for the Eckardt score. Future studies will address these limitations.

5 | CONCLUSION

The APRO is a reliable and valid patient-reported outcome measure to evaluate symptom severity in patients with achalasia in both clinical practice and research. The scale adds measures of dietary change to assess possible burden for those opting to modify their diet. The APRO also adds a 'weight gain' option missing in the Eckardt score to determine if patients who gain weight have their own unique outcomes. Additional studies are necessary to compare the APRO to the Eckardt score as a measure of achalasia outcomes, especially among patients undergoing surgery. The APRO's relationship to hypervigilance and symptom-specific anxiety, and the four possible phenotypes in achalasia, also warrant further investigation.

AUTHOR CONTRIBUTIONS

John E. Pandolfino: Conceptualization (lead); data curation (supporting); formal analysis (supporting); funding acquisition (lead); methodology (equal); project administration (equal); resources (lead); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Dustin A. Carlson:** Conceptualization (equal); data curation (lead); formal analysis (supporting); methodology (supporting); project administration (equal); writing – original draft (supporting); writing – review and editing (equal). **Josie McGarva:** Writing – original draft (supporting); writing – review and editing (supporting). **Peter J. Kahrilas:** Data curation (supporting); writing – original draft (supporting); writing – review and editing (equal). **Michael Vaezi:** Conceptualization (equal); methodology (equal); writing – original draft (supporting); writing – review and editing (equal). **David Katzka:** Conceptualization (equal); methodology (equal); writing – original draft (supporting); writing – review and editing (equal). **Tiffany H. Taft:** Conceptualization (lead); data curation (equal); formal analysis (lead); methodology (lead); validation (lead); visualization (lead); writing – original draft (lead); writing – review and editing (equal).

AUTHORSHIP

Tiffany H. Taft is acting as the submission's guarantor and takes responsibility for the integrity of the work, from inception to published article. All authors approved the final version of the manuscript.

Guarantor of article: Tiffany H. Taft

FUNDING INFORMATION

This study is funded by NIH-NIDDK 1P01DK117824-01.

CONFLICT OF INTEREST

THT: Ownership interest in Oak Park Behavioural Medicine LLC, consulting for Takeda, GastroGirl; DAC: Speaking/consulting for Medtronic; Consulting for Phathom Pharmaceuticals; DK: Grant Funding from Shire; MV: Advisory for Ironwood, Phathom, ISOThrive, Sanofi, Bayer, Medtronic, Diversatek; PJK: Consulting for Ironwood, Reckitt Benchiser, Johnson & Johnson; Grant Funding from Ironwood; JM: None; JEP: Consulting/Speaking for Medtronic, Diversatek, Ethicon/J&J, Endogastric Solutions, Ironwood, Astra Zeneca, Takeda, Phathom, Neurogastrx, Medtronic- Shared patent.

ETHICS STATEMENT

The study was approved by the Institutional Review Board of Northwestern University (Study ID: STU00201372). A full waiver of consent was approved as the questionnaires are minimal risk and given as standard of care within the GI clinic. Participant privacy was ensured by assignment of a unique study ID for all data analyses.

ORCID

Dustin A. Carlson https://orcid.org/0000-0002-1702-7758
Peter J. Kahrilas https://orcid.org/0000-0002-8260-1250
David Katzka https://orcid.org/0000-0003-2307-2317
Tiffany H. Taft https://orcid.org/0000-0002-4670-2441

REFERENCES

1. Shaw MJ, Talley NJ, Beebe TJ, Rockwood T, Carlsson R, Adlis S, et al. Initial validation of a diagnostic questionnaire for gastroesophageal reflux disease. Am J Gastroenterol. 2001;96:52–7.
2. Taft TH, Carlson DA, Triggs J, Craft J, Starkey K, Yadlapati R, et al. Evaluating the reliability and construct validity of the Eckardt symptom score as a measure of achalasia severity. Neurogastroenterol Motil. 2018;30:e13287.
3. Eckardt VF, Aignherr C, Bernhard G. Predictors of outcome in patients with achalasia treated by pneumatic dilation. Gastroenterology. 1992;103:1732–8.
4. Spiegel BM, Hays RD, Bolus R, Melmed GY, Chang L, Whitman C, et al. Development of the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) gastrointestinal symptom scales. Am J Gastroenterol. 2014;109:1804–14.
5. Schoepfer AM, Straumann A, Panczak R, Coslovsky M, Kuehni CE, Maurer E, et al. Development and validation of a symptom-based activity index for adults with eosinophilic esophagitis. Gastroenterology. 2014;147:1255–1266.e21.

6. Lazarescu A, Karamanolis G, Aprile L, De Oliveira RB, Dantas R, Sifrim D. Perception of dysphagia: lack of correlation with objective measurements of esophageal function. Neurogastroenterol Motil. 2010;22:1292–7. e336-7.
7. Schoepfer AM, Simko A, Bussmann C, Safroneeva E, Zwahlen M, Greuter T, et al. Eosinophilic esophagitis: relationship of subepithelial eosinophilic inflammation with epithelial histology, endoscopy, blood eosinophils, and symptoms. Am J Gastroenterol. 2018;113:348–57.
8. Safroneeva E, Straumann A, Schoepfer AM. Latest insights on the relationship between symptoms and biologic findings in adults with eosinophilic esophagitis. Gastrointest Endosc Clin N Am. 2018;28:35–45.
9. Taft TH, Riehl M, Sodikoff JB, Kahrilas PJ, Keefer L, Doerfler B, et al. Development and validation of the brief esophageal dysphagia questionnaire. Neurogastroenterol Motil. 2016;28:1854–60.
10. Newberry C, Vajravelu RK, Pickett-Blakely O, Falk G, Yang YX, Lynch KL. Achalasia patients are at nutritional risk regardless of presenting weight category. Dig Dis Sci. 2018;63:1243–9.
11. Taft TH, Carlson DA, Simons M, Zavala S, Hirano I, Gonsalves N, et al. Esophageal hypervigilance and symptom-specific anxiety in patients with eosinophilic esophagitis. Gastroenterology. 2021;161:1133–44.
12. Guadagnoli L, Yadlapati R, Taft T, Pandolfino JE, Tye M, Keefer L. Esophageal hypervigilance is prevalent across gastroesophageal reflux disease presentations. Neurogastroenterol Motil. 2021;33:e14081.
13. Carlson DA, Gyawali CP, Roman S, Vela M, Taft TH, Crowell MD, et al. Esophageal hypervigilance and visceral anxiety are contributors to symptom severity among patients evaluated with high-resolution esophageal manometry. Am J Gastroenterol. 2020;115:367–75.
14. Wong MW, Liu TT, Yi CH, Lei WY, Hung JS, Cock C, et al. Oesophageal hypervigilance and visceral anxiety relate to reflux symptom severity and psychological distress but not to acid reflux parameters. Aliment Pharmacol Ther. 2021;54:923–30.
15. Pandolfino JE, Gawron AJ. Achalasia: a systematic review. JAMA. 2015;313:1841–52.
16. Bedell A, Taft TH, Keefer L, Pandolfino J. Development of the Northwestern esophageal quality of life scale: a hybrid measure for use across esophageal conditions. Am J Gastroenterol. 2016;111:493–9.
17. Taft TH, Triggs JR, Carlson DA, Guadagnoli L, Tomasino KN, Keefer L, et al. Validation of the oesophageal hypervigilance and anxiety scale for chronic oesophageal disease. Aliment Pharmacol Ther. 2018;47:1270–7.
18. Yadlapati R, Kahrilas PJ, Fox MR, Bredenoord AJ, Prakash Gyawali C, Roman S, et al. Esophageal motility disorders on high-resolution manometry: Chicago classification version 4.0([c]). Neurogastroenterol Motil. 2021;33:e14058.
19. Kahrilas PJ, Bredenoord AJ, Fox M, Gyawali CP, Roman S, Smout AJPM, et al. The Chicago classification of esophageal motility disorders, v3.0. Neurogastroenterol Motil. 2015;27:160–74.
20. Dorsey YC, Song EJ, Leiman DA. Beyond the Eckardt score: patient-reported outcomes measures in esophageal disorders. Curr Gastroenterol Rep. 2021;23:29.
21. Kamal AN, Kathpalia P, Otaki F, Bredenoord AJ, Castell DO, Clarke JO, et al. Development of quality indicators for the diagnosis and management of achalasia. Neurogastroenterol Motil. 2021;33:e14118.
22. Zhong C, Tan S, Ren Y, Lü M, Peng Y, Fu X, et al. Quality of life following peroral endoscopic myotomy for esophageal achalasia: a systematic review and meta-analysis. Ann Thorac Cardiovasc Surg. 2020;26:113–24.
23. Slone S, Kumar A, Jacobs J, Velanovich V, Richter JE. Accuracy of achalasia quality of life and Eckardt scores for assessment of clinical improvement post treatment for achalasia. Dis Esophagus. 2021;34:doaa080.
24. Allison PJ, Locker D, Feine JS. Quality of life: a dynamic construct. Soc Sci Med. 1997;45:221–30.
25. Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. Qual Life Res. 1993;2:441–9.
26. Cisternas D, Taft T, Carlson DA, Glasinovic E, Monrroy H, Rey P, et al. The Spanish version of the esophageal hypervigilance and anxiety score shows strong psychometric properties: results of a large prospective multicenter study in Spain and Latin America. Neurogastroenterol Motil. 2021;33:e14102.
27. Horenstein A, Heimberg RG. Anxiety disorders and healthcare utilization: a systematic review. Clin Psychol Rev. 2020;81:101894.

How to cite this article: Pandolfino JE, Carlson DA, McGarva J, Kahrilas PJ, Vaezi M, Katzka D, et al. Validation of the Achalasia Patient-Reported Outcomes Questionnaire. Aliment Pharmacol Ther. 2022;56:1434–1443. https://doi.org/10.1111/apt.17230