Psychometric evaluation of the Dysphagia Symptom Questionnaire for adults and adolescents with eosinophilic esophagitis

Eilish McCann, PhD,[a] Kathryn A. Peterson, MD,[b] Diane Whalley, PhD,[c] Shanshan Qin, PhD,[d] Sarette T. Tilton, PharmD,[e] Siddhesh Kamat, MSc,[a] Xian Sun, PhD,[a] and Evan S. Dellon, MD[f] *Sleepy Hollow, NY; Salt Lake City, Utah; Manchester, United Kingdom; Research Triangle Park and Chapel Hill, NC; and Bridgewater, NJ*

Background: Eosinophilic esophagitis (EoE) is a chronic, inflammatory disease of the esophagus leading to symptoms of esophageal dysfunction; dysphagia is the most common symptom experienced by adults and adolescents.
Objective: We sought to perform a psychometric evaluation of the Dysphagia Symptom Questionnaire (DSQ), a patient-reported outcome measure for patients with EoE.
Methods: Using baseline and week 24 data from the randomized, interventional, multinational phase 3 R668-EE-1774 trial (NCT03633617), the measurement properties of the DSQ—including reliability, construct and known-groups validity, responsiveness, and interpretation of change—were evaluated.
Results: The analysis population comprised 239 patients with EoE (age [mean ± SD], 28.1 ± 13.14 years; 63.6% male; 90.4% White). Intraclass correlation coefficients of 0.92 and 0.97 exceeded the acceptable reliability threshold (≥0.70). Construct validity correlations with EoE symptom and impact measures were moderate at baseline ($|r|$ = 0.44-0.55) and week 24 ($|r|$ = 0.55-0.69), and the DSQ biweekly total score discriminated among groups defined by disease severity. Analyses exploring interpretation of change from baseline on the DSQ biweekly total score indicated thresholds for within-patient improvement ranging from 9 to 23 points; a within-patient improvement from baseline of 13 points or greater could be considered clinically meaningful.
Conclusions: This analysis confirmed that the DSQ has acceptable distributional properties, test-retest reliability, construct validity, and ability to detect change. Therefore, the DSQ is a valid and reliable measure to assess the patient-reported symptom of dysphagia among adult and adolescent patients with EoE in the context of a clinical trial setting. (J Allergy Clin Immunol Global 2024;3:100302.)

Key words: *Eosinophilic esophagitis, patient-reported outcomes measure, dysphagia, symptoms, validation*

Eosinophilic esophagitis (EoE) is a chronic, inflammatory disease of the esophagus characterized by local eosinophilic inflammation leading to symptoms of esophageal dysfunction.[1-4] EoE may be either sporadic or familial, with an estimated 7% of individuals with EoE having a family member also affected by EoE.[5] More than 1 in 1000 people currently live with EoE in Europe and North America, where prevalence is the highest; the incidence rate ranges from 5 to 20 new cases per 100,000 people annually in both adults and children.[6] Symptoms of EoE are detrimental to the health-related quality of life of patients, who report significant physical, psychological, and social burdens.[7] As EoE progresses, chronic inflammation can lead to fibrosis; this fibrosis in turn contributes to food impactions,[8-10] which can require emergent endoscopic removal to relieve.[3,8,9,11] Thus, diagnosis in the early stages of disease is key.[12] Although there is no cure for EoE, treatment options include dietary changes, swallowed topical steroids (fluticasone or budesonide), a biologic (dupilumab), and endoscopic therapy with esophageal dilation.[13] Anti-inflammatory agents have been investigated as more effective treatments for EoE.[13]

Although younger children with EoE experience a range of symptoms, dysphagia is the most prevalent EoE symptom in both adults and adolescents,[13-17] and as such is an important patient-reported end point in clinical trials evaluating therapeutics for EoE. A number of disease-specific patient-reported outcome (PRO) measures have been used to measure EoE symptoms in a clinical trial setting, including the Straumann Dysphagia Index, the Eosinophilic Esophagitis Activity Index, the Dysphagia Numeric Rating Scale, and the Dysphagia Symptom Questionnaire (DSQ).[16,18-20] The DSQ was developed in accordance with guidance from the US Food and Drug Administration[21] to measure the daily frequency and severity of dysphagia associated with EoE.[16,20] Since the initial development of the DSQ in 2013,[16] its content validity and psychometric properties have been evaluated and described.[20,22] The DSQ has been applied successfully in several clinical trials to evaluate patient outcomes.[23-27] Strengths of the DSQ include its daily recall period, its development with both adult and adolescent patients, and its assessment of both frequency and severity of dysphagia.

The randomized, placebo-controlled, phase 3 clinical trial R668-EE-1774 (NCT03633617) was conducted to investigate the efficacy and safety of dupilumab in adults and adolescents with EoE.[26] Dupilumab is a humanized mAb against IL-4 and IL-13, 2 cytokines that are elevated in patients with EoE and play a role in pathogenesis.[28,29] The absolute change in DSQ biweekly total

From [a]Regeneron Pharmaceuticals, Inc, Sleepy Hollow; [b]the University of Utah Health, Salt Lake City; [c]RTI Health Solutions, Manchester; [d]RTI Health Solutions, Research Triangle Park; [e]Sanofi Global, Bridgewater; and [f]the University of North Carolina School of Medicine, Chapel Hill.

Received for publication September 18, 2023; revised March 27, 2024; accepted for publication April 28, 2024.
Available online 16 July 2024.

Corresponding author: Eilish McCann, PhD, Regeneron Pharmaceuticals, Inc, 1 Rockwood Rd, Sleepy Hollow, NY 10591. E-mail: eilish.mccann@regeneron.com.

The CrossMark symbol notifies online readers when updates have been made to the article such as errata or minor corrections

2772-8293

© 2024 The Authors. Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.jacig.2024.100302

Abbreviations used

DSQ: Dysphagia Symptom Questionnaire
EoE: Eosinophilic esophagitis
EoE-EREFS: Eosinophilic Esophagitis Endoscopic Reference Score
EoE-IQ: Eosinophilic Esophagitis Impact Questionnaire
EoE-SQ: Eosinophilic Esophagitis Symptom Questionnaire
eos/hpf: Eosinophils/hpf
ICC: Intraclass correlation coefficient
PGIC: Patient Global Impression of Change
PGIS: Patient Global Impression of Severity
PRO: Patient-reported outcome

score from baseline to week 24 was included as a coprimary symptom outcome in the R668-EE-1774 study, along with the proportion of patients who achieved a peak esophageal intraepithelial eosinophil count of less than or equal to 6 eosinophils/hpf (eos/hpf) at week 24. The goal of the present analysis was to use the data gathered in the R668-EE-1774 trial to assess the psychometric properties of the DSQ in adult and adolescent patients with EoE.

METHODS

Study design and participants

The randomized, phase 3 clinical trial enrolled adults and adolescents with EoE who were 12 years or older. The trial was conducted in 3 parts and a follow-up period: part A (n = 81) and part B (n = 240) (each consisting of a 24-week double-blind treatment period), part C (a 28-week extended active treatment period), and a 12-week follow-up period after the end of the extended active treatment period. The DSQ was administered daily throughout the trial and the follow-up period.[26] For trial enrollment, patients had to have a baseline DSQ biweekly total score of 10 or higher. For this evaluation, DSQ data from part A were used for preliminary analyses, whereas DSQ data from part B were used for the primary psychometric analyses. The PRO analysis population included all randomized patients who completed at least 1 of 3 PRO measures (DSQ, Eosinophilic Esophagitis Symptom Questionnaire [EoE-SQ], or Eosinophilic Esophagitis Impact Questionnaire [EoE-IQ]) at baseline in the respective study part. In the present analysis, data were pooled across treatment arms within each study part, because the goal was to evaluate the psychometric properties of the DSQ irrespective of treatment allocation. Findings presented here are for part B unless otherwise specified; results for part A were similar but are not presented.

Study measures

Study participants completed the DSQ each day following their final meal using an electronic diary. Questions on the DSQ have a daily recall period; question 1 asks whether or not the patient ate solid food that day, and, for patients who did eat solid food that day, questions 2 and 3 assess the frequency and severity of any dysphagia that occurred. A fourth DSQ item assessing pain when swallowing food (odynophagia) was also administered in the study but did not contribute to the DSQ scoring. The daily DSQ score ranges from 0 to 6; item scoring details are presented in Table I. The DSQ biweekly total score was calculated over a 14-day period as the sum of the daily DSQ scores divided by the number of days with electronic diary data, multiplied by 14 to convert the score to a biweekly total score. A minimum of 8 diary completion days were required over that period for the biweekly total score to be calculated.[20] The DSQ biweekly total score can range from 0 to 84, where higher scores indicate a greater dysphagia burden:

$$\text{DSQ biweekly total score} = \frac{\text{Sum of points from questions 2 and 3 from daily DSQ electronic diary}}{\text{Number of electronic diary days reported with nonmissing data}} \times 14 \text{ days.}$$

Several supporting PRO measures were used in the psychometric evaluation of the DSQ: the Patient Global Impression of Severity (PGIS), the Patient Global Impression of Change (PGIC), the EoE-SQ, and the EoE-IQ. The peak esophageal intraepithelial eosinophil count and the Eosinophilic Esophagitis Endoscopic Reference Score (EoE-EREFS) were also included as supporting clinical measures. Table II provides the supporting study measures in detail.

Statistical analyses

Quality of completion and score distribution. Quality of completion of the DSQ was assessed by the number and percentage of patients with missing daily assessments over the course of the study. Descriptive statistics were summarized for DSQ biweekly total scores at baseline and week 24 and also for the change between baseline and week 24. Floor (worst outcome) and ceiling (best outcome) effects were defined as there being more than 20% of patients with the worst DSQ biweekly total score or the best DSQ biweekly total score, respectively.

Using data from patients in part A and a subset of patients in part B of the study, simulation analyses were conducted to evaluate the missing-data rule for the DSQ that allows up to 6 days of missing data from the 14-day calculation period. Using baseline data from patients with all 14 daily DSQ scores in a 14-day period, daily scores were randomly set to "missing" for each patient for missing levels ranging from 1 to 13 days during the 14-day period. At each missing level, the score was considered stable if, over 500 simulation replications, the 95% CI of the SD of the DSQ biweekly total score was within a reference range defined as the SD ± 0.1 SD of the DSQ biweekly total score derived from the complete data.

Reliability. To evaluate test-retest reliability, or the reproducibility of the DSQ biweekly total scores over time, intraclass correlation coefficients (ICCs) for scores at week 20 (test) and week 24 (retest) were computed using 2-way, mixed-effects

TABLE I. DSQ items and scoring

Question no.*	Response option	Score
1. Since you woke up this morning, did you eat solid food?	No	No score assigned
	Yes	No score assigned
1a. Please select the reason for not eating solid food since you woke up this morning.†	Because of your problems with swallowing solid food	No score assigned Sensitivity analysis: score of 6
	Because of a reason NOT related to your problems with swallowing solid food	No score assigned
2. Since you woke up this morning, has food gone down slowly or been stuck in your throat?	No	0
	Yes	2
3. For the most difficult time you had swallowing food today (during the past 24 h), did you have to do anything to make the food go down or to get relief?	No, it got better or cleared up on its own	0
	Yes, I had to drink liquid to get relief	1
	Yes, I had to cough and/or gag to get relief	2
	Yes, I had to vomit to get relief	3
	Yes, I had to seek medical attention to get relief	4

The DSQ was modified from DSQ v 4.0, J. Patient-Reported Outcomes (2017) 1:3, © S. Hudgens, C. Evans, E. Phillips, and M. Hill, licensed under Creative Commons 4.0. http://creativecommons.org/licenses/by/4.0.

*Question 4 ("What was the worst pain you had while swallowing food over the past 24 h?") measures pain related to swallowing as an exploratory outcome and is not included in the DSQ daily score calculation.

†Question 1a was added to the DSQ on the basis of the Health Authority recommendation.

TABLE II. Study measures

Outcome measure	Measurement concept	Scoring	Recall period
Global assessment measures			
PGIS	Overall difficulty in swallowing food	4-point scale: 1 = none; 2 = mild; 3 = moderate; 4 = severe Lower scores indicate lower symptom severity	Past week
PGIC	Overall change in difficulty swallowing food	7-point scale: 0 = very much better; 1 = moderately better; 2 = a little better; 3 = no change; 4 = a little worse; 5 = moderately worse; 6 = very much worse Lower scores indicate greater improvement in difficulty swallowing food	Change since started taking study injection
EoE-specific PRO measures			
EoE-IQ	Emotional, social, work and school, and sleep impact of EoE	Range, 1 to 5 Higher scores indicate worse HRQoL	Past 7 d
EoE-SQ	Frequency and severity of 5 EoE symptoms other than dysphagia: chest pain, stomach pain, burning feeling in the chest, food or liquid coming back up into the throat, and throwing up	EoE-SQ frequency: range, 5 to 25 EoE-SQ severity: range, 0 to 30 Higher scores indicate more frequent/more severe symptoms	Past 7 d
Endoscopic/histologic measures			
EoE-EREFS	Endoscopic scoring measure based on inflammatory and remodeling features of EoE for proximal and distal regions of the esophagus[30]	Range, 0 to 18 Higher scores indicate greater disease activity	Current
Peak esophageal intraepithelial eosinophil count	Histologic measure based on esophageal biopsies	Maximum of the quantities of eos/hpf Categorized into 3 levels: 1 = ≤6 eos/hpf; 2 = >6 to <15 eos/hpf; 3 = ≥15 eos/hpf	Current

HRQoL, Health-related quality of life.

ANOVA with absolute agreement for single measures.[31] Two "stable" patient subsets were defined: patients with the same PGIS score at test and retest and patients with the same PGIC score at test and retest. Acceptable reliability was defined as an ICC of 0.70 or higher.[32,33]

Construct validity. Correlations between DSQ biweekly total scores and scores on the supporting measures at baseline and week 24 were determined to assess convergent and divergent validity. The strength of correlations was evaluated (correlations <0.3, weak or small; ≥0.3 to <0.7, moderate; ≥0.7 to <0.9, strong; and ≥0.9, very strong).[34,35] For the validation of the DSQ, we hypothesized the following correlations for the DSQ biweekly total scores:

- Moderate to strong correlations with the PGIS, EoE-SQ Frequency and Severity, and EoE-IQ scores;
- Small to moderate correlations with the peak esophageal intraepithelial eosinophil count; and
- Small correlations with EoE-EREFS.

In addition, ANOVAs and *t* tests were used to compare categories of PGIS, the presence or absence of EoE symptoms (assessed by EoE-SQ Frequency items), and peak eosinophil count at baseline and week 24 to determine the ability of the DSQ biweekly total score to differentiate between known groups defined by the supporting measures (known-group validity).

Responsiveness. To evaluate the DSQ's ability to detect change, correlations were calculated between the change in DSQ biweekly total score from baseline to week 24 and the corresponding changes assessed by supporting measures. In addition, ANOVAs were performed to compare changes in DSQ biweekly total scores from baseline to week 24 by responsiveness groups on the basis of PGIS, PGIC, and peak esophageal intraepithelial eosinophil count. Patients were considered improved if they had a more than or equal to 1-point improvement on the PGIS or responded at least "A little better" to the PGIC, and they were considered worsened if they had a more than or equal to 1-point deterioration on the PGIS or responded at least "A little worse" to the PGIC. Standardized effect size statistics were calculated for within- and between-group changes; a standardized effect size of 0.20 to 0.49 was considered small, 0.50 to 0.79 moderate, and more than or equal to 0.80 large.[34]

Interpretation of change. To explore meaningful within-patient change in DSQ biweekly total scores, anchor-based analyses were performed using part B data from the PGIS and the PGIC as candidate anchor measures. To be considered appropriate, the anchor measure was required to have a responsiveness correlation (of $|r| \geq 0.371$) with DSQ biweekly total scores.[36] Descriptive statistics of the changes in the DSQ biweekly total score were computed by the levels of the PGIS change from baseline to week 24 and the PGIC response at week 24. Candidate anchor levels included a 1-point improvement on the PGIS and a response of "A little better" for the PGIC. Given the potential for greater recall bias with the longer recall period of the PGIC,[37] the PGIS was considered the primary anchor measure. Empirical cumulative distribution function and probability density function plots were generated by anchor levels to provide visual support. The 0.5 SD of baseline scores and the standard error of measurement using the test-retest ICC between weeks 20 and 24 as a reliability estimate were computed, which yielded supportive, distribution-based estimates of meaningful change that quantify measurement error.

RESULTS

Patient characteristics

The PRO population of part B of the R668-EE-1774 study comprised 239 patients with a mean age of 28.1 ± 13.14 years (Table III). Most patients were male (63.6%) and White (90.4%). Of the 239 patients in the PRO population, 213 remained in the study at week 24.

DSQ completion and score distribution

At baseline, 100% of patients in the part B PRO population had sufficient diary completion to calculate a DSQ biweekly total score; by week 24, 175 patients (82% of the 213 patients) had a DSQ biweekly total score. Mean biweekly total score was 36.67 ± 11.22 at baseline and improved to 17.36 ± 18.05 at week 24 (Table IV). The proportion of patients answering "No" to question 1 ("Since you woke up this morning, did you eat solid food?") was 0.5% to 1.8% at baseline and 0% to 1.8% at week 24. On any given day during the 14-day assessment period, fewer than 12% of patients had missing DSQ daily scores at baseline, and fewer than 28% had missing DSQ daily scores at week 24 (see Table E1 in this article's Online Repository at www.jaci-global.org).

For the missing-data simulation analysis, the reference range for the SD of the DSQ biweekly total score at baseline was 11.75 to 14.36. The simulation showed that the 95% CI of the mean of the SDs from the repeat simulations remained within the reference range when up to 10 days of DSQ data were missing, supporting the rule of up to 6 missing days during each 14-day assessment period for computation of the biweekly total score.

There were no floor or ceiling effects in DSQ biweekly total scores at baseline (ie, no patient had the best [0] or the worst [84] score at baseline [Table IV]), thereby allowing the demonstration of both improvement and worsening in symptoms over time. At baseline, more than 80% of patients reported food going down slowly or getting stuck each day, although more than 60% of patients responded that they did not need to do anything, or only needed a drink, to get relief from dysphagia. DSQ scores decreased (improved) over time, and by week 24, 23.4% of patients reported the best score of 0.

Reliability

For the subsample of patients determined to be stable between weeks 20 and 24 on the basis of PGIS and PGIC scores, the test-retest ICCs for DSQ biweekly total scores were 0.92 and 0.97, respectively, which exceed the minimum value of 0.70 for acceptable reliability.[33] Mean score differences indicated no significant change in scores over the test-retest period ($P > .05$).

Construct validity

Convergent and divergent validity. Table V shows the construct validity correlations for DSQ biweekly total scores and supporting measures at baseline and week 24 (see also Table E2 in this article's Online Repository at www.jaci-global.org). The correlations were higher at week 24 than at baseline, and the patterns of association were as expected. Specifically, correlations with PGIS, EoE-SQ Frequency and Severity, and EoE-IQ scores were moderate at baseline ($|r| = 0.44$-0.55) and week 24 ($|r| = 0.55$-0.69), indicating convergent validity (ie, correlation between measures assessing similar EoE concepts). Correlations

TABLE III. Patient demographic and disease characteristics at baseline

Characteristics	Adolescents aged ≥12 to <18 y (n = 79)	Adults aged ≥18 y (n = 160)	Total (N = 239)
Age (y)			
Mean ± SD	15.0 ± 1.62	34.7 ± 11.32	28.1 ± 13.14
Median	15.0	35.0	24.0
Q1:Q3	14.0:16.0	24.0:41.5	16.0:38.0
Minimum:maximum	12:17	18:68	12:68
Sex, n (%)			
Female	22 (27.8)	65 (40.6)	87 (36.4)
Male	57 (72.2)	95 (59.4)	152 (63.6)
Race, n (%)			
Asian	2 (2.5)	3 (1.9)	5 (2.1)
Black or African American	7 (8.9)	1 (0.6)	8 (3.3)
White	64 (81.0)	152 (95.0)	216 (90.4)
Other	6 (7.6)	1 (0.6)	7 (2.9)
Not reported	0	3 (1.9)	3 (1.3)
Ethnicity, n (%)			
Hispanic or Latino	3 (3.8)	10 (6.3)	13 (5.4)
Not Hispanic or Latino	76 (96.2)	149 (93.1)	225 (94.1)
Unknown	0	1 (0.6)	1 (0.4)
Comorbidity, n (%)			
Allergic rhinitis	57 (72.2)	96 (60.0)	153 (64.0)
Asthma	45 (57.0)	62 (38.8)	107 (44.8)
Atopic dermatitis	33 (41.8)	29 (18.1)	62 (25.9)

TABLE IV. Descriptive statistics for DSQ biweekly total score at baseline and week 24

Statistics	Baseline (N* = 239)	Week 24 (n = 213)	Change baseline to week 24 (n = 213)
N	239	175	175
Mean ± SD	36.67 ± 11.220	17.36 ± 18.051	−18.77 ± 15.57
Median	38.77	9.00	−18.92
Q1:Q3	29.17:44.33	2.00:33.83	−29.27:−7.82
Minimum:maximum	8.4:70.0	00:70.0	−56.2:19.4
%† with most severe DSQ biweekly total score (floor effect)	0.0	0.0	NA
%† with least severe DSQ biweekly total score (ceiling effect)	0.0	23.4	NA
%‡ missing	0.0	17.8	17.8

NA, Not applicable.
*The number of patients in the PRO population who remained in the study at the corresponding time point.
†Percentage is based on n.
‡Percentage is based on N.

with EoE-EREFS at baseline and week 24 and with peak esophageal intraepithelial eosinophil count at week 24 were small ($|r| = 0.12$-0.19), indicating divergent validity (ie, weaker correlation between DSQ and measures assessing different EoE concepts [histology and endoscopic severity]).

Known-groups validity. Mean DSQ biweekly total scores for known groups on the basis of (1) PGIS and presence or absence of 5 EoE symptoms at baseline and week 24 and (2) peak esophageal intraepithelial eosinophil count at week 24 are presented in Table E2. Patterns of mean DSQ biweekly total scores by levels of PGIS met expectations, with higher mean scores observed with increasingly severe disease at both time points. Similarly, the DSQ biweekly total score was able to discriminate between the presence and absence of all 5 EoE symptoms and between categories of peak esophageal intraepithelial eosinophil count (all $P < .05$).

Responsiveness

Table V (see also Table E2) presents ANOVAs for mean DSQ biweekly total change scores by responsiveness groups at week 24. The highest negative DSQ biweekly total change scores (indicating the greatest improvement) were observed for patients who had improved on PGIS and PGIC. The omnibus tests, as well as the pairwise comparisons of mean change scores between patients who had improved and patients who had not changed, were statistically significant for both PGIS and PGIC (all $P \leq .0001$). There were no significant differences in DSQ biweekly total change scores between patients with peak esophageal intraepithelial eosinophil count less than or equal to 6 eos/hpf and those with greater than or equal to 6 eos/hpf at week 24. The pattern of responsiveness correlations was generally consistent with the cross-sectional correlations. DSQ biweekly total change scores showed moderate correlations with change scores on other

TABLE V. Summary of key measurement properties of the DSQ at baseline and week 24

Measurement property	Category	Part B
Reliability		
ICC (n) for scores at week 20 and week 24	Patients with no change on the PGIS from week 20 to week 24	0.92 (78)
	Patients with no change on the PGIC from week 20 to week 24	0.97 (85)
Construct validity		
Pearson correlation coefficient (n) at baseline/week 24	PGIS	0.47 (228)/0.69 (175)
	EoE-SQ Frequency	0.55 (228)/0.62 (174)
	EoE-SQ Severity	0.46 (228)/0.55 (174)
	EoE-IQ	0.44 (224)/0.68 (175)
	EoE-EREFS	–0.12 (237)/0.18 (180)
	Peak esophageal intraepithelial eosinophil count	Not conducted/0.19 (182)
Ability to detect change		
Pearson correlation of change (n) at week 24	PGIS	0.53 (166)
	PGIC	0.54 (177)
DSQ change by PGIS change between baseline and week 24	Mean change score (n)	Improved: –24.93 (102) No change: –8.43 (55) Worsened: –9.47 (9) $P \leq .0001$
	Within-group SES	Improved: –1.80 No change: –0.63 Worsened: –0.48
	Between-group SES	Improved vs no change: –1.21 Improved vs worsened: –1.08 No change vs worsened: 0.07

SES, Standardized effect size.

TABLE VI. Summary of anchor-based estimates to interpret change from baseline to week 24 in DSQ biweekly total score

	DSQ biweekly total change score from baseline to week 24			
Global assessment measure	**N**	**Mean ± SD**	**95% CI**	**Median**
PGIS				
Worsening (≥1-point worsening)	9	–9.47 ± 19.587	–24.53 to 5.59	0.00
0-point change	55	–8.43 ± 13.285	–12.03 to –4.84	–7.82
1-point improvement	77	–22.42 ± 13.202	–25.41 to –19.42	–23.46
2-point improvement	22	–30.43 ± 12.075	–35.78 to –25.07	–32.14
3-point improvement	3	–49.24 ± 6.529	–65.46 to –33.02	–48.22
PGIC				
Worsening (a little, moderately, very much)	5	–7.28 ± 13.122	–23.58 to –9.01	–8.17
No change	26	–3.02 ± 11.062	–7.49 to 1.45	0.00
A little better	39	–11.80 ± 14.470	–16.49 to –7.11	–9.33
Moderately better	44	–22.69 ± 14.385	–27.07 to –18.32	–23.06
Very much better	63	–28.40 ± 12.474	–31.54 to –25.26	–28.00

EoE-specific PRO measures and small correlations with changes in the endoscopic and histology measures (Table E2).

Interpretation of change

The adequacy of the PGIS and PGIC as anchor measures was confirmed by correlations with DSQ biweekly total change scores of 0.53 and 0.54, respectively (Table V), which exceed the required minimum of 0.371.[36] Mean and median changes in DSQ biweekly total scores across the levels of each anchor measure also were as expected, with greater reductions in DSQ biweekly total scores being generally associated with greater levels of improvement on the PGIS and PGIC at week 12 (see Table E3 in this article's Online Repository at www.jaci-global.org) and at week 24 (Table VI). The empirical cumulative distribution function and probability density function plots showed adequate separation of the anchor groups for the PGIS and to a lesser extent for the PGIC (see Fig E1-E4 in this article's Online Repository at www.jaci-global.org), supporting the use of PGIS and PGIC response categories to determine meaningful change on the DSQ biweekly total score.

Using the PGIS anchors for change from baseline to week 24, the median DSQ biweekly total change scores were –23.46 (n = 77; mean, –22.42) for a 1-point improvement on the PGIS and –32.14 (n = 22; mean, –30.43) for a 2-point improvement on the PGIS (Table VI; results for age subgroups are presented in Tables E4 and E5 in this article's Online Repository at www.jaci-global.org). The median DSQ biweekly total change scores

at week 24 using the PGIC anchors were –9.33 (n = 39; mean, –11.80) for "A little better" and –23.06 (n = 44; mean, –22.69) for "Moderately better." In addition, the lower limits of the 95% CI for patients with no change on the PGIS and PGIC were –12.03 and –7.49, respectively (Table VI). The results based on the PGIS (considered to be the primary anchor measure) indicated a threshold range of –13 to –24 for change on the DSQ biweekly total score. The corresponding estimates from both age groups fell into this range, with the adolescents' magnitude at the smaller side.

As additional sensitivity analyses, the distribution of DSQ score changes at week 24 by baseline PGIS ratings and also by age groups (adolescents and adults) were reviewed within participants who achieved a 1-category improvement on the PGIS at week 24 (see Table E6 in this article's Online Repository at www.jaci-global.org). Most participants with a 1-category improvement on the PGIS (n = 94) reported mild (n = 37) or moderate (n = 52) severity at baseline. The median change in the DSQ biweekly total score was –24.08 (mean, –21.88) for participants with a moderate baseline PGIS and –23.92 (mean, –25.20) for participants with a mild baseline PGIS. The corresponding estimates by baseline PGIS were similar across adolescents and adults, except for adolescents with mild PGIS at baseline, whose change scores were lower (median, –11.80; mean, –18.48; n = 8). Most estimates from the sensitivity analyses were within the threshold range of –13 to –24, estimated from the main analyses. Therefore, we propose a threshold range of a 13- to 24-point improvement to reflect meaningful within-patient change on the DSQ biweekly total score.

In the distribution-based analyses, the overall standard error of measurement was 3.07 on the basis of an ICC of 0.92 (PGIS stable sample) and 2.09 on the basis of an ICC of 0.97 (PGIC stable sample); the 0.5 SD of the baseline score was 5.61. The sensitivity results from 2 age groups deviated by 1 point or lower. As expected, these estimates were lower than the anchor-based estimates, thereby supporting the estimation of the meaningful within-patient change from the anchor-based analyses.

DISCUSSION

This psychometric analysis using data from the R668-EE-1774 clinical trial confirmed that the DSQ is a valid and reliable measure to assess patient-reported dysphagia among adult and adolescent patients with EoE. No floor or ceiling effects were found at baseline, and the random missing-data simulation supported the DSQ scoring algorithm, allowing up to 6 missing days across the 14-day score calculation period. This analysis also showed high test-retest reliability and adequate convergent and divergent and known-groups validity for the DSQ. The construct validity correlations tended to be stronger at week 24, because the study inclusion criteria required patients to be symptomatic at baseline, thereby limiting the dynamic range in baseline DSQ scores.

In addition, the DSQ had moderate to high responsiveness correlations with other EoE-specific patient-reported measures; as expected, correlations with clinical measures (ie, EoE-EREFS and peak esophageal intraepithelial eosinophil count) were weaker, consistent with previous findings that dysphagia symptom indices do not correlate strongly with histologic measures.[23,38] Patterns of change consistent with expectations and large between-group effect sizes in the comparison between the improved and no-change subgroups defined by the PGIS and PGIC demonstrated the ability of the DSQ to detect change. Anchor-based analyses indicated thresholds for improvement in the DSQ biweekly total score ranging from 13 to 24 points, with this level of improvement or greater considered clinically meaningful within-patient change. The present study population included patients who were nonresponsive to proton-pump inhibitors, many of whom were previously treated with topical steroids, and with substantial rates of steroid nonresponse and previous esophageal dilation; thus, the meaningful change threshold described here may not be directly applicable to other patient populations that differ in disease severity and/or previous treatment history.

The DSQ has been used in previous clinical trials to show significant improvements in dysphagia symptoms with EoE treatment, supporting its fitness for purpose in this context of use.[23,25] Results of this analysis, which included patients receiving either placebo or active treatment, further support that the DSQ has acceptable measurement properties, in addition to establishing thresholds for clinically meaningful within-patient change. The proposed threshold of 13 to 24 was generally higher than the previous mean changes of –6.5 and –13.5 from patients reporting "A little better" and "Better" PGIC ratings in the study by Hudgens et al.[20] This difference may be attributable to the use of different primary anchor measures (PGIS in our study vs PGIC in the study by Hudgens et al) and assessment period (week 24 in our study vs week 12 in the study by Hudgens et al). When similarly assessing the week 12 time point, we observed mean changes of –8.6 and –17.6 from patients reporting "A little better" and "Moderately better" on the current version of the PGIC (Table E3), which are close to the mean changes reported by Hudgens et al.[20] Given that the PGIS is a commonly preferred anchor measure and week 24 was the most relevant time point in our analyses, we suggest a point threshold of 13- to 24-point reduction for meaningful within-patient improvement on the DSQ biweekly total score.

Some limitations must be noted. The maximum possible score on the DSQ is 84 points, achievable only if patients had impactions severe enough to require emergency medical care every day throughout the 14-day period. Given that the maximum score of the DSQ is practically impossible, the question of whether its dynamic range is sufficient to show potential change must be considered when interpreting a meaningful within-patient change threshold. Qualitative patient input on noticeable and meaningful DSQ score changes to complement the anchor-based analyses would strengthen our interpretation of meaningful change.

The DSQ psychometric evaluation based on data from the R668-EE-1774 trial confirmed the measure's construct validity, reliability, and ability to detect change. The findings confirm those of previous studies and further strengthen the evidence that the DSQ is a psychometrically valid measure to assess symptoms of dysphagia among adults and adolescents with EoE.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The analysis presented here used data from the pivotal, randomized, placebo-controlled, multinational phase 3 R668-EE-1774 trial (NCT03633617). The local institutional review board or ethics committee at each trial center oversaw the conduct and documentation of the trial. Written informed consent or

assent (or both) was obtained from all the patients or their parent or legal guardian before enrollment.

DISCLOSURE STATEMENT

Regeneron Pharmaceuticals, Inc, and Sanofi Global provided the financial support for the study. RTI Health Solutions, an independent nonprofit research organization, received funding under a research contract with Regeneron Pharmaceuticals, Inc, and Sanofi Global to conduct this study and provide publication support in the form of manuscript writing, styling, and submission.

Disclosure of potential conflict of interest: E. S. Dellon has received research funding from Adare/Ellodi, Allakos, Arena/Pfizer, AstraZeneca, GlaxoSmithKline, Meritage, Miraca, Nutricia, Celgene/Receptos/BMS, Regeneron, Revolo, and Shire/Takeda; is a consultant for Abbott, AbbVie, Adare/Ellodi, Aimmune, Akesobio, Alfasigma, ALK, Allakos, Amgen, Aqilion, Arena/Pfizer, Aslan, AstraZeneca, Avir, Biorasi, Calypso, Celgene/Receptos/BMS, Celldex, Eli Lilly, EsoCap, Eupraxia, Fering, GlaxoSmithKline, Gossamer Bio, Holoclara, Invea, Knightpoint, Landos, LucidDx, Morphic, Nexstone Immunology, Nutricia, Parexel/Calyx, Phathom, Regeneron, Revolo, Robarts/Alimentiv, Salix, Sanofi, Shire/Takeda, Target RWE, and Upstream Bio; and has received educational grants from Allakos, Holoclara, and Invea. K. A. Peterson is a consultant or advisor for AGA, Alladapt, AstraZeneca, Allakos, Bristol Meyers Squibb, Ellodi, Invea, Lucid, Nexstone, Peerview, Regeneron, Takeda, and WebMD; has received research fees from AstraZeneca, Allakos, Adare, Regeneron-Sanofi, Revolo, repSpeaker, AGA, Regeneron, Peerview, Takeda, and WebMD; has received unrestricted grant support from Allakos and Chobani; and has equity in Nexeos Bio. D. Whalley and S. Qin are employees of RTI Health Solutions. S. T. Tilton is an employee of Sanofi Global and may hold stocks and/or shares in the company. E. McCann, S. Kamat, and X. Sun are employees of Regeneron Pharmaceuticals, Inc, and may hold shares and/or stock options in the company.

We thank Sara Musetti Jenkins and Kate Lothman of RTI Health Solutions for medical writing assistance.

Data availability: Qualified researchers may request access to study documents (including the clinical study report, study protocol with any amendments, blank case report form, and statistical analysis plan) that support the methods and findings reported in this article. Individual anonymized participant data will be considered for sharing once the indication has been approved by a regulatory body, if there is legal authority to share the data and there is not a reasonable likelihood of participant re-identification. Please submit requests on https://vivli.org/.

REFERENCES

1. Furuta GT, Aceves SS. The National Biome Initiative: an allergy perspective. J Allergy Clin Immunol 2017;139:1131-4.
2. DeBrosse CW, Franciosi JP, King EC, Butz BK, Greenberg AB, Collins MH, et al. Long-term outcomes in pediatric-onset esophageal eosinophilia. J Allergy Clin Immunol 2011;128:132-8.
3. Falk GW. Clinical presentation of eosinophilic esophagitis in adults. Gastroenterol Clin North Am 2014;43:231-42.
4. Liacouras CA, Furuta GT, Hirano I, Atkins D, Attwood SE, Bonis PA, et al. Eosinophilic esophagitis: updated consensus recommendations for children and adults. J Allergy Clin Immunol 2011;128:3-20. e6; quiz 1-2.
5. Sherrill JD, Rothenberg ME. Genetic dissection of eosinophilic esophagitis provides insight into disease pathogenesis and treatment strategies. J Allergy Clin Immunol 2011;128:23-32. quiz 3-4.
6. Arias A, Lucendo AJ. Epidemiology and risk factors for eosinophilic esophagitis: lessons for clinicians. Expert Rev Gastroenterol Hepatol 2020;14:1069-82.
7. Pokrzywinski RM, Harding G, Brooks A, Goodwin B, Williams J. Documenting the journey of patients with eosinophilic esophagitis and the impact of the disease on patients and their caregivers: a cross-sectional, qualitative research study. Adv Ther 2020;37:4458-78.
8. Straumann A. The natural history and complications of eosinophilic esophagitis. Gastrointest Endosc Clin N Am 2008;18:99-118; ix.
9. Straumann A, Spichtin HP, Grize L, Bucher KA, Beglinger C, Simon HU. Natural history of primary eosinophilic esophagitis: a follow-up of 30 adult patients for up to 11.5 years. Gastroenterology 2003;125:1660-9.
10. Lucendo AJ, Santander C, Savarino E, Guagnozzi D, Perez-Martinez I, Perello A, et al. EoE CONNECT, the European Registry of Clinical, Environmental, and Genetic Determinants in Eosinophilic Esophagitis: rationale, design, and study protocol of a large-scale epidemiological study in Europe. Therap Adv Gastroenterol 2022;15:17562848221074204.
11. Chehade M, Jones SM, Pesek RD, Burks AW, Vickery BP, Wood RA, et al. Phenotypic characterization of eosinophilic esophagitis in a large multicenter patient population from the Consortium for Food Allergy Research. J Allergy Clin Immunol Pract 2018;6:1534-44.e5.
12. Muir AB, Brown-Whitehorn T, Godwin B, Cianferoni A. Eosinophilic esophagitis: early diagnosis is the key. Clin Exp Gastroenterol 2019;12:391-9.
13. Dellon ES, Gonsalves N, Hirano I, Furuta GT, Liacouras CA, Katzka DA, et al. ACG clinical guideline: evidenced based approach to the diagnosis and management of esophageal eosinophilia and eosinophilic esophagitis (EoE). Am J Gastroenterol 2013;108:679-92; quiz 93.
14. Lucendo AJ, Molina-Infante J, Arias A, von Arnim U, Bredenoord AJ, Bussmann C, et al. Guidelines on eosinophilic esophagitis: evidence-based statements and recommendations for diagnosis and management in children and adults. United European Gastroenterol J 2017;5:335-58.
15. Fujiwara Y. Symptom-based diagnostic approach for eosinophilic esophagitis. J Gastroenterol 2020;55:833-45.
16. Dellon ES, Irani AM, Hill MR, Hirano I. Development and field testing of a novel patient-reported outcome measure of dysphagia in patients with eosinophilic esophagitis. Aliment Pharmacol Ther 2013;38:634-42.
17. US Food and Drug Administration. Eosinophilic esophagitis: developing drugs for treatment guidance for industry. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/eosinophilic-esophagitis-developing-drugs-treatment-guidance-industry. Accessed August 5, 2024.
18. Straumann A, Conus S, Degen L, Felder S, Kummer M, Engel H, et al. Budesonide is effective in adolescent and adult patients with active eosinophilic esophagitis. Gastroenterology 2010;139:1526-37, 37.e1.
19. 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;147:1255-66.e21.
20. Hudgens S, Evans C, Phillips E, Hill M. Psychometric validation of the Dysphagia Symptom Questionnaire in patients with eosinophilic esophagitis treated with budesonide oral suspension. J Patient Rep Outcomes 2017;1:3.
21. US Food and Drug Administration. Patient-reported outcome measures: use in medical product development to support labeling claims. Available at: https://www.fda.gov/media/77832/download. Accessed August 5, 2024.
22. Kamat S, Hartford C, Guillemin I, Chao J, Eckert L, Symonds T, et al. Qualitative patient interviews to assess content validity of the modified Dysphagia Symptom Questionnaire. Gastroenterology 2020;158:S831.
23. Dellon ES, Collins MH, Katzka DA, Hudgens S, Lan L, Williams J, et al. Improvements in dysphagia and pain with swallowing in patients with eosinophilic esophagitis receiving budesonide oral suspension. Clin Gastroenterol Hepatol 2021;19:699-706.e4.
24. Dellon ES, Woosley JT, Arrington A, McGee SJ, Covington J, Moist SE, et al. Efficacy of budesonide vs fluticasone for initial treatment of eosinophilic esophagitis in a randomized controlled trial. Gastroenterology 2019;157:65-73.e5.
25. Hirano I, Collins MH, Katzka DA, Mukkada VA, Falk GW, Morey R, et al. Budesonide oral suspension improves outcomes in patients with eosinophilic esophagitis: results from a phase 3 trial. Clin Gastroenterol Hepatol 2022;20:525-34.e10.
26. Dellon ES, Rothenberg ME, Collins MH, Hirano I, Chehade M, Bredenoord AJ, et al. Dupilumab in adults and adolescents with eosinophilic esophagitis. N Engl J Med 2022;387:2317-30.
27. NCT04322708: a study of lirentelimab (AK002) in patients with active eosinophilic esophagitis (KRYPTOS). ClinicalTrials.gov; 2023. Available at: https://clinicaltrials.gov/ct2/show/NCT04322708. Accessed June 8, 2023.
28. Gupta SK, Fitzgerald JF, Kondratyuk T, HogenEsch H. Cytokine expression in normal and inflamed esophageal mucosa: a study into the pathogenesis of allergic eosinophilic esophagitis. J Pediatr Gastroenterol Nutr 2006;42:22-6.
29. Blanchard C, Stucke EM, Rodriguez-Jimenez B, Burwinkel K, Collins MH, Ahrens A, et al. A striking local esophageal cytokine expression profile in eosinophilic esophagitis. J Allergy Clin Immunol 2011;127:208-17. 17.e1-7.

30. Hirano I, Moy N, Heckman MG, Thomas CS, Gonsalves N, Achem SR. Endoscopic assessment of the oesophageal features of eosinophilic oesophagitis: validation of a novel classification and grading system. Gut 2013;62:489-95.
31. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1:30-46.
32. Fink A, Litwin MS. How to measure survey reliability and validity. Thousand Oaks (CA): Sage; 1995.
33. Nunnally JC, Bernstein I. The role of university in the development of entrepreneurial vocations: a Spanish study. Psychometr Theory 1978;387-405.
34. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.
35. Hinkle DE, Wiersma W, Jurs SG. Applied statistics for the behavioral sciences. Boston: Houghton Mifflin College Division; 2003.
36. Hays RD, Brodsky M, Johnston MF, Spritzer KL, Hui KK. Evaluating the statistical significance of health-related quality-of-life change in individual patients. Eval Health Prof 2005;28:160-71.
37. US Food and Drug Administration. Patient-focused drug development (PFDD) guidance, incorporating clinical outcome assessments into endpoints for regulatory decision-making. Discussion document for guidance 4. 2019. Available at: https://www.fda.gov/media/132505/download. Accessed January 19, 2023.
38. Cotton CC, Woosley JT, Moist SE, McGee SJ, Iuga A, Shaheen NJ, et al. Determination of a treatment response threshold for the Eosinophilic Esophagitis Endoscopic Reference Score. Endoscopy 2022;54:635-43.