Received: 15 July 2019 | Revised: 5 February 2020 | Accepted: 6 February 2020

DOI: 10.1111/nmo.13827

ORIGINAL ARTICLE

Neurogastroenterology & Motility NGM WILEY

Fair reliability of eckardt scores in achalasia and non-achalasia patients: Psychometric properties of the eckardt spanish version in a multicentric study

Daniel Cisternas[1] | Hugo Monrroy[2] | Arnoldo Riquelme[2] | Oslando Padilla[2] | Eduardo Fuentes-López[3] | Arturo Valle[1] | Ricardo Mejia[2] | Albis Hani[4] | Andres F. Ardila-Hani[4] | Ana Maria Leguizamo[4] | Claudio Bilder[5] | Andres Ditaranto[5] | Jose Maria Remes-Troche[6] | Antonio Ruiz de León[7] | Julio Pérez de la Serna[7] | Ingrid Marin[8] | Jordi Serra[8]

[1]Facultad de Medicina Clínica Alemana, Clínica Alemana de Santiago, Universidad del Desarrollo. Santiago, Chile

[2]Pontificia Universidad Católica de Chile, Santiago, Chile

[3]Department of Health Sciences, Faculty of Medicine, Pontificia Universidad Catolica de Chile, Santiago, Chile

[4]San Ignacio Hospital, Pontificia Universidad Javeriana, Bogota, Colombia

[5]School of Medicine, Hospital Universitario, Fundación Favaloro, Buenos Aires, Argentina

[6]Universidad Veracruzana, Veracruz, México

[7]Hospital Clínico San Carlos, Universidad Complutense Madrid, Madrid, Spain

[8]Motility and Functional Gut Disorders Unit, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), University Hospital Germans Trias i Pujol, Universitat Autonoma de Barcelona, Badalona, Spain

Correspondence
Daniel Cisternas, Clínica Alemana de Santiago, Avenida Vitacura 5951, 6681920 - Santiago, Chile.
Email: dcisternasc@alemana.cl

Funding information
No funding declared.

Abstract

Background: Eckardt symptom score (ESS) is the most used tool for the evaluation of esophageal symptoms. Recent data suggest that it might have suboptimal reliability and validity. The aims of this study were as follows: (a) Develop and validate an international Spanish ESS version. (b) Perform psychometric ESS evaluation in patients with achalasia and non-achalasia patients.

Methods: Eckardt symptom score translation was performed by Delphi process. ESS psychometric evaluation was done in two different samples of patients referred for manometry. First sample: 430 dysphagia non-achalasia patients. Second sample: 161 achalasia patients. Internal consistency was evaluated using Cronbach's α and Guttman coefficient (<0.5 = unacceptable. 0.5-0.7 = fair. >0.7 = acceptable).

Key Results: Our data show that in patients without and with achalasia, ESS behaves similarly. Both show a fair reliability with Cronbach's α of 0.57 and 0.65, respectively. Based on our results, we recommend interpretation of the Spanish ESS be done with caution. The psychometric quality of the ESS could not be improved by removal of any items based on the single-factor structure of the scale and no items meeting criteria for elimination.

Conclusions and Inferences: Eckardt symptom score Spanish translation was developed. ESS showed a fair reliability for the evaluation of patients with any causes of dysphagia. Our results highlight the need for development and psychometric validation of new dysphagia scoring tools.

KEYWORDS

dysphagia, Eckardt symptom score, psychometric evaluation, reliability, validity

1 | INTRODUCTION

Achalasia is one of the most studied causes of dysphagia, with a prevalence of 10 in 100 000 worldwide.[1] Therapeutic options include Heller myotomy, pneumatic dilation (PD), and peroral endoscopic myotomy (POEM). All these options are effective but are invasive and could result in secondary effects.[2-4] A therapeutic decision is basically taken when symptoms burden justifies the procedure-related risks. Accordingly, scoring patient's symptoms is crucial. Achalasia´s predominant symptoms are dysphagia and regurgitation. Fisichella et al tested a sample of untreated achalasia patients and reported that dysphagia was present in 94%, regurgitation in 76%, heartburn in 52%, chest pain in 41%, and weight loss in 35% of patients.[5] Remarkably, all these symptoms are non-specific for achalasia.

In 1992, Eckardt et al published a prospective study of predictors of a good response to PD in 54 achalasia patients. To define clinical response, they designed a score that evaluated the frequency or severity of dysphagia, regurgitation, chest pain, and weight loss. There is no description of the criteria used to define which symptoms were included and how they were graded.[6] This score has been subsequently referred to as Eckardt symptom score (ESS). It is the most used score in achalasia studies, including its use as response criterion for different treatments,[7-10] in the evaluation of a novel manometric[11] and FLIP-derived[12] variables and as response predictor.[13]

High-resolution manometry is now considered the gold standard for diagnosing esophageal motility disorders.[14] The Chicago v3.0 classification[15] (CC3.0) provides a standardized and hierarchical approach to the analysis and categorization of abnormalities that has led to a significant increase in our knowledge of motility disorders, especially the recognition that a defining feature of major esophageal motility disorders is obstructive physiology, either at the gastro-esophageal junction, in the distal esophagus, or both.[16]

Recently, it has become a standard that the adequate use of a certain patient-related outcome (PRO) requires an evaluation of its validity and reliability.[17,18] Validity refers to the extent to which a PRO measures what it purports to measure. Reliability refers to the extent to which it provides stable and consistent results.[18] ESS was described long before these recommendations, and it was not until 2018 that the first ESS psychometric evaluation was published.[19] In that study, Taft et al evaluated ESS in 107 achalasia patients. They showed only a fair reliability (with both Cronbach α and Guttmann statistic below 0.7) and validity (with Pearson's correlation r between 0.15 and 0.41 for physiological variables and between 0.43 and 0.78 for other PROs).

The FDA suggests that the attributes of a certain PRO score (like its reliability and validity) cannot be assumed to be relevant in all the populations in which the instrument is used.[17] In the case of populations from other regions of the world, it recommends the translation and cultural adaptations of the instruments.[20] To our knowledge, no formal ESS Spanish translation/cultural adaptation has been published. On the other hand, FDA recommendations also apply to ESS use in different disease populations. Taft´s study evaluated ESS only in achalasia patients. Nevertheless, it has been used as a generic score for symptomatic burden evaluation in different esophageal conditions, including ineffective esophageal motility (IEM),[21] esophagogastric junction outflow obstruction (EGJOO),[22] and different spastic/hyper-contractile disorders.[23-25]

Key Points

- The Spanish translation of the Eckardt's Symptom Score (ESS) was developed.
- ESS showed fair reliability for the evaluation of patients with any cause of dysphagia.
- Development and psychometric validation of new dysphagia scoring tools is needed.

Thus, the aims of this study are as follows: (a). Perform a formal translation/cultural adaptation of ESS to Spanish, intended to be widely used in Spanish-speaking populations. (b). Perform a formal psychometric evaluation of ESS in achalasia patients from Spain and Latin America. (c). Perform a formal psychometric evaluation of ESS in non-achalasia patients complaining of dysphagia from Spain and Latin America.

2 | MATERIALS AND METHODS

2.1 | Eckardt symptom score translation

An initial Spanish ESS translation was performed by one English-speaker researcher (DC). Gastroenterologists and surgeons working in the field of esophageal disorders were recruited from different countries in Latin America and Spain. They participated in a Delphi process,[26-28] evaluating every phrase of this translation in terms of: (a) Its congruency with the original English ESS version. (b) Its comprehensibility in wide populations in their own country. Agreement with the translation of every phrase was evaluated using a 5-point Likert scale (0 = Complete disagreement. 1 = Disagreement. 2 = Neither agree nor disagree. 3 = Agreement. 4 = Complete agreement). The translation was approved when 80% of responses had a score ≥3. Researchers were asked to comment/propose changes in the translation for every phase that they score 2 or less. Using these comments, a revision of the translation was done. New Delphi process was performed until all phrases were approved. The final Spanish ESS version was used to develop a reverse translation (Spanish into English) by an independent professional translator (PC). Both original ESS and reverse translation English versions were compared for each participant to be sure that the meaning of each item was maintained.

2.2 | Sample size

Using the STATA software v16 (StataCorp LLC), considering an alpha of 0.05, a power of 80% and an expected mean difference of 0.45 points for a two-sample paired-means test as in the study by Taft

TABLE 1 Final ESS Spanish translation

Score *Score*	Weight loss (Kg) *Pérdida de peso (Kg)*	Dysphagia *Dificultad para tragar (dificultad en el paso de la comida)*	Retrosternal pain *Dolor en el pecho (retroesternal)*	Regurgitation *Sensación que la comida regresa hacia la boca (regurgita)*
0	None	None	None	None
0	*Nada*	*Nunca*	*Nunca*	*Nunca*
1	< 5Kg	Occasional	Occasional	Occasional
1	*<5Kg*	*Ocasionalmente*	*Ocasionalmente*	*Ocasionalmente*
2	5-10 Kg	Daily	Daily	Daily
2	*5-10 Kg*	*Diariamente*	*Diariamente*	*Diariamente*
3	>10 Kg	Each meal	Each meal	Each meal
3	*>10 Kg*	*En cada comida*	*En cada comida*	*En cada comida*

Note: Final Spanish version. Each cell shows original English version (plain) and final Spanish version (italic).

et al[29] (obtained by weight loss item), it was estimated that 41 people should be recruited. Furthermore, as ESS consists of four items, a minimum of 80 (item x 20) patients was required for reliability analysis.[30] Kaiser-Meyer-Olkin (KMO) was used to determine the appropriateness of sample size for the exploratory factor analysis. Finally, a power test based on Bonett's proposal was applied for determining sampling adequacy for reliability analysis (Cronbach's alpha).

2.3 | Patient selection

Patients were prospectively recruited among adults derived for a high-resolution esophageal manometry (HRM) due to dysphagia. Participant centers were as follows: Clínica Alemana de Santiago (Chile-CAS), Hospital Clínico de la Pontificia Universidad Católica de Chile (Chile-UC), Hospital San Ignacio-Pontificia Universidad Javeriana de Bogotá (Colombia), Fundación Favaloro (Argentina), Universidad Veracruzana (México), Hospital Clínico San Carlos (Spain-Madrid), and Hospital Universitari Germans Trias i Pujol-Badalona (Spain-Badalona). Recruitment occurred in two periods: between May-July 2016 and September 2018-December 2019. All the study forms were filled by patients in front of a participant researcher. Participant´s registered information included gender, age, ESS evaluated immediately prior to the HRM, HRM variables (integrated relaxation period (IRP) and distal contractile integral (DCI) and final diagnosis). All studies were analyzed using ManoView ESO 3.0 analysis software (Medtronic. Minneapolis, MN, USA). HRM diagnosis was assigned according to CC3.0. Patients with incomplete clinical information or previous esophagogastric procedures were excluded from the analysis.

Analysis was performed separately for achalasia and non-achalasia diagnosis. A subgroup of achalasia patients from Clínica Alemana and Pontificia Universidad Católica had a second manometric and clinical evaluation after POEM treatment. This subgroup was used to describe ESS response to treatment.

2.4 | Statistical analysis

Continuous variables are represented as the mean ± standard deviation. The Shapiro-Wilk test showed that ESS score did not follow a normal distribution ($P < .001$). Using the D'Agostino et al[31] test corrected by Royston,[32] it was determined that the distribution is asymmetric ($P < .001$). Therefore, non-parametric tests were applied to compare the ESS scores and to estimate the correlations between them. The Wilcoxon signed-rank test was used to evaluate changes in variables pre- and post-treatment and Mann-Whitney test to compare unmatched samples. Internal consistency was evaluated using Cronbach's α and split-half Guttman coefficient (for both scores, cutoffs were as follows: <0.5 = unacceptable. 0.5-0.7 = fair. >0.7 = acceptable).[33] Inter-item, item-total ESS, and ESS manometric variables correlations were determined using Spearman's rho. Principal component factor analysis (PCFA) with varimax rotation was used to perform a factor analysis of the correlation matrix. For determination of factors number, we used scree plots and Kaiser´s criterion of eigenvalue >1.0.[34,35] For determination of the convenience of deleting an item, we used the following criteria: (a) Item-total ESS correlation $r < .3$. (b) After deletion of the item, Cronbach´s $\alpha > .7$. A $P < .05$ was considered significant.

The determination coefficient (R^2) was estimated through a confirmatory factor analysis, quantifying the variable's percentage of variance, explained by the factor identified in the exploratory factor analysis. The correlation matrix used when executing the principal component factor analysis (PCFA) and confirmatory factorial analysis was constructed on the basis of polychoric correlations. Additionally, weighted least squares mean and variance adjusted method (WLSMV) was applied for estimating the confirmatory analysis, as it is appropriate when modeling ordinal data.

The statistical analysis was performed by STATA V.16 software (StataCorp LP) and MPLUS V.7.3 (StatModel).

3 | RESULTS

3.1 | Eckardt symptom score translation

Thirteen translators (eight gastroenterologist experts in esophageal diseases, one psychometrist, one advanced endoscopist, one general gastroenterologist, and two digestive surgeons) (29-61 years old) were recruited from six countries (Chile five, Argentina two, Spain two, Colombia two, Mexico one, and Ecuador one). The first round of the Delphi process occurred in April 2014 and consisted of nine sentences (four describing symptoms and five evaluating frequency/severity). 8/9 sentences met the criterion for acceptance. A second round took place in May 2014 and the revised sentence got acceptance. A reverse translation was independently performed. It was considered to have an excellent agreement with the original English version. Table 1 shows the final ESS Spanish version.

After its use in a total of 591 individuals (430 non-achalasia and 161 achalasia patients) from all the participant center's countries, no patient had any difficulty understanding and filling this final Spanish ESS version. All the forms were correctly and fully filled.

3.2 | Eckardt symptom score psychometric evaluation

3.2.1 | Non-achalasia

Four hundred and thirty patients (61.1% female) were recruited (Chile-CAS 149, Chile-UC 133, Spain-Badalona 70, Colombia 54, Mexico 19, and Spain-Madrid 5). Mean age was 52 ± 17 years old. All individuals suffered from dysphagia, with or without reflux symptoms. According to CC3.0, manometric results were as follows: normal 217 (50.4%), IEM: 145 (33.7%), absent contractility 30 (6.9%), EGJOO: 34 (7.9%), hypertensive peristalsis (Jackhammer's esophagus) 3 (0.6%), and distal esophageal spasm 1 (0.2%). Total ESS had a median of 3 (IQR: 2-5.25). Individual ESS items scored: dysphagia 1 (IQR: 0-2), regurgitation 1 (IQR: 0-2), chest pain 1 (IQR 0-1), weight loss 0 (IQR: 0-1) (Figure 1).

In the evaluation of reliability, ESS showed fair internal consistency, with Cronbach α = 0.57 and Guttman split-half coefficient (Lambda 2) = 0.58. All items showed a moderate item-total ESS correlation (Table 2). ESS dysphagia explained 61% total ESS variance, whereas ESS regurgitation and ESS chest pain explained less than 20% each (Table 2). The deletion of no item increased the Cronbach´s alpha to a level >0.7 (Table 2). Regarding validity, all inter-item correlations were significant, but very small in magnitude (rho range 0.10-0.41) (Table 3). There was a significant correlation between each item from the ESS and the BEDQ overall score (rho range 0.34 to 0.76), including a rho = 0.76 between ESS dysphagia and BEDQ total score (Table 3). No significant correlation was found between total or any ESS item and IRP nor DCI.

Using a scree plot and Kaiser´s criterion, PCFA suggested the existence of one factor, which explained only 47% variance of the overall ESS score. All factor loadings were between 0.4 and 0.5, and the KMO index varied between 0.7 and 0.6 (Table 4). The confirmatory factor analysis showed a maximum R^2 of .61 and a minimum of 0.15 for the dysphagia and chest pain items, respectively. This can be interpreted as the 61% and 15% of variance of the dysphagia and chest pain items, respectively, are explained by the underlying factor.

3.2.2 | Achalasia

Pretreatment

One hundred and sixty-one patients with achalasia were recruited (50.3% female). Mean age was 51 ± 17 years old. Using CC3.0, 27/161 (16.8%) corresponded to Achalasia I, 117/161 (72.7%) to Achalasia II, and 17/161 (10.5%) to Achalasia III. In all cases, dysphagia was the reason for referral. Each item median ESS was higher than non-achalasia patients: dysphagia 3 (IQR: 2-3) (P < .001 vs non-achalasia), regurgitation 2 (IQR: 1-3) (P < .001 vs non-achalasia), chest pain 1 (IQR: 0-2) (P = .003 vs non-achalasia), weight loss 2 (IQR 0-3) (P < .001 vs non-achalasia), and total ESS 7 (IQR: 4-9) (P < .001 vs non-achalasia).(Figure 1).

Eckardt symptom score showed a fair reliability, with a Cronbach´s α = 0.65 and Guttman split-half score (Lambda 2) = 0.66. Except for chest pain, all inter-item correlations were moderate, higher when compared to non-achalasia patients (Table 3). There was a significant correlation between each item from the ESS and the BEDQ overall score (rho range 0.35-0.51), and a rho = 0.63 between ESS total and BEDQ total score. The deletion of no item increased Cronbach´s alpha to a level >0.7. No significant correlation was found between total or any ESS item and IRP.

Using a scree plot and Kaiser´s criterion, PCFA suggested the existence of one factor, which explained 55% variance of the overall ESS score. All factor loadings were over 0.7 except for the chest pain item with a loading of 0.5. The KMO index varied between 0.7 and 0.6 (Table 4). The confirmatory factor analysis showed a maximum R^2 of 0.58 and a minimum of 0.14 for the weight loss and chest pain items, respectively. This can be interpreted as the 58% and 14% of variance of the weight loss and chest pain items, respectively, are explained by the underlying factor.

Postsurgical assessment

In 68/161 cases, post-treatment ESS was also available. All of them were treated with POEM and reevaluated after a mean of 4.3 months (range 2-11 months). Total and all individual ESS items showed a significant reduction after treatment (Table 5). There was a significant reduction in IRP (27.1 ± 12.6 vs 9.9 ± 4.1 mm Hg.; P < .001) and basal LES pressure (37.8 ± 15.1 vs 16.7 ± 8.0 mm Hg.; P < .001). No significant correlation could be found between any individual nor total ESS and delta IRP.

4 | DISCUSSION

In the present study, we present a Spanish ESS version which was produced after a careful translation process in a wide spectrum of

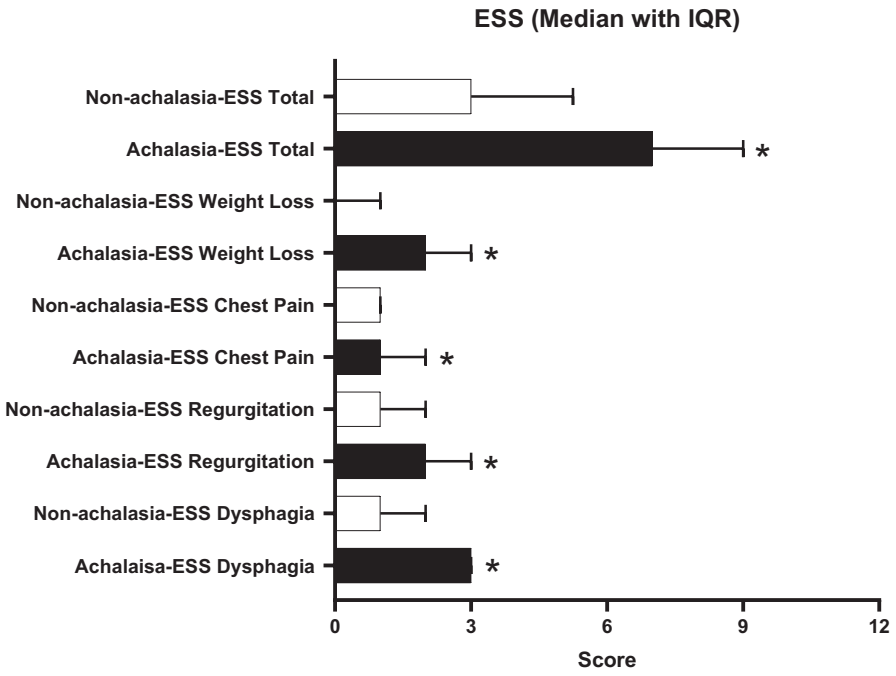

FIGURE 1 Eckardt symptom score (ESS) results according to manometric diagnosis (achalasia vs non-achalasia). Total = total ESS. *=$P < .05$

TABLE 2 Item statistics in different patients' samples

	Corrected item-total Eckardt symptom score (ESS) correlation	Square multiple correlation (R^2)	Alfa Cronbach if item deleted
Non-achalasia (n = 430). Alfa Cronbach 0.57			
ESS dysphagia	0.74	.61	0.48
ESS regurgitation	0.60	.16	0.41
ESS chest pain	0.55	.15	0.56
ESS weight loss	0.65	.39	0.53
Achalasia pretreatment (n = 161). Alfa Cronbach 0.65			
ESS dysphagia	0.70	.57	0.56
ESS regurgitation	0.77	.42	0.53
ESS chest pain	0.54	.14	0.68
ESS weight loss	0.77	.58	0.53

Spanish-speaking countries in Latin America and Spain. This reduces the potential effect of subtle local variations in the Spanish use, making this ESS version widely usable.

Our data show that ESS behaves similarly in patients without and with achalasia, with a fair reliability (Cronbach`s α of 0.57 and 0.65, respectively). These results are remarkably similar to Taft´s study,[19] suggesting that this suboptimal reliability is expected to occur in widely different populations. It could be hypothesized that this is a score inherent characteristic and is less influenced by translation or cultural differences. It implies that in clinical practice and research, ESS results should be interpreted with caution. For example, with a reliability of 0.57, there is 0.68 error variance (random error) in the scores (0.57 x 0.57 = 0.32; 1.00-0.32 = 0.68).[36] Our data also suggest that the ESS reliability cannot be improved by deleting specific items. However, it must be said that the score is able to detect clinical changes after treatment in the subgroup of achalasia, although in none of the subgroups analyzed there was a relationship with manometric variables.

Regarding the factors that explain the suboptimal reliability of the score, it can be inferred that one of them is the low number of items, which is evident when remembering Cronbach's α equation calculation formula.[37] Nevertheless, this small number of items makes ESS very easy to be widely applicable. Also inferred from the formula, the slightly better performance in achalasia can be explained by a better inter-item correlation in this scenario. ESS was designed specifically for achalasia, and the symptoms evaluated are typically present in this condition, which is not always the case in non-achalasia. This can be specifically important in case of normal HRM studies. It should also be noted that a moderate correlation with BEDQ was observed, especially in the dysphagia item.

Dysphagia is the most frequent achalasia symptom[5] and the main determinant of symptom severity and treatment response. Nevertheless, ESS only evaluates its frequency, not its severity and it makes no difference regarding different food consistencies. Adaptive behaviors and food impaction events are also not included. Regurgitation is also a frequent achalasia complain, but it is widely non-specific. This could be relevant as an achalasia patient can have a post-treatment ESS = 3 only due to gastro-esophageal reflux disease (GERD), which is a frequent treatment complication.[38,39] Weight loss is referred by a minority of achalasia patients.[5] Its criterion validity can be questioned, as it shows a very low correlation with other items in our and Taft ´s studies.[29] Patel et al could not find any difference in any other ESS symptoms when comparing achalasia patients with and without weight loss.[40] On the other hand, only 1/68 of our patients

TABLE 3 Inter-item correlations of Eckardt symptom score (ESS) score and the Brief Esophageal Dysphagia Questionnaire (BEDQ)

	ESS dysphagia	ESS regurgitation	ESS chest pain	ESS weight loss	BEDQ total score
Non-achalasia (n = 430)					
ESS dysphagia	1.0				
ESS regurgitation	0.20*	1.0			
ESS chest pain	0.26*	0.18*	1.0		
ESS weight loss	0.41*	0.22 *	0.10*	1.0	
BEDQ total score	0.76**	0.34**	0.35**	0.50**	1.0
Achalasia pretreatment (n = 161)					
ESS dysphagia	1.0				
ESS regurgitation	0.37*	1.0			
ESS chest pain	0.35*	0.16*	1.0		
ESS weight loss	0.37*	0.54*	0.16	1.0	
BDEQ total score	0.51**	0.51**	0.35*	0.40**	1.0

*=$P < .05$.

**=$P < .001$.

TABLE 4 Summary of principal component factor analysis (component matrix) and reliability for ESS in different patient samples

	Factor loading	Kaiser-Meyer-Olkin (KMO)	R^2
Non-achalasia (n = 430). Cronbach`s alpha 0.57			
ESS dysphagia	0.58	0.6	.61
ESS regurgitation	0.45	0.7	.16
ESS chest pain	0.42	0.6	.15
ESS weight loss	0.54	0.6	.39
% variance	47		
Achalasia pretreatment (n = 161). Cronbach's alpha 0.65			
ESS dysphagia	0.80	0.7	.57
ESS regurgitation	0.80	0.7	.42
ESS chest pain	0.54	0.6	.14
ESS weight loss	0.79	0.7	.58
% variance	55		

TABLE 5 Change in ESS pre- and post-treatment in achalasia patients (n = 68)

Variable	Pretreatment median (p25-p75)	Post-treatment median (p25-p75)	*P value*
ESS total score	7.5 (6-19)	1 (0-2)	<.001
ESS dysphagia	2 (1-3)	0 (0-0)	<.001
ESS regurgitation	3 (2-3)	0 (0-1)	<.001
ESS chest pain	1 (1-2)	0 (0-1)	<.001
ESS weight loss	2 (1-3)	0 (0-0)	<.001

Note: Median pre- and post-treatment ESS items in achalasia. P25: percentile 25. P75 = percentile 75.

showed some weight loss while 67/68 gained weight after treatment. This last phenomenon could not be detected using ESS. Nevertheless, ESS weight loss did not meet our criteria for elimination convenience.

Using manometric and fluoroscopic variables, no good correlation between ESS and objective criteria of esophageal dysfunction has been previously demonstrated. Krieger-Grubel et al found no difference in impedance nor timed barium swallow parameters when comparing achalasia patients with ESS > or <3 after treatment.[41] Carlson et al[42] found no difference in responder rate (defined as ESS < 3) when comparing patients with complete vs incomplete bolus transit after treatment. This lack of correlation between bolus perception/dysphagia and objective measurements has also been shown using other ways of sensation evaluation (like Likert scale) in healthy and symptomatic individuals.[43-46] This

highlights the fact that dysphagia is a complex phenomenon that goes far beyond our current methods to evaluate esophageal function. Nevertheless, ESS has only shown moderate correlation with other esophageal symptoms scores, with correlation coefficients around 0.5.[29,41] So, it is still possible to better explain dysphagia in terms of objective measurements, using other tools to evaluate symptoms. In 2005, Urbach et al described a disease-specific dysphagia score (Measure of Achalasia Disease Severity, MADS).[47] It is considered to have robust measurement properties,[48] as it shows significant reliability (Cronbach´s α = 0.83) and good data-to-model fit. Nevertheless, it has hardly been used in the literature. More recently, Taft et al described the Brief Esophageal Dysphagia Questionnaire (BEDQ). It is a 10-item generic dysphagia score that showed very good reliability and construct validity.[49] Interestingly, BEDQ showed significant correlations with some manometric variables in Taft´s and ESS dysphagia item in our study. Future studies should specifically address if these new scores help us to better study esophageal dysfunction.

Latin America is a multicultural society that inherited its language from Spain. Although there are variations in multiple terms, the core is highly preserved for common diseases, as well as basic descriptive concepts such as those used in the EES.[50] The major variations are identified in popular constructs of syndromes that despite sharing proper folk names, the variety of symptoms and popular explanations can be very wide.[51] To overcome these difficulties, clinicians from each center participated in the process of constructing the Spanish version of the EES considering these various inter-cultural variations. Despite this effort, part of the results may be due to potential problems in linguistic and cultural factors.

Our study has some weakness: We had only a few manometric variables and no radiological studies or other symptoms scores to correlate with ESS, so we could not evaluate in depth construct validity. Measurements were not repeated to evaluate temporal stability. Nevertheless, we have enough number of patients from different areas of the Spanish-speaking spectrum to evaluate ESS reliability.

In conclusion, we have developed and validated a Spanish ESS version, suitable to wide use in different regions of the world using Spanish as the native language. In our population, ESS showed fair reliability for the evaluation of patients with any causes of dysphagia. Our results highlight the need for development and psychometric validation of new dysphagia scoring tools.

ACKNOWLEDGMENTS

Translation team: Daniel Cisternas, Arnoldo Riquelme, Arturo Valle, Allan Sharp, Fernando Pimentel, Antonio Ruiz de Leon, Jose Remes-Troche, Ramiro Coello, Jordi Serra, Albis Hani, Ana María Leguizamo, Claudio Bilder, Andres Ditranto, and Paula Cisternas (independent professional translator). Administrative assistance: Lorena Olivero and Soledad Quesada.

CONFLICT OF INTERESTS

No competing interests declared.

AUTHOR CONTRIBUTIONS

DC contributed to conception and study design, recruitment, data acquisition, analysis and interpretation, manuscript drafting, editing, critical revision, and final approval; HM contributed to study design, recruitment, data acquisition, analysis and interpretation, manuscript drafting, editing, critical revision, and final approval; AR contributed to data analysis and interpretation, manuscript critical revision, and final approval; OP and EFL contributed to data analysis and interpretation, and manuscript final approval; AV and RM contributed to patient recruitment and treatment and manuscript final approval; AH, AAF, AL, CB, AD, JRT, ARL, PJD, IM, and JS contributed to patient recruitment, manuscript critical revision, and final approval; all authors approved the final manuscript version.

ORCID

Daniel Cisternas https://orcid.org/0000-0003-0909-5192
Hugo Monrroy https://orcid.org/0000-0002-4740-3945
Jordi Serra https://orcid.org/0000-0003-2120-6270

REFERENCES

1. Podas T, Eaden J, Mayberry M, Mayberry J. Achalasia: a critical review of epidemiological studies. *Am J Gastroenterol*. 1998;93(12):2345–2347.
2. Lynch KL, Pandolfino JE, Howden CW, Kahrilas PJ. Major complications of pneumatic dilation and Heller myotomy for achalasia: single-center experience and systematic review of the literature. *Am J Gastroenterol*. 2012;107(12):1817–1825.
3. Nabi Z, Reddy DN, Ramchandani M. Adverse events during and after per-oral endoscopic myotomy: prevention, diagnosis, and management. *Gastrointest Endosc*. 2018;87(1):4–17.
4. Haito-Chavez Y, Inoue H, Beard KW, et al. comprehensive analysis of adverse events associated with per oral endoscopic myotomy in 1826 patients: An International Multicenter Study. *Am J Gastroenterol*. 2017;112(8):1267–1276.
5. Fisichella PM, Raz D, Palazzo F, Niponmick I, Patti MG. Clinical, radiological, and manometric profile in 145 patients with untreated achalasia. *World J Surg*. 2008;32(9):1974–1979.
6. Eckardt VF, Aignherr C, Bernhard G. Predictors of outcome in patients with achalasia treated by pneumatic dilation. *Gastroenterology*. 1992;103(6):1732–1738.
7. Boeckxstaens GE, Annese V, Varannes SBD, et al. Pneumatic dilation versus laparoscopic Heller's myotomy for idiopathic achalasia. *N Engl J Med*. 2011;364(19):1807–1816.
8. Akintoye E, Kumar N, Obaitan I, Alayo QA, Thompson CC. Peroral endoscopic myotomy: a meta-analysis. *Endoscopy*. 2016;48(12):1059–1068.
9. Marano L, Pallabazzer G, Solito B, et al. Surgery or peroral esophageal myotomy for achalasia: a systematic review and meta-analysis. *Medicine (Baltimore)*. 2016;95(10):e3001.
10. Kumagai K, Kjellin A, Tsai JA, et al. Toupet versus Dor as a procedure to prevent reflux after cardiomyotomy for achalasia: results of a randomised clinical trial. *Int J Surg*. 2014;12(7):673–680.
11. Carlson DA, Lin Z, Kahrilas PJ, et al. High-resolution impedance manometry metrics of the esophagogastric junction for the assessment of treatment response in achalasia. *Am J Gastroenterol*. 2016;111(12):1702–1710.
12. Pandolfino JE, de Ruigh A, Nicodème F, Xiao Y, Boris L, Kahrilas PJ. Distensibility of the esophagogastric junction assessed with the Functional Lumen Imaging Probe (FLIP™) in achalasia patients. *Neurogastroenterol Motil*. 2013;25(6):496–501.

13. Ren Y, Tang X, Chen Y, et al. Pre-treatment Eckardt score is a simple factor for predicting one-year peroral endoscopic myotomy failure in patients with achalasia. *Surg Endosc*. 2017;31(8):3234–3241.
14. Rohof WOA, Bredenoord AJ. Chicago classification of esophageal motility disorders: lessons learned. *Curr Gastroenterol Rep*. 2017;19(8):37.
15. Kahrilas PJ, Bredenoord AJ, Fox M, et al. The Chicago Classification of esophageal motility disorders, v3. 0. *Neurogastroenterol Motil*. 2015;27(2):160-174.
16. Kahrilas PJ, Bredenoord AJ, Carlson DA, Pandolfino JE. Advances in management of esophageal motility disorders. *Clin Gastroenterol Hepatol*. 2018;16(11):1692–1700.
17. Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4:79.
18. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. *Am J Health Syst Pharm*. 2008;65(23):2276–2284.
19. Taft TH, Carlson DA, Triggs J, et al. Evaluating the reliability and construct validity of the Eckardt symptom score as a measure of achalasia severity. *Neurogastroenterol Motil*. 2018;30(6):e13287.
20. US Department of Health and Human Services FDA Center for Drug Evaluation and Research, US Department of Health and Human Services FDA Center for Biologics Evaluation and Research, US Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4:1–20.
21. Jain A, Baker JR, Chen JW. In ineffective esophageal motility, failed swallows are more functionally relevant than weak swallows. *Neurogastroenterol Motil*. 2018;30(6):e13297.
22. Song BG, Min YW, Lee H, et al. Clinicomanometric factors associated with clinically relevant esophagogastric junction outflow obstruction from the Sandhill high-resolution manometry system. *Neurogastroenterol Motil*. 2018;30(3):e13221.
23. Khashab MA, Messallam AA, Onimaru M, et al. International multicenter experience with peroral endoscopic myotomy for the treatment of spastic esophageal disorders refractory to medical therapy (with video). *Gastrointest Endosc*. 2015;81(5):1170–1177.
24. Kristensen HO, Bjerregaard NC, Rask P, Mortensen FV, Kunda R. Peroral endoscopic myotomy (POEM) for nutcracker esophagus. Three cases with 12 months follow-up. *Scand J Gastroenterol*. 2014;49(11):1285–1289.
25. Khan MA, Kumbhari V, Ngamruengphong S, et al. Is POEM the answer for management of spastic esophageal disorders? A systematic review and meta-analysis. *Dig Dis Sci*. 2017;62(1):35–44.
26. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000;32(4):1008–1015.
27. Dalkey N, Helmer O. An experimental application of the DELPHI method to the use of experts. *Manag Sci*. 1963;9(3):458–467.
28. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm*. 2016;38(3):655–662.
29. Taft TH, Carlson DA, Triggs J, et al. Evaluating the reliability and construct validity of the Eckardt symptom score as a measure of achalasia severity. *Neurogastroenterol Motil*. 2018;30(6):e13287.
30. Cabrera-Nguyen P. Author guidelines for reporting scale development and validation results in the Journal of the Society for Social Work and Research. *J Soc Soc Work Res*. 2010;1(2):99–103.
31. D'agostino RB, Belanger A, D'agostino RB. A suggestion for using powerful and informative tests of normality. *Am Statist*. 1990;44(4):316–321.
32. Royston PJSTB. Comment on sg3. 4 and an improved D'Agostino test. 1992;1(3).
33. Nunnally J, CaBIH. *Psychometric Theory*. 3 ed: McGraw-Hill: New York; 1994.
34. Kaiser H. The application of electronic computers to factor analysis. *Educ Psychol Measur*. 1960;20(1):141–151.
35. Fabrigar L, Wegener D, MacCallum R, Strahan E. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods*. 1999;4:272.
36. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53.
37. Ritter NL. Understanding a widely misunderstood statistic. Cronbach's.2010.
38. Kahrilas PJ, Katzka D, Richter JEJG. Clinical practice update: the use of per-oral endoscopic myotomy in achalasia: expert review and best practice advice from the AGA Institute. *Gastroenterology*. 2017;153(5):1205–1211.
39. Kumbhari V, Familiari P, Bjerregaard N, et al. Gastroesophageal reflux after peroral endoscopic myotomy: a multicenter case–control study. *Endoscopy*. 2017;49(07):634–642.
40. Patel DA, Naik R, Slaughter JC, Higginbotham T, Silver H, Vaezi MF. Weight loss in achalasia is determined by its phenotype. *Dis Esophagus*. 2018;31(9):doy046.
41. Krieger-Grübel C, Tutuian R, Borovicka J. Correlation of esophageal clearance and dysphagia symptom assessment after treatment for achalasia. *United Eur Gastroenterol J*. 2016;4(1):55–61.
42. Carlson DA, Beveridge CA, Lin Z, et al. Improved assessment of bolus clearance in patients with achalasia using high-resolution impedance manometry. *Clin Gastroenterol Hepatol*. 2018;16(5):672–680. e671.
43. Chen C, Yi CJN. Clinical correlates of dysphagia to oesophageal dysmotility: studies using combined manometry and impedance. *Neurogastroenterol Motil*. 2008;20(6):611–617.
44. 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(12):1292-e1337.
45. Xiao Y, Kahrilas PJ, Nicodème F, Lin Z, Roman S, Pandolfino JE. Lack of correlation between HRM metrics and symptoms during the manometric protocol. *Am J Gastroenterol*. 2014;109(4):521.
46. Cisternas D, Scheerens C, Omari T, et al. Anxiety can significantly explain bolus perception in the context of hypotensive esophageal motility: results of a large multicenter study in asymptomatic individuals. *Neurogastroenterol Motils*. 2017;29(9):e13088.
47. Urbach DR, Tomlinson GA, Harnish JL, Martino R, Diamant NE. A measure of disease-specific health-related quality of life for achalasia. *Am J Gastroenterol*. 2005;100(8):1668.
48. Patel DA, Sharda R, Hovis KL, et al. Patient-reported outcome measures in dysphagia: a systematic review of instrument development and validation. *Dis Esophagus*. 2017;30(5):1.
49. Taft TH, Riehl M, Sodikoff JB, et al. Development and validation of the brief esophageal dysphagia questionnaire. *Neurogastroenterol Motil*. 2016;28(12):1854–1860.
50. Weller SC, Baer RD. Intra- and intercultural variation in the definition of five illnesses: AIDS, diabetes, the common cold, Empacho, and Mal de ojo. *Cross Cultural Res*. 2001;35(2):201–226.
51. Weller SC, Pachter LM, Trotter RT, et al. Empacho in four Latino groups: A study of intra- and inter-cultural variation in beliefs. *Med Anthropol*. 1993;15(2):109–136.

How to cite this article: Cisternas D, Monrroy H, Riquelme A, et al. Fair reliability of eckardt scores in achalasia and non-achalasia patients: Psychometric properties of the eckardt spanish version in a multicentric study. *Neurogastroenterol Motil*. 2020;32:e13827. https://doi.org/10.1111/nmo.13827