Qual Life Res (2015) 24:1597–1605
DOI 10.1007/s11136-014-0886-2

Confirmatory factor analysis of the Patient Assessment of Constipation-Symptoms (PAC-SYM) among patients with chronic constipation

Luca Neri · Paul Maurice Conway · Guido Basilisco · Laxative Inadequate Relief Survey (LIRS) Group

Accepted: 3 December 2014 / Published online: 19 December 2014

Abstract

Background and aim PAC-SYM is widely adopted to asses constipation severity. However, it has been validated in a small sample, few items have been included based on expert opinion and not on empirical grounds, and its factor structure has never been replicated. We aimed at evaluating the psychometric properties of PAC-SYM in patients with chronic constipation.

Methods We enrolled 2,203 outpatients with chronic constipation in two waves. We used wave I sample to test the psychometric properties of the PAC-SYM and wave II sample to cross-validate its factor structure, to assess criterion validity, responsiveness to clinical change, and its minimal clinically important difference.

Results Only a minority of patients reported any rectal tearing (38 %). Deletion of such item leads to a 11-item version (M:PAC-SYM). The remaining items in the rectal domain were moderately correlated with the stool domain. Exploratory factor analysis and confirmatory factor analysis revealed a bifactor structure with two subscales (stool and abdominal symptoms) and a general severity factor. The M:PAC-SYM demonstrated excellent reliability, moderate correlation with SF-12 and treatment satisfaction ($r = 0.28$–0.45), discrimination across Rome III criteria for functional constipation and abdominal pain, and responsiveness to clinical change ($\beta = -0.49$; $\omega^2 = 0.25$). M:PAC-SYM minimal clinically important difference was 0.24.

Conclusion Our analysis shows that the rectal domain may not represent a relevant cluster of symptoms for patients with chronic constipation. We developed a modified version of the PAC-SYM which might better represent symptom severity of most patients seeking care in gastroenterology referral centers.

Keywords Constipation severity · Quality of life · Chronic non-organic constipation · Chronic constipation

Introduction

Assessment of patient-reported outcomes is a key issue in functional gastrointestinal disorders since the diagnosis of these conditions, the assessment of disease severity and treatment outcomes is based on symptoms. Multi-item symptoms questionnaires are often used to assess the effectiveness of medications and to predict the improvement of quality of life in randomized controlled trials. However, these questionnaires are prone to a number of bias which may result in fallacious conclusions with regard to treatment efficacy and tolerability [1].

The Patient Assessment of Constipation-Symptoms (PAC-SYM) is a 12-item self-report questionnaire

Members of the LIRS Study Group are listed in Electronic Supplementary Material and should be regarded as collaborators for indexing purposes.

Electronic supplementary material The online version of this article (doi:10.1007/s11136-014-0886-2) contains supplementary material, which is available to authorized users.

L. Neri (✉)
Dipartimento di Scienze Cliniche e di Comunità, Università degli Studi di Milano, Via San Barnaba, 8, 20122 Milan, Italy
e-mail: luca.neri@unimi.it

P. M. Conway
Department of Psychology, University of Copenhagen, Copenhagen, Denmark

G. Basilisco
Gastroenterology and Endoscopy Unit, Fondazione IRCCS "Ca' Granda – Ospedale Maggiore Policlinico", Milan, Italy

subdivided in three symptom subscales (i.e., abdominal, stool, and rectal). The initial psychometric evaluation provided evidence that the PAC-SYM is a reliable and valid instrument assessing the severity of constipation in adult patients, in which it can detect clinically meaningful changes over time and distinguish between responders and non-responders to treatment [2].

The PAC-SYM has recently been used in the integrated analysis of three double-blind placebo-controlled trials with prucalopride in women who reported inadequate relief from laxatives at trial entry [3]. In line with the results in the original studies [4–6], the effect of prucalopride was smaller for rectal than for abdominal and stool symptoms subscales and the proportion of patients with a PAC-SYM severity score >2 at baseline was 50 and 71 % for abdominal and stool symptoms, but only 15 % for rectal symptoms; in the analysis of individual item scores, the smallest effect of prucalopride was observed for two items of the rectal symptom subscale, namely "rectal burning" and "rectal bleeding/tearing." Items "rectal bleeding/tearing" and "rectal burning" have been included in the original questionnaire [2] based on clinical judgment, but not on empirical findings. In fact, such symptoms were fairly infrequent in the validation sample [2], and thus captured non-relevant complains for the majority of patients.

Although some evidence of invariance of PAC-SYM in different samples has been obtained in older adults in long-term care facilities [2] and in patients with low back pain and opioid-induced constipation [7], no attempts to replicate the factor structure, responsiveness to clinical change and construct validity of the PAC-SYM questionnaire have been carried out in patients with chronic constipation. Replication addresses how well factors generalize across samples drawn from the same population [8]. Hence, we aimed at evaluating the psychometric properties of the PAC-SYM questionnaire with its subscales and characterizing its relationship with key outcomes such as quality of life and treatment satisfaction.

Methods

The present study is a part of the Laxative Inadequate Relief Survey (LIRS), aiming at evaluating quality of life, treatments satisfaction, activity impairment, and health care utilization among patients with chronic constipation in 39 Italian referral centers for functional gastrointestinal disorders. The study consisted of two waves: In LIRS I, a cross-sectional survey, 878 consecutive outpatients were enrolled from September through December 2011. In LIRS II, a repeated measure survey, 1,325 outpatients from the same centers were enrolled, of whom 45 were assessed twice or more from March 2012 to May 2013.

Patients' eligibility was ascertained by a gastroenterologist during an outpatient clinical examination in both waves. Patients reporting at least two of the Rome III criteria for functional constipation [9] were included after exclusion of any organic cause of gastrointestinal symptoms. Abdominal pain or discomfort lasting at least 3 days/month in the past 3 months, a symptoms associated with irritable bowel syndrome with constipation (IBS-C), was not used as criterion of exclusion given the wide overlap between IBS-C and functional constipation [10]. A gastroenterologist recorded three possible features of abdominal pain in a data collection form: (1) relieved by defecation; (2) changed after the meal; and (3) represented the most bothersome complain for the patient. The same gastroenterologist recorded patients' age, sex, BMI, smoking habit, daily intake of water, time since constipation onset, pregnancies, difficult deliveries, previous abdominal and extra-abdominal surgical intervention, concomitant diseases, and treatment regimens during a regular outpatient visit.

We matched clinical data with a self-administered questionnaire completed by each patient.

Measures

The questionnaires included the PAC-SYM [2], the Italian version of the RAND SF-12 [11] and the Treatment Satisfaction Questionnaire for Medication (TSQM-v2) [12]. The PAC-SYM questionnaire consists of 12 items in three domains: abdominal symptoms [abdominal, ABD: 4 items: (1) discomfort in your abdomen, (2) pain in your abdomen, (3) bloating in your abdomen, and (4) stomach cramps], rectal symptoms [rectal, REC: 3 items: (5) painful bowel movements, (6) rectal burning during or after a bowel movement, (7) rectal bleeding or tearing during or after bowel movement] and stool symptoms [stool, STO: 5 items: (8) incomplete bowel movement like you did not finish, (9) bowel movement that were too hard, (10) bowel movement that were too small, (11) straining or squeezing to try to pass bowel movements, (12) feeling like you had to pass a bowel movement but you could not]. Ratings occur along a five-point Likert scale (form 0 = absence of symptoms to 4 = very severe). The TSQM-v2 assesses patients' satisfaction with treatment effectiveness, side effects, and convenience. Summary scores are calculated using a 0–100 scale, with higher scores corresponding to higher treatment satisfaction [12].

Additionally, the LIRS II questionnaire included a section inquiring about therapy switching in the past month. Patients were asked to rate clinical change after therapy switch with a bipolar global rating of change scale (GRC, 15 point, from −7 = extremely deteriorated to +7=extremely improved) [13]. It has been suggested that the

minimal clinically important difference (MCID) for the GRC scale is two point [13]. For this reason, we classified patients' ratings as follows: −7 to −5 = very deteriorated; −4 to −2: deteriorated; −1 to 1 = unchanged; +2 to +4 = improved; and +5 to +7 = very improved.

Analysis

All analyses have been conducted with SAS 9.2®. We used the LIRS I sample to test basic psychometric properties and construct validity of the PAC-SYM questionnaire (item ceiling and floor effect, convergent and divergent validity, factor structure, and internal consistency) and the LIRS II sample to test criterion validity (association with other relevant self-reported outcomes, i.e., treatment satisfaction, quality of life, and global rating of change scores) and assess the minimal clinically important difference for the PAC-SYM scores. Using principal component analysis, the original validation study revealed the existence of three subscales of symptoms making up the PAC-SYM (abdominal, rectal and stool) [2]. Hence, we adopted a multi-trait approach to test items convergent and discriminant validity across the proposed three subscales. We generated an inter-item correlation matrix: All correlations ≤ 0.35 among items belonging to the same hypothesized scale suggest lack of convergent validity (i.e., items might not represent the same construct). Internal consistency was evaluated by computing Cronbach's α. We also tabulated the item-total correlations, indicating the association between individual items and the total scores (all subscales). Correlations between sub-scale scores and any item which do not belong to that scale ≥ 0.40 suggest lack of divergent validity (i.e., items might not discriminate across the hypothesized constructs). We then carried out a confirmatory factor analysis (CFA) and compared the fit of five different models thought to explain patients' responses to PAC-SYM items.

Model 1 (*Appendix 1 of Electronic Supplementary Material, panel 1*): Since PAC-SYM total score is considered an overall measure of constipation severity and it is often used as outcome measure in RCTs or predictor measure in observational studies [3–6], Model 1 hypothesized a single-factor structure, i.e., all items belonging to a general severity scale with no abdominal, rectal, or stool subscales.

Model 2 and 3 (*Appendix 1, panels 2–3*): To account for the three-dimensional structure empirically observed in the validation study by Frank et al. [2], we tested both an uncorrelated and an inter-correlated three-factor structure (i.e., abdominal, rectal, and stool factors).

Model 4 (*Appendix 1, panel 4*): We tested the hypothesis that a higher-order PAC-SYM factor can account for the three lower-order factors, i.e., a general severity explaining three lower-order factors to which all items belong.

Model 5: We tested a bi-factor structure [14] which assumes that the three sub-scale scores add unique information beyond (i.e., after adjustment for) the general severity factor. This model would reconcile both the empirical factor structure and current use in clinical research.

Factor loadings were required to be >0.3 and statistically significant ($P < 0.05$) for each item to be considered as an adequate indicator of the respective hypothesized construct. We adopted the root mean square error of approximation (RMSEA), the goodness of fit index (GFI), and the normed χ^2 index for evaluating model fit: RMSEA < 0.10; GFI ≥ 0.90 and normed $\chi^2 < 5$ are considered to indicate acceptable fit [15]. To compare the fit of the hypothesized models, we used the Akaike information criteria (AIC), the Swartz Bayesian criteria (SBC), and the χ^2 test (the latter for nested models). Lower values of the AIC and SBC indicate that a given model has a better fit compared with a competing one, taking also criteria of parsimony into account.

We calculated Spearman's correlations to evaluate the association between PAC-SYM score and other relevant outcome measures. We evaluated whether PAC-SYM score and its subscales discriminate across clinically different subgroups (defined by abdominal pain and items of the Rome III criteria) by reporting absolute difference and effect size (Cohen's d or ω^2 where appropriate). For large samples, the use of conventional p value thresholds is of little meaning, while clinical relevance is to be preferred as decisional criterion. We assessed PAC-SYM responsiveness to change by evaluating its ability to discriminate across patients reporting improvement, stability, or deterioration on the GRC scale. The minimal clinical important difference (MCID) for PAC-SYM score was calculated by using a random intercept model to evaluate longitudinal variation in PAC-SYM score and GRC scale in 45 patients who were repeatedly assessed in the LIRS II survey. PAC-SYM score was included in the model as time-varying covariate. The MCID corresponded to the PAC-SYM coefficient estimate associated with a two-point increase in the GRC scale. This anchor-based method was compared with a distribution-based method to derive MCID: According to distribution-based criteria, the MCID for a patient-reported outcome ranges between 0.2 and 0.5 effect size [16–22].

Results

Study sample

Characteristics of patients enrolled in the LIRS I and LIRS II studies are described in Table 1. LIRS I and LIRS II samples were substantially similar relative to patients'

Table 1 Characteristics of patients with chronic constipation enrolled in the LIRS I and LIRS II studies and in the whole sample

Characteristics	Mean (SD) or *N* (%)		
	Whole sample, *N* = 2,203	LIRS I, *N* = 878	LIRS II, *N* = 1,325
Mean age (years)	50.1 (16.7)	50.3 (16.6)	49.9 (16.9)
Women	1,808 (82.1)	706 (80.4)	1,102 (83.2)
Employed	1,090 (49.5)	370 (42.1)	720 (54.3)
Rome III criteria			
Lumpy/hard stools	1,638 (74.4)	659 (75.1)	979 (73.9)
Incomplete evacuation	1,604 (72.8)	650 (74.0)	954 (72.1)
Obstruction	346 (40.4)	346 (40.4)	507 (38.3)
Manual maneuvers	539 (24.5)	220 (25.1)	319 (24.1)
<3 Defecations/week	1,501 (68.2)	565 (64.4)	936 (70.7)
Strain	1,812 (82.3)	723 (82.4)	1,089 (82.3)
Mean time since disease onset (years)	11.9 (13.7)	17.3 (15.0)	7.0 (10.2)
Abdominal pain	369 (16.8)	149 (17.0)	220 (16.6)
Therapy			
Diet/other	266 (12.2)	125 (14.6)	141 (10.7)
Bulking/osmotic	157 (7.2)	43 (5.0)	114 (8.6)
Stimulant/herbal	113 (5.2)	52 (6.1)	61 (4.6)
Enema	58 (2.7)	23 (2.7)	35 (2.6)
Multi-drug	1,352 (62.0)	563 (65.7)	789 (59.6)
None	107 (4.9)	51 (5.9)	56 (4.2)
Prucalopride[a]	–	N/A[a]	151 (11.4)
Water >1 L/day	1,325 (60.3)	537 (61.7)	788 (59.5)
Current smoker (Y)	806 (36.6)	325 (38.1)	477 (36.0)
Comorbidity index	2.4 (2.0)	2.5 (2.1)	2.3 (1.9)
BMI	23.6 (4.1)	23.8 (4.0)	23.5 (4.3)

Data are mean (SD) or *N* (%)

[a] Prucalopride was not available in Italy during LIRS I

characteristics. The average age was 50 years (SD 16.7) and 80 % of patients were women. In both samples, the most prevalent complain was straining followed by reports of lumpy or hard stools and sensation of incomplete evacuation in more than one quarter of defecations. Half of the patients had more than two comorbid medical conditions. Most patients received multiple therapies for constipation in the 3 months prior to the interview.

Ceiling and floor effect

PAC-SYM items have been completed in full by 97.5 % of patients in the LIRS I sample. Ceiling (i.e., scores at the top of the scale) and floor (i.e., scores at the bottom of the scale) effects for the total score were negligible. On the contrary, we observed a sizeable floor effect in the items concerning the rectal domain (items 5–7) and in item 4 (Table 2).

Table 2 Ceiling and floor effect for the individual items of the PAC-SYM questionnaire

Items	Ceiling (%)	Floor (%)	Mean (SD)
1. Discomfort in your abdomen	5.54	15.64	1.77 (1.10)
2. Pain in your abdomen	3.94	25.06	1.47 (1.13)
3. Bloating in your abdomen	14.25	7.60	2.23 (1.14)
4. Stomach cramps	2.17	50.30	0.91 (1.09)
5. Painful bowel movement	4.58	37.76	1.24 (1.21)
6. Rectal burning during or after a bowel movement	5.00	37.26	1.23 (1.21)
7. Rectal tearing or bleeding after a bowel movement	2.86	61.70	0.71 (1.08)
8. Incomplete bowel movement, like you didn't "finish"	11.91	10.97	2.09 (1.16)
9. Bowel movement that were too hard	13.70	11.10	2.20 (1.18)
10. Bowel movement that were too small	9.28	19.29	1.84 (1.24)
11. Straining or squeezing to try to pass bowel movements	20.43	7.08	2.44 (1.16)
12. Feeling like you had to pass a bowel movement you couldn't	11.11	24.35	1.76 (1.33)

Ceiling and floor effects represent the proportion of patients reporting the highest possible and lowest possible score for each item. Mean values for the single items are also reported

Convergent validity

Inter-item correlations ($r_{3–4} = 0.33$, $r_{5–7} = 0.31$, $r_{8–9} = 0.34$) reveal that items 3 and 4, 5 and 7, 8 and 9 might not represent the same constructs (Table 3).

Divergent validity

Both the correlations of items 5 (painful bowel movement) and 6 (rectal burning during or after a bowel movement) with the STO score were $r = 0.40$, thus suggesting that answers to these questions thought to belong to the REC scale might be partially explained by defecation-/stool-related problems.

Internal consistency

The standardized Cronbach's α of each subscale exceeded $\alpha > 0.70$ (ABD 0.80; REC 0.72; STO 0.80). Standardized α slightly increased after the deletion of item 4 ($\alpha = 0.81$) and item 7 ($\alpha = 0.74$) from the scales to which they

Table 3 Multi-trait analysis of item convergent and divergent validity

Dimension	Item	ABD				REC				STO				Total scores		
		1	2	3	4	5	6	7	8	9	10	11	12	ABD	REC	STO
ABD	1	–												**0.84**	0.20	0.31
	2	**0.69**	–											**0.83**	0.26	0.30
	3	**0.60**	**0.50**	–										**0.77**	0.23	0.32
	4	**0.40**	**0.47**	0.33	–									**0.69**	0.32	0.28
REC	5	0.27	0.38	0.25	0.34	–								0.38	**0.80**	**0.40**
	6	0.19	0.20	0.21	0.29	**0.58**	–							0.27	**0.86**	**0.40**
	7	0.01	0.05	0.08	0.13	0.31	**0.49**	–						0.08	**0.72**	0.22
STO	8	0.32	0.28	0.32	0.25	0.32	**0.35**	0.19	–					0.36	0.36	**0.70**
	9	0.14	0.14	0.16	0.15	0.28	0.28	0.16	0.34	–				0.18	0.30	**0.75**
	10	0.20	0.18	0.23	0.19	0.25	0.23	0.18	**0.37**	**0.57**	–			0.25	0.28	**0.74**
	11	0.22	0.23	0.23	0.18	0.33	0.34	0.16	**0.47**	**0.59**	**0.44**	–		0.26	0.35	**0.78**
	12	0.27	0.29	0.28	0.28	0.32	0.30	0.15	**0.47**	**0.36**	**0.37**	**0.48**	–	0.35	0.32	**0.73**

Table displays inter-item and item-total Pearson's correlation coefficients. Items 3, 4, 5, 7, 8, 9 might not represent their hypothesized construct

belong. Standardized α for the total PAC-SYM score was 0.83.

Confirmatory/exploratory factor analysis

Competing models are illustrated in Appendix 1 and described in the "Methods" section.

Model 1 (*single-factor model, Appendix 1, panel 1*): Fit indexes (RMSEA = 0.17; normed $\chi^2 = 25.6$; GFI = 0.75; AIC = 1,277 and SBC = 1,022) indicated that this model did not fit the data appropriately.

Model 2 (*uncorrelated three-factor model, Appendix 1, panel 2*): Model 2 significantly improved model fit (RMSEA = 0.12; normed $\chi^2 = 12.3$; GFI = 0.87; AIC = 558 and SBC = 302; χ^2 test for comparison of nested models: $p < 0.01$) and all item loadings were >0.40. However, general model fit was still unsatisfactory.

Model 3 (*correlated three-factor model, Appendix 1, panel 3*): Model 3 allowed that inter-correlations among the three factors (i.e., ABD, REC, STO) would be freely estimated. This new specification slightly improved fit compared with Model 2 (RMSEA = 0.09; normed $\chi^2 = 7.7$; GFI = 0.92; AIC = 289 and SBC = 48; χ^2 test for comparison of nested models: $p < 0.01$). Parameter estimates indicated moderate inter-correlations among factors (range $r = 0.36$–0.57), which may indicate that a second-order factor exists.

Model 4 (*one second-order factor and three first-order factors, Appendix 1, panel 4*): Model 4 hypothesized that one second-order underlying factor (i.e., constipation severity) can account for the three first-order factors (i.e., ABD, REC, STO). This model slightly improved fit compared with Model 3 (RMSEA = 0.09; normed $\chi^2 = 6.8$; GFI = 0.93; AIC = 246 and SBC = 11; χ^2 test for comparison of nested models: $p < 0.01$). However, modification indexes for Model 4 showed that direct paths from the second-order factor to the PAC-SYM items would improve model fit, suggesting a structure consistent with a bi-factor model.

Model 5 (*bi-factor model*): we tested a bi-factor model with individual items loading on both a general factor (i.e., constipation severity) and the three first-order factors (i.e., ABD, REC, STO). Model 5a improved fit compared with Model 4 (RMSEA = 0.06; normed $\chi^2 = 3.6$; GFI = 0.97; AIC = 69 and SBC = −96; χ^2 test for comparison of nested models: $p < 0.01$). Modification indexes also suggested that error terms between items 8, 9, and 12 should be allowed to covary, which is clinically plausible. A new model (Model 5b) further improved model fit (RMSEA = 0.04; normed $\chi^2 = 2.4$; GFI = 0.98; AIC = 15 and SBC = −141; χ^2 test for comparison of nested models: $p < 0.01$).

Model 6 (*revised bi-factor model* Fig. 1): Even though the fit of Model 5b to the data appeared satisfactory, our results showed that divergent validity was questionable for items grouped in the REC subscale. Items 5 and 6 had high correlations with the STO subscale, while item 7 had a strong floor effect which may impact scale responsiveness to clinical change. Hence, we removed item 7 and ran an exploratory factor analysis which revealed two factors: the ABD (items 1–4), which was identical to the original one and the M-STO (items 5, 6, 8–12) which incorporated items 5 and 6. Such results suggests that the rectal scale reproducibility is questionable. We then tested a bi-factor model with two first-order factors (Model 6a). Even though this model fitted the data adequately, it did not provide a significant improvement compared with Model 5a (RMSEA = 0.08; normed $\chi^2 = 6.1$; GFI = 0.96;

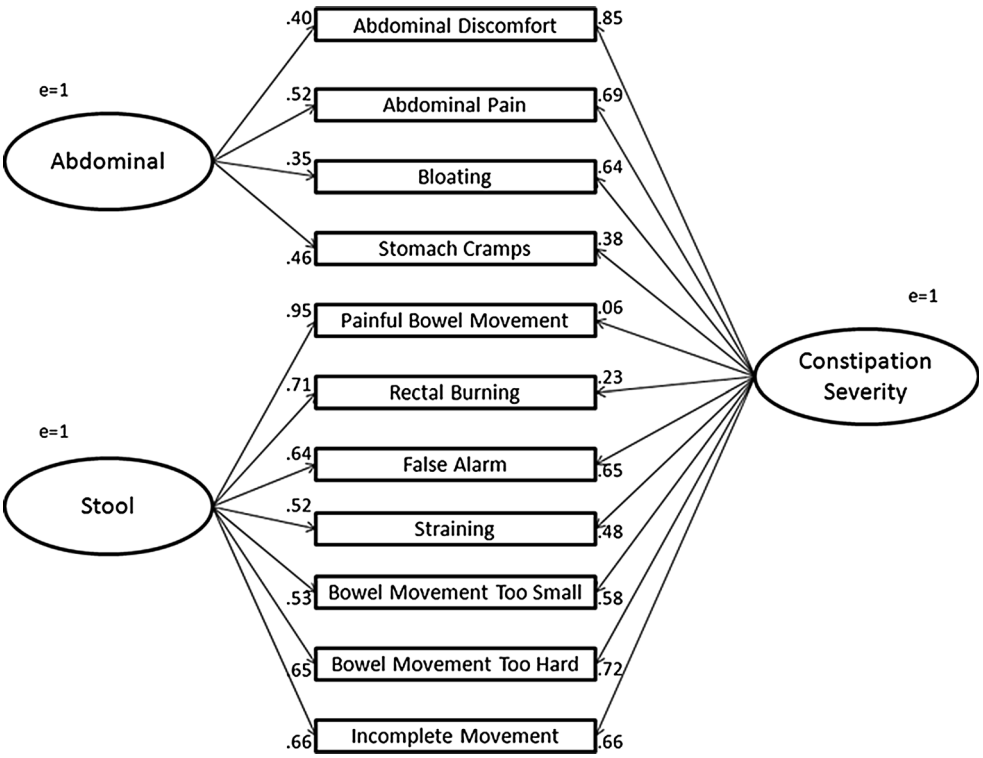

Fig. 1 Bi-factor model with two first-order factors. *Arrows* represents paths from latent factors and observed responses to questionnaire items. Coefficients represents standardized loadings of observed responses to questionnaire items on latent factors. *Note*: the term *e* represent an error term

AIC = 123 and SBC = −14). Modification indexes suggested that error terms of: Items 1 (abdominal discomfort) and 3 (bloating); 4 (stomach cramps) and 5 (painful bowel movement); and 9 (bowel movement too hard), 10 (bowel movement too small) and 11 (squeeze/strain to pass bowel movement) should be allowed to correlate, possibly indicating a co-occurrence of such symptoms. Since these covariations are clinically plausible, we tested a new model (Fig. 1) allowing the estimation of such error covariance parameters. The new specification (Model 6b) strongly improved model fit (RMSEA = 0.04; normed $\chi^2 = 2.4$; GFI = 0.98; AIC = 17 and SBC = −107; χ^2 test for comparison of nested models: $p < 0.01$).

Cross-validation. To assess whether the new bi-factor model with two first-layer factors is stable across samples of the same population, we tested the same Model 6b in the LIRS II sample obtaining similar fit statistics (RMSEA = 0.04; normed $\chi^2 = 3.6$; GFI = 0.99; AIC = 43 and SBC = −95). The internal consistency of the M-PAC-SYM score ($\alpha = 0.88$) and its subscales (ABD; $\alpha = 0.83$; M-STO; $\alpha = 0.87$) was excellent.

Discriminant validity

Discrimination of the modified PAC-SYM total (M-PAC-SYM) and its subscales scores (ABD and M-STO) across clinically relevant groups defined by Rome III criteria was tested in the LIRS II sample and presented in Table 4. The average M-PAC-SYM, ABD, and M-STO scores were 1.56 ± 0.82, 1.55 ± 0.92, and 1.57 ± 0.94, respectively. The ABD scale was more discriminative of constipated patients with abdominal pain, while the M-STO scale was more discriminative of constipated patients that satisfy the Rome III criteria for functional constipation, but who did not report abdominal pain. Among these criteria, reduced bowel frequency was associated with the smallest difference in scale scores (Table 4). Each additional Rome III symptom ascertained by a gastroenterologist corresponded to an increase in M-PAC-SYM ($\beta = 0.31$; $\omega^2 = 0.16$), ABD ($\beta = 0.19$; $\omega^2 = 0.05$), M-STO ($\beta = 0.37$; $\omega^2 = 0.18$) scores.

Concurrent validity

Correlations of SF-12 physical composite score with M-PAC-SYM total score and subscales ranged from −0.33 (M-STO) to −0.39 (M-PAC-SYM) while correlations of SF-12 mental composite score ranged from −0.32 (M-STO) to −0.37 (M-PAC-SYM). Correlations of M-PAC-SYM and satisfaction with treatment effectiveness, side effects, and convenience ranged from −0.28 ($r_{\text{TSQM:side–STO}}$), to −0.45 ($r_{\text{TSQM: Effectiveness–PAC-SYM: M}}$).

Responsiveness to clinical change

We tested responsiveness to change in 413 patients who initiated a new treatment during the 3 months preceding

Table 4 Discrimination of modified PAC-SYM scores across clinically different groups

Clinical characteristics[a]	Δ (σ)			
	M-PAC-SYM	M-ABD	M-STO	REC
Rome III				
Lumpy/hard stools	**0.32 (0.40)**	0.16 (0.17)	**0.39 (0.42)**	**0.16 (0.19)**
Incomplete evacuation	**0.44 (0.63)**	**0.29 (0.36)**	**0.52 (0.65)**	**0.25 (0.27)**
Obstruction	**0.57 (0.74)**	**0.29 (0.33)**	**0.68 (0.78)**	**0.53 (0.68)**
Manual maneuvers	**0.26 (0.33)**	0.04 (0.04)	**0.39 (0.43)**	**0.38 (0.45)**
<3 defecations/week	0.03 (0.04)	0.13 (0.15)	0.13 (0.14)	0.06 (0.07)
Strain	**0.29 (0.37)**	**0.22 (0.25)**	**0.33 (0.36)**	**0.16 (0.18)**
Abdominal pain	**0.26 (0.32)**	**0.44 (0.48)**	0.16 (0.17)	**0.19 (0.21)**
Pain—improves after bowel movement	**0.23 (0.29)**	**0.46 (0.50)**	0.10 (0.11)	**0.16 (0.18)**
Pain—changes after eating	**0.25 (0.31)**	**0.53 (0.60)**	0.09 (0.10)	**0.16 (0.18)**
Pain—is the most important symptom	**0.57 (0.74)**	**0.82 (0.98)**	**0.42 (0.47)**	**0.36 (0.41)**

Figures represent differences (Δ) and effect sizes (σ) in scores between patients with and without characteristics

Bold font indicates statistically significant differences at $p < 0.05$

[a] Data collected by a gastroenterologists during a regular outpatient visit

Table 5 Adjusted mean M-PAC-SYM score and effect size across classes of clinical change after initiating a new therapy

	Global rating of change					
	Strongly worsened	Worsened	Unchanged	Improved	Strongly improved	β (ω^2)
M-PAC-SYM	2.73	2.24	1.75	1.26	0.77	−0.49 (.25)
M-ABD	2.54	2.14	1.74	1.34	0.94	−0.40 (.15)
M-STO	2.83	2.29	1.75	1.21	0.67	−0.54 (.23)

Results represent estimated means, regression coefficients, and effect sizes for each M-PAC-SYM score (total and sub-scales) by classes of clinical change after therapy switch. Each general linear model has been adjusted by treatment regimen. Regression coefficient β represent the change in M-PAC-SYM score associated with an one class increase in global rating of change. Effect size estimates ω^2 represent the strength of the association between M-PAC-SYM scores and global rating of change classes

the interview. We observed a strong association between the M-PAC-SYM scales and patients' assessment of clinical change after initiating a new therapy in the 3 months prior to the interview independent of the treatment regimen (Table 5). Effect size estimates for difference across single classes of improvement ranged from $d = 0.43$ to 0.60 (moderate effect size), while effect sizes of regression estimates ranged from $\omega^2 = 0.15$ to 0.25 (large effect size).

Minimal clinically important difference

We assessed anchor-based MCID in 42 patients assessed longitudinally. Of them, 8 have been assessed three times. Overall, the longitudinal dataset included 92 valid observations. We observed an association between longitudinal variations in global rating of change scores and variations in M-PAC-SYM scores ($p < 0.01$ for all scales). Anchor-based MCID were 0.24, 0.38, and 0.16 for M-PAC-SYM, ABD, and M-STO scores, respectively, corresponding to $d = 0.30$, 0.39, 0.20 effect size estimates for differences across classes of improvement. This difference in the total score translates into a one-point difference in three (or bigger changes in a smaller number of symptoms) of the 11 items of the M-PAC-SYM general severity scale.

Discussion

PAC-SYM is a key outcome measure in patients with chronic constipation. [3–6]. The PAC-SYM is used to assess both general constipation severity and the severity of specific aspects of this condition, namely abdominal, rectal, and stool/defecation symptoms. These sub-scales imply that each cluster of symptoms independently provides a unique, non-overlapping, piece of information beyond a general severity factor. However, the original validation sample of the PAC-SYM was fairly small, and the proposed factor structure and scoring algorithm have never been replicated. Additionally, it should be pointed out that the previous

attempts to link pathophysiological abnormalities to symptom clusters lead to limited results [23, 24].

Our study confirmed that the PAC-SYM questionnaire is a reliable tool in third-level of care consulters with chronic non-organic constipation, and we have found evidence partially supporting construct validity through confirmatory factor analysis. However, the construct validity and suitability of the rectal subscale as an outcome measure in this population was questionable. The analysis of inter-item correlation matrix revealed that items belonging to the rectal domain might indeed be explained by different phenomena rather than representing rectal symptoms as a stand-alone concept. This fact was coupled with the high floor effect of item 7. Overall, these findings suggest that items grouped in the PAC-SYM rectal domain may not be relevant for the majority of patients or may represent different clinical problems, thus reducing scale responsiveness to change, discrimination of clinically different groups, and construct validity; these findings might explain the small effect sizes on this scale observed in recent RCTs [3].

The results of our study offer an evidence-based and easy solution to such problem. After deletion of item 7 (rectal bleeding or tearing during or after bowel movement), we were able to replicate the new bi-factor model with two first-order factors (i.e., the original abdominal and a new stool domains) in both LIRS waves, thus lending support to the stability of this structure in the population of interest. The proposed modified version of the PAC-SYM provided excellent fit to the data, discriminated across clinically different subgroups, showed strong concurrent validity with measures of well-being and treatment satisfaction and was sensitive to clinical changes.

In contrast with the Rome III definition for functional constipation, abdominal pain or discomfort was not adopted as a criterion of exclusion from our study, in line with the recent evidence showing that functional constipation and IBS-C are not distinct entities [10]. In this regard, the ABD subscale was more discriminative of patients reporting abdominal pain, whereas the new M-STO subscale was more discriminative of patients reporting symptoms related to defecation. These results suggest that within the continuum of patients with chronic constipation or with IBS-C, clinical differences may exist and be revealed by the modified PAC-SYM subscales but not by the original REC scale which was less discriminative than either the ABD and M-STO scales, respectively, for each symptom cluster (Table 4). In addition, the Rome III criterion of reduced bowel frequency was poorly associated with the severity of constipation, in line with previous studies among patients who self-reported constipation in whom bowel frequency was normal in about half of the cases [25, 26] and infrequent bowel movements was scored only as the fifth most bothersome symptom [25].

A further important aim of our study was to establish anchor-based minimal clinically important differences for PAC-SYM general factor and its subscales, a key property allowing the interpretation of scores in research and clinical practice. The Food and Drug Administration and EMEA have recommended that investigators classify responders for the primary outcome in RCTs based upon a priori rules reflecting clinically meaningful symptom improvements. We have shown that anchor-based MCID relative to a global rating of change scale corresponds to a weak–moderate effect size, consistent with previous recommendations in this field of research [16–22]. In previous RCTs, authors defined clinical response based on a one-point change in PAC-SYM score [3], which is much greater than the anchor-based MCID observed in our study (0.24 in repeated measure analysis), thus possibly underestimating the response rate [3–6]. Our data provide evidence-based minimal clinically important differences for interpreting differences in M-PAC-SYM total score observed in clinical trials and epidemiologic studies. The MCID found in our study translates into a one-point difference in 3 (or bigger changes in a smaller number of symptoms) of the 11 items of the M-PAC-SYM general severity scale.

Our study has several strengths. We were able to comprehensively assess the psychometric properties of the PAC-SYM questionnaire in a large sample of patients with chronic non-organic constipation and cross-validate our results in an independent sample of the same population lending support to the stability of our results.

However, we must also acknowledge some limitations. Since, we enrolled patients with a long history of constipation from third level of care outpatient clinics, our results might not generalize to all subjects with constipation. In particular, we cannot exclude that issues relative to the rectal scale such as rectal bleeding or tearing during or after bowel movement might be more pronounced in other clinical settings, such as primary health care or coloproctology units, where patients with acute symptoms are possibly more common. The invariance of the factor structure observed in our samples should be tested in patients with different clinical profiles and health care seeking behavior. Additionally, our MCID assessment relied on a small sample of patients who were assessed longitudinally. As a consequence, we cannot exclude that selection bias and classification bias have occurred.

In conclusion, our study confirmed that the PAC-SYM questionnaire is a reliable tool for the assessment of constipation severity and partially supports its construct validity. However, we showed that the validity of the rectal symptoms sub-scale is questionable in patients with chronic constipation, indicating that this scale should not be used to assess constipation severity in this population. A global constipation severity score and its subscales defining abdominal and stool-related symptoms derived from a

modified 11-item version might better represent symptom severity of most patients seeking care in gastroenterology referral centers and should be used instead of the original PAC-SYM version in this population.

Acknowledgments The LIRS study has been partially funded by Shire Pharmaceutical Inc. Data collection has been carried out by DoxaPharma SRL.

Conflict of interest Dr. Neri has received an unrestricted research grant by DoxaPharma SRL to support the production of this manuscript.

References

1. Spiegel, B., Camilleri, M., Bolus, R., Andresen, V., Chey, W. D., Fehnel, S., et al. (2009). Psychometric evaluation of patient-reported outcomes in irritable bowel syndrome randomized controlled trials: A Rome Foundation report. *Gastroenterology, 137*(6), 1944–1953.
2. Frank, L., Kleinman, L., Farup, C., Taylor, L., & Miner, P, Jr. (1999). Psychometric validation of a constipation symptom assessment questionnaire. *Scandinavian Journal of Gastroenterology, 348*, 70–77.
3. Tack, J., Stanghellini, V., Dubois, D., Joseph, A., Vandeplassche, L., & Kerstens, R. (2014). Effect of prucalopride on symptoms of chronic constipation. *Neurogastroenterology and Motility, 1*, 21–27.
4. Camilleri, M., Kerstens, R., Rykx, A., et al. (2008). A placebo-controlled trial of prucalopride for severe chronic constipation. *New England Journal of Medicine, 358*, 2344–2354.
5. Quigley, E. M., Vandeplassche, L., Kerstens, R., & Ausma, J. (2009). Clinical trial: The efficacy, impact on quality of life, and safety and tolerability of prucalopride in severe chronic constipation—A 12-week, randomized, double-blind, placebo-controlled study. *Alimentary Pharmacology and Therapeutics, 29*, 315–328.
6. Tack, J., van Outryve, M., Beyens, G., Kerstens, R., & Vandeplassche, L. (2009). Prucalopride (Resolor) in the treatment of severe chronic constipation in patients dissatisfied with laxatives. *Gut, 58*, 357–365.
7. Slappendel, R., Simpson, K., Dubois, D., & Keininger, D. L. (2006). Validation of the PAC-SYM questionnaire for opioid-induced constipation in patients with chronic low back pain. *European Journal of Pain, 10*(3), 209–217.
8. Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
9. Longstreth, G. F., Thompson, W. G., Chey, W. D., Houghton, L. A., Mearin, F., & Spiller, R. C. (2006). Functional bowel disorders. *Gastroenterology, 130*, 1480–1491.
10. Wong, R. K., Palsson, O. S., Turner, M. J., et al. (2010). Inability of the Rome III criteria to distinguish functional constipation from constipation-subtype irritable bowel syndrome. *American Journal of Gastroenterology, 105*, 2228–2234.
11. Apolone, G., Mosconi, P., & Quattrocchi, L. (2005). *Questionario sullo stato di salute SF-12, versione italiana*. Milan: Guerini e Associati Editore. ISBN 8883352017.
12. Atkinson, M. J., Kumar, R., Cappelleri, J. C., et al. (2005). Hierarchical construct validity of the treatment satisfaction questionnaire for medication (TSQM version II among outpatient pharmacy consumers. *Value Health, 8*(Suppl 1), S9–S24.
13. Kamper, S. J., Maher, C. G., & Mackay, G. (2009). Global rating of change scales: A review of strengths and weaknesses and considerations for design. *The Journal of Manual and Manipulative Therapy, 17*(3), 163–170.
14. Chen, F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research, 41*(2), 189–225.
15. Hu, L.-T., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
16. Crosby, R. D., Kolotkin, R. L., & Williams, G. R. (2003). Defining clinically meaningful change in health-related quality of life. *Journal of Clinical Epidemiology, 56*, 395–407.
17. Samsa, G., Edelman, D., Rothman, M. L., et al. (1999). Determining clinically important differences in health status measures: A general approach with illustration to the health utilities index mark II. *Pharmacoeconomics, 15*, 141–155.
18. Norman, G. R., Sloan, J. A., & Wyrwich, K. W. (2003). Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Medical Care, 41*, 582–592.
19. Beaton, D. E. (2003). Simple as possible? Or too simple? Possible limits to the universality of the one-half standard deviation (comment). *Medical Care, 41*, 593–596.
20. Farivar, S. S., Kiu, H., & Hays, R. D. (2004). Another look at the half standard deviation estimate of the minimally important difference in health-related quality of life scores. *Expert Review of Pharmacoeconomics and Outcomes Research, 4*(5), 521–529.
21. Cella, D., Eton, D. T., Lai, J., Peterman, A. H., & Merkel, D. E. (2002). Combining anchor and distribution-based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy (FACT) anemia and fatigue scales. *Journal of Pain and Symptom Management, 24*, 547–561.
22. Maringwa, J. T., Quinten, C., King, M., Ringash, J., Osoba, D., Coens, C., et al. (2011). Minimal important differences for interpreting health-related quality of life scores from the EORTC QLQ-C30 in lung cancer patients participating in randomized controlled trials. *Supportive Care in Cancer, 19*(11), 1753–1760.
23. Mertz, H., Naliboff, B., & Mayer, E. A. (1999). Symptoms and physiology in severe chronic constipation. *American Journal of Gastroenterology, 94*, 131–138.
24. Dinning, P. G., Jones, M., Hunt, L., Fuentealba, S. E., Kalanter, J., King, D. W., et al. (2011). Factor analysis identifies subgroups of constipation. *World Journal of Gastroenterology, 17*, 1468–1474.
25. Ashraf, W., Park, F., Lof, J., & Quigley, E. M. M. (1966). An examination of the reliability of reported stool frequency in the diagnosis of idiopathic constipation. *American Journal of Gastroenterology, 91*(1), 26–32.
26. Johanson, J. F., & Kralstein, J. (2007). Chronic constipation: A survey of the patient perspective. *Alimentary Pharmacology and Therapeutics, 25*, 599–608.