Translation and international validation of the Colostomy Impact score

Kristensen, Helle; Thyø, Anne; Jøssing Emmertsen, Katrine; Smart, Neil J.; Pinkney, Thomas; Warwick, Andrea M.; Pang, Dong; Furnée, Edgar J.B.; Verkuijl, Sanne J.; José Rama, Nuno

Published in:
Colorectal Disease

DOI:
10.1111/codi.15635

Received: 24 January 2021 | Revised: 3 March 2021 | Accepted: 10 March 2021

DOI: 10.1111/codi.15635

ORIGINAL ARTICLE

Translation and international validation of the Colostomy Impact score

Helle Ø Kristensen[1] | Anne Thyø[1] | Katrine Jøssing Emmertsen[1,2] | Neil J. Smart[3] | Thomas Pinkney[4] | Andrea M. Warwick[5] | Dong Pang[6] | Edgar J. B. Furnée[7] | Sanne J. Verkuijl[7] | Nuno José Rama[8] | Hugo Domingos[9] | João Maciel[10] | Alejandro Solis-Peña[11] | Eloy Espín Basany[11] | Marta Hidalgo-Pujol[12] | Sebastian Biondo[12] | Annika Sjövall[13,14] | Peter Christensen[1]

[1]Department of Surgery, Danish Cancer Society Centre for Research on Survivorship and Late Adverse Effects After Cancer in the Pelvic Organs, Aarhus University Hospital, Aarhus N, Denmark

[2]Surgical Department, Randers Regional Hospital, Aarhus N, Denmark

[3]Royal Devon and Exeter NHS Foundation Trust, Exeter, UK

[4]Academic Department of Surgery, University of Birmingham, Birmingham, UK

[5]Brisbane Academic Functional Colorectal Unit, QEII Hospital, Brisbane, QLD, Australia

[6]Peking University School of Nursing, Beijing, China

[7]Department of Surgery, Division of Abdominal Surgery, University Medical Center, University of Groningen, Groningen, The Netherlands

[8]Surgery – Colorectal Unit, Centro Hospitalar de Leiria, Leiria, Portugal

[9]Colorectal Surgery Unit, Champalimaud Foundation, Lisboa, Portugal

[10]Colorectal Surgery Unit, Instituto Português de Oncologia- Lisbon, Lisboa, Portugal

[11]Colorectal Surgery Unit, General Surgery Department, Hospital Vall d'Hebron, Universitat Autonoma de Barcelona, Barcelona, Spain

[12]Colorectal Surgery Unit, Bellvitge University Hospital, Barcelona, Spain

[13]Division of Coloproctology, Department of Pelvic Cancer, Karolinska University Hospital, Stockholm, Sweden

[14]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden

Correspondence
Helle Ø Kristensen, Department of Surgery, Danish Cancer Society Centre for Research on Survivorship and Late Adverse Effects After Cancer in the Pelvic Organs, Aarhus University Hospital, Palle Juul-Jensens Blvd 35, C-119-139, 8200 Aarhus N, Denmark.
Email: hellkren@rm.dk

Funding information
Working hours of HØK were funded by the Danish Cancer Society and the Department of Clinical Medicine at Aarhus University Hospital. The Bengt-Ihre Foundation funded work-time for data collection in Sweden.

Abstract

Aim: Optimal oncological resection in cancers of the lower rectum often requires a permanent colostomy. However, in some patients a colostomy may have a negative impact on health-related quality of life (HRQoL). The Colostomy Impact (CI) score is a simple questionnaire that identifies patients with stoma dysfunction that impairs HRQoL by dividing patients into 'minor' and 'major' CI groups. This aim of this study is to evaluate construct and discriminative validity, sensitivity, specificity and reliability of the CI score internationally, making it applicable for screening and identification of patients with stoma-related impaired HRQoL.

Method: The CI score was translated in agreement with WHO recommendations. Cross-sectional cohorts of rectal cancer survivors with a colostomy in Australia, China, Denmark, the Netherlands, Portugal, Spain and Sweden were asked to complete the CI score, the European Organization for Research and Treatment of Cancer (EORTC) quality of life 30-item core questionnaire, the stoma-specific items of the EORTC quality of life 29-item colorectal-specific questionnaire and five anchor questions assessing the impact of colostomy on HRQoL.

Results: A total of 2470 patients participated (response rate 51%–93%). CI scores were significantly higher in patients reporting reduced HRQoL due to their colostomy than in patients reporting no reduction. Differences in EORTC scale scores between patients with minor and major CI were significant and clinically relevant. Sensitivity was high regarding dissatisfaction with a colostomy. Regarding evaluation of discriminative validity, the CI score relevantly identified groups with differences in HRQoL. The CI score proved reliable, with equal CI scores between test and retest and an intraclass correlation coefficient in the moderate to excellent range.

Conclusion: The CI score is internationally valid and reliable. We encourage its use in clinical practice to identify patients with stoma dysfunction who require further attention.

KEYWORDS

Colostomy, health-related quality of life, rectal cancer, stoma

What does this paper add to the literature?

The Colostomy Impact (CI) score is a patient-reported outcome measure identifying patients with stoma dysfunction that impairs health-related quality of life. This study evaluates construct and discriminative validity, sensitivity, specificity and reliability of the CI score in cross-sectional cohorts of rectal cancer survivors with a colostomy in Australia, China, Denmark, the Netherlands, Portugal, Spain and Sweden.

INTRODUCTION

Surgery is still the cornerstone in curative treatment of cancer of the rectum. In tumours close to the anal verge, formation of a permanent end colostomy is often necessary to obtain an optimal oncological resection. However, a colostomy may have a negative impact on health-related quality of life (HRQoL) in some patients. The boundaries for sphincter-preserving surgery are continuously being pushed, with intersphincteric resections reducing the number of patients requiring an end colostomy. Nevertheless, it is well known that sphincter-preserving surgery carries a risk of low anterior resection syndrome and that the risk of reduced HRQoL from bowel dysfunction increases with decreasing tumour height from the anal verge [1,2]. There is no conclusive evidence that HRQoL in patients with a stoma is inferior to that of patients who have a low anterior resection [3–5]. With the improved 5-year survival rate, there are an increasing number of long-term survivors who have to live with the late effects of their cancer treatment. Increasingly HRQoL has gained acceptance as an important outcome in cancer treatment, and studies have shown that in 19%–23% of patients with a permanent colostomy after rectal cancer surgery the stoma impairs their HRQoL [6,7]. A number of questionnaires have been developed to study stoma-related quality of life [8–10]. However, in view of the need for a short, simple and valid screening tool to identify patients with reduced HRQoL due to stoma dysfunction, the Colostomy Impact (CI) score was developed in 2016 [6]. The unidimensional CI score enables clinicians to identify patients with stoma dysfunction in an efficient, standardized and systematic way in order to initiate targeted measures to improve HRQoL. Once validated, the score will also be valuable for research purposes in standardizing and simplifying the reporting of stoma dysfunction.

The CI score was developed in a Danish population in 2016 and identified patients with reduced HRQoL due to stoma dysfunction with a sensitivity of 85.7% and specificity of 59.5%. The CI score has recently been validated in Danish patients with a colostomy after surgery for benign conditions [11] and is now applicable to a Danish population with a colostomy regardless of the underlying condition. Since cultural and geographical differences may affect the impact of colostomy on patients, the CI score should be meticulously translated and validated thoroughly before translated versions are taken into use.

This validation study reports the translation of the CI score and aims to evaluate construct and discriminative validity as well as sensitivity and specificity of the CI score in Australia, China, Denmark, the Netherlands, Portugal, Spain and Sweden; in addition the test–retest reliability of the CI score was evaluated in Denmark, the Netherlands, Spain and Sweden.

METHOD

Participants

Cross-sectional cohorts of rectal cancer survivors with a permanent colostomy were identified in the participating countries, as shown in Table 1. Inclusion criteria were curative surgery for rectal cancer resulting in a permanent end colostomy and at least 12 months of follow-up. Exclusion criteria were recurrence, inability to complete

TABLE 1 Patient characteristics

	Australia	China	Denmark	The NL	Portugal	Spain	Sweden
n	95	110	1583	117	97	207	258
Population/inclusion method	Hospital database	Hospital's medical record system	National register (DCCG)	Dutch Surgical Colorectal Audit	Hospital database	Hospital database	National register (SCRCR)
Response rate	51%	93%	74%	82%	80%	67%	61%
Mode of administration							
Web-based, *n* (%)	1 (2)	24 (22)	942 (60)	43 (37)	0	0	0
Pen and paper, *n* (%)	53 (94)	83 (78)	641 (40)	73 (63)	29 (30)	46 (23)	258 (100)
Interview, *n* (%)	2 (4)	0	0	0	67 (69)	156 (77)	0
Sex							
Male, *n* (%)	67 (68)	73 (66)	985 (63)	71 (62)	63 (65)	136 (66)	153 (59)
Female, *n* (%)	31 (32)	37 (34)	584 (37)	44 (38)	34 (35)	70 (34)	105 (41)
BMI (kg/m^2), mean (range)	26.9 (16–41)	23.5 (16–32)	27.0 (14–68)	27.1 (16–55)	26.9 (16–35)	26.9 (15–66)	25.9 (15–61)
Age (years), mean (range)	69.7 (35–93)	67.0 (31–93)	74.4 (30–96)	70.7 (37–94)	71.9 (35–97)	76.1 (47–96)	72.7 (36–91)
Follow-up (years), mean (range)	4.5 (1.0–15.0)	4.6 (1.0–39.2)	6.2 (2.1–12.2)	5.7 (1.8–10.7)	5.6 (1.6–16.6)	6.1 (1.0–15.1)	4.1 (1.0–7.9)
Access to stoma nurse							
Yes, *n* (%)	86 (85)	120 (98)	1,180 (77)	101 (89)	77 (87)	160 (78)	190 (77)
No, *n* (%)	10 (10)	3 (2)	47 (3)	7 (6)	4 (4)	32 (16)	24 (10)
Don't know, *n* (%)	5 (5)	0	308 (20)	6 (5)	8 (9)	12 (6)	34 (13)
Stage (IUCC)							
0, *n* (%)	3 (3)	5 (5)	0	0	11 (11)	10 (5)	18 (7)
I, *n* (%)	35 (39)	30 (27)	523 (39)	18 (15)	30 (31)	42 (21)	83 (32)
II, *n* (%)	28 (31)	33 (30)	426 (31)	39 (34)	28 (29)	59 (30)	68 (26)
III, *n* (%)	22 (24)	28 (25)	399 (29)	59 (51)	24 (25)	79 (40)	71 (27)
IV, *n* (%)	2 (2)	14 (13)	5 (1)	0	4 (4)	8 (4)	20 (8)
Procedure							
APE, *n* (%)	70 (76)	105 (97)	1237 (78)	79 (68)	80 (82)	167 (83)	258 (100)
Hartmann's, *n* (%)	6 (7)	2 (2)	346 (22)	21 (18)	9 (9)	14 (7)	0
Pelvic exenteration, *n* (%)	15 (16)	0	0	16 (14)	1 (1)	5 (2)	0
Setting							
Acute, *n* (%)	1 (1)	0	13 (1)	3 (3)	6 (6)	6 (3)	0
Elective, *n* (%)	87 (99)	95 (86)	1570 (99)	113 (97)	90 (93)	195 (97)	258 (100)
ASA score							
1–2, *n* (%)	67 (77)	99 (80)	1356 (86)	108 (93)	57 (59)	131 (66)	179 (69)
≥3, *n* (%)	20 (23)	10 (8)	214 (14)	8 (7)	39 (40)	68 (34)	78 (31)

(Continues)

TABLE 1 (Continued)

	Australia	China	Denmark	The NL	Portugal	Spain	Sweden
Complications							
No complications, *n* (%)	25 (26)	101 (83)	1250 (79)	75 (65)	64 (66)	119 (59)	152 (58)
Clavien–Dindo I–II, *n* (%)	51 (53)	0	29 (2)	21 (18)	16 (17)	50 (25)	71 (27)
Clavien–Dindo III, *n* (%)	9 (9)	6 (5)	57 (4)	18 (15)	14 (14)	25 (12)	33 (13)
Clavien Dindo IV, *n* (%)	2 (2)	0	19 (1)	2 (2)	0	6 (3)	2 (1)
Unknown, *n* (%)	9 (9)	15 (12)	228 (14)	0	3 (3)	1 (1)	2 (1)
Oncological therapy							
Chemotherapy, *n* (%)	7 (8)	1 (1)	225 (35)	1 (1)	5 (5)	10 (5)	12 (5)
Radiotherapy, *n* (%)	1 (1)	1 (1)	39 (6)	30 (26)	7 (8)	14 (7)	119 (46)
Chemo- and radiotherapy, *n* (%)	53 (61)	99 (98)	375 (57)	70 (60)	47 (52)	140 (71)	76 (29)
None, *n* (%)	26 (30)	0	15 (2)	15 (13)	32 (35)	33 (17)	52 (20)

Abbreviations: APE, abdominoperineal excision; ASA, American Society of Anesthesiologists; BMI, body mass index; NL, Netherlands.

questionnaires (dementia, impaired vision, inability to read or write the specific language), age under 18 years and lacking consent. In Denmark, patients were identified from a national registry (Table 1). Eligible patients were sent an electronic invitation and a web-based questionnaire or a paper version of the questionnaire by regular mail with a prepaid envelope, depending on patient preference. Reminders were sent to nonresponders after approximately 3 weeks.

A similar approach was used in Sweden. In the Netherlands, patients treated at the participating hospitals were identified from a national database and approached as described above. Patients in Australia, Spain and Portugal and 91 of the patients from China were identified from local hospital databases or the hospitals' electronic medical record systems and approached in person, by mail or by telephone. In China, convenient inclusion from the hospital's stoma clinic was employed for inclusion of 29 patients.

For test–retest analysis, subgroups of patients in Denmark, Sweden, the Netherlands and Spain were randomly selected and asked to complete the CI score a second time 2–6 weeks after completion of the primary questionnaire, along with a question on changes in stoma function or stoma care since the primary questionnaire. Patients reporting changes in stoma function or returning the second test after more than 6 weeks were excluded.

Questionnaire/booklet

Patients were asked to complete questions regarding sociodemographic factors, lifestyle, stoma care and stoma complications, five anchor questions on the overall impact of colostomy on HRQoL, the CI score, the European Organization for Research and Treatment of Cancer quality of life 30-item core questionnaire (EORTC QLQ-C30) and the stoma-specific items of the EORTC quality of life 29-item colorectal-specific questionnaire (EORTC QLQ-CR29). Disease- and treatment-specific information was collected from hospital charts (Australia, China and the Netherlands) or registers/databases (Denmark Portugal, Spain and Sweden).

The primary anchor question was 'Overall, do you think that the colostomy impairs your quality of life?' The options were 'not at all', 'a little', 'some' and 'a lot'. This question is unvalidated but was part of the basic stoma questionnaire used for the development of the CI score. For the purpose of construct validation, the four categories of this anchor question were merged into two categories: colostomy impact 'none at all/a little' and colostomy impact 'some/a lot'. Four additional anchor questions regarding satisfaction, embarrassment, adaptation and restrictions in daily life were added to explore the construct of the CI score. Each question had four answer options that were dichotomized as described above to calculate the sensitivity and specificity of the CI score of each anchor.

The CI score is a seven-item patient reported outcome measure (PROM) concerning stoma-related problems, symptoms and complications. The CI score was developed using the primary anchor

question to identify factors associated with reduced HRQoL and to obtain a simple unidimensional measure of stoma dysfunction related to reduced HRQoL [6]. Each item has two to five answer options, and a weighted scoring system provides a total sum-score ranging from 0 to 38, with a higher score representing greater stoma dysfunction. A CI score of 0–9 points is categorized as minor CI and a score of 10–38 points is categorized as major CI (Appendix S1 and S2). As missing items hinder calculation of the sum-score, patients with missing items were excluded from further analysis.

The EORTC QLQ-C30 v.3.0 is a 30-item multidimensional generic HRQoL measure for cancer patients. It provides five multi-item functional scales and one multi-item global health status/QoL scale where a higher score represents better functioning, along with three multi-item symptom scales and six single-item measures where a higher score represents worse symptoms. Participants also completed the stoma-specific items of the EORTC QLQ-CR29 – a module for colorectal cancer patients consisting of 29 questions, seven of which concern patients with a stoma. Using linear transformation, scale scores of the EORTC QLQ-C30 were calculated ranging from 0 to 100 according to the scoring manual [12] and the stoma-specific items of the EORTC QLQ CR29 were summarized using linear transformation as a symptom scale ranging from 0 to 100. Missing data were handled according to the scoring manual. The clinical relevance of difference in EORTC QLQ-C30 scale scores between the minor and major CI groups were assessed according to Cocks et al. [13].

Translation and adaptation

The CI score has previously been translated into English [6]. The English version of the CI score, anchor questions and background questions were translated into Chinese, Dutch, Portuguese, Spanish and Swedish following WHO recommendations for translation and cross-cultural adaptation of instruments performed by professional translators. Final versions were reviewed by a colorectal surgeon at each participating centre and adapted where needed.

Validation

This validation study adheres to the Cosmin taxonomy of measurement properties [14]. Since items in the CI score are independent variables of the construct, stoma dysfunction, the conceptual framework is based on a formative model. Content validity was meticulously ensured in the development process, leaving the following measurement properties to be investigated: construct validity, discriminative validity and retest reliability.

Construct validity was assessed by testing hypotheses formulated a priori:

1. Patients reporting inferior HRQoL in the anchor question have a significantly higher CI score than patients reporting better HRQoL.
2. Patients with major CI have significantly lower scores than patients with minor CI on all functional EORTC QLQ-C30 scales and significantly higher scores on all symptom scales. Differences are clinically relevant in global health status, role functioning, social functioning and fatigue.
3. Patients with major CI have significantly higher scores than patients with minor CI on the stoma-specific EORTC QLQ-CR29 scale.
4. The CI score has a sensitivity of 85% and a specificity of 50% for identifying patients with 'some/a lot' CI as measured by the anchor questions.

Discriminative validity was assessed by an a priori formulated hypothesis: any differences in HRQoL between groups as measured by the primary anchor question will be reflected as a significant difference in CI score. Differences in CI score between the following groups were studied: age, sex, body mass index (<25 kg/m^2 or ≥25 kg/m^2), whether costs of stoma care products confer a financial burden on the household, surgical setting (acute or elective), neoadjuvant/adjuvant oncological treatment and postoperative complications (Clavien–Dindo <IIIb or ≥IIIb).

Retest reliability was analysed on an item level and on sum-score. For each item, a change between two adjacent answer categories resulted in a one-point change. A change of more than one category resulted in a two- or three-point change according to the number of categories. The item-level score was calculated as the sum of changed categories. The weighted sum-score was calculated according to the scoring instructions.

Statistical analysis

Analyses of retest reliability and construct validity were performed per country. Discriminative validity was analysed on the cumulated cohort.

The CI scores are presented as medians and interquartile ranges (IQRs) considering skewness of the data. The EORTC data are presented as mean (SD), as is the convention [15]. For significance testing, chi-square tests or Mann–Whitney *U*-tests were performed according to the type of data handled. An a priori power calculation was performed based on detecting a five-point difference in the scales of the EORTC QLQ-C30 with a significance level of 0.05 and power of 80%. Centres including patients corresponding to the power calculation or whose analysis reached significance were included in the analysis of EORTC data. All countries were included in the remaining analysis.

Receiver operating curves (ROCs) were plotted, yielding sensitivity and specificity.

Discriminative validity was assessed using the primary anchor question and the CI score was considered able to detect potential differences if tests of difference in CI score corresponded to the findings using the primary anchor question. Retest reliability was assessed by the Wilcoxon signed-rank test for significance between test and

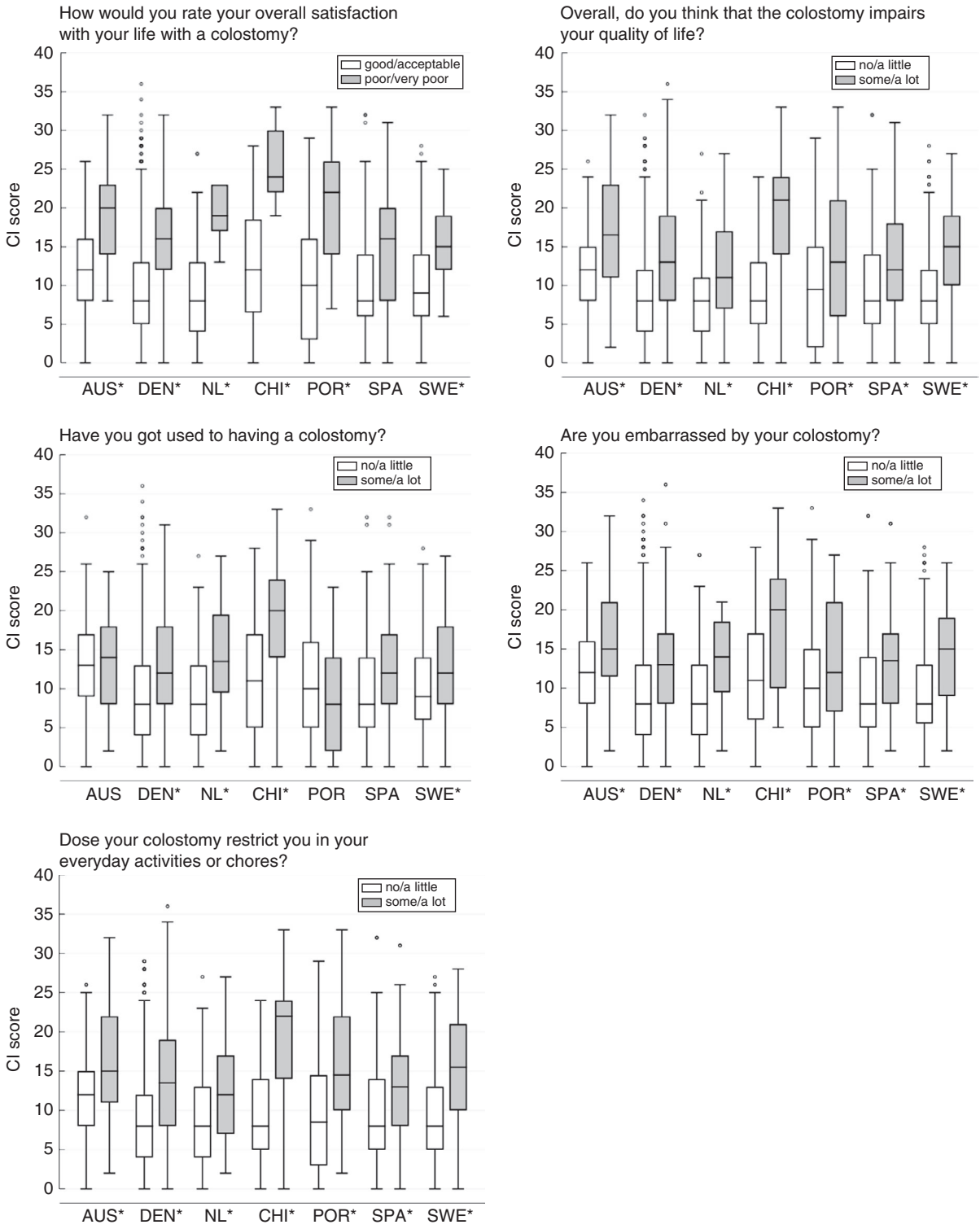

FIGURE 1 Construct validity; Median CI scores and IQR in HRQoL groups determined by anchor questions. Countries where the difference in CI score is significant between anchor groups are marked with (*)

retest, by intraclass correlation coefficients (ICCs) and Bland–Altman plots with limits of agreement (LoA). An ICC less than 0.5 was considered poor reliability, 0.5–0–75 moderate reliability, 0.75–0.9 good reliability and an ICC >0.9 was considered excellent [16].

Data were collected and managed using the REDCap electronic data collection tools hosted at Aarhus University [17,18]. All statistical analyses were performed using Stata 16 (StataCorp LCC). A significance level of 0.05 was chosen.

TABLE 2 Mean scale scores of the EORTC QLQ-C30. The size of the clinical relevance of the differences was assessed according to Cocks et al. [13] and reported as trivial (none), small (light grey ○), medium (grey ●) or large (dark grey ●)

EORTC QLQ-C30	China		Denmark		Spain		Sweden	
	Minor CI, mean (SD)	Major CI, mean (SD)	Minor CI, mean (SD)	Major CI, mean (SD)	Minor CI, mean (SD)	Major CI, mean (SD)	Minor CI, mean (SD)	Major CI, mean (SD)
Global QoL	78.1 (11.6)	62.1 (13.5)	80.2 (19.7)	66.6 (22.3)	72.0 (20.6)	62.0 (22.9)	78.2 (18.1)	63.4 (20.9)
Physical function (PF)	89.1 (6.6)	74.1 (18.0)	87.7 (17.2)	76.1 (22.8)	83.1 (23.2)	70.5 (29.4)	89.0 (15.2)	76.6 (20.8)
Role function (RF)	79.8 (14.3)	66.6 (20.4)	88.2 (20.7)	74.9 (28.4)	86.4 (27.8)	74.6 (33.8)	87.0 (21.8)	74.6 (25.8)
Emotional function (EF)	85.5 (10.9)	76.1 (14.3)	93.6 (12.2)	84.4 (19.5)	86.9 (18.9)	76.7 (24.6)	91.1 (13.6)	78.7 (22.7)
Cognitive function (CF)	94.5 (8.6)	80.0 (12.2)	91.6 (14.5)	83.1 (20.5)	86.0 (19.1)	82.1 (23.4)	92.2 (11.7)	81.4 (23.0)
Social function (SF)	72.8 (15.4)	68.7(19.8)*	91.7 (16.4)	80.9 (24.9)	88.6 (21.1)	77.1 (28.8)	86.8 (21.0)	72.6 (27.3)
Fatigue (FA)	6.2 (8.1)	26.8 (16.4)	16.7 (19.4)	30.8 (25.3)	17.5 (22.7)	29.8 (26.6)	17.3 (18.1)	33.1 (24.8)
Nausea and vomiting (NV)	.77 (5.0)	2.3 (7.7)*	1.8 (8.1)	5.6 (14.1)	2.9 (8.8)	3.8 (31.8)*	1.3 (5.4)	7.6 (16.7)
Pain (PA)	2.7 (7.2)	14.1 (15.9)	5.8 (15.6)	18.6 (25.2)	16.6 (25.5)	25.6 (25.7)	7.5 (17.4)	20.5 (25.1)
Dyspnoea (DY)	.77 (5.0)	12.3 (16.2)	9.8 (20.0)	19.3 (26.5)	12.6 (23.7)	18.6(26.6)*	16.8 (21.4)	30.1 (30.1)
Insomnia (SL)	8.5 (17.9)	23.5 (20.1)	12.6 (21.1)	24.4 (30.2)	22.6 (27.8)	34.0 (31.8)	13.8 (19.9)	26.8 (30.8)
Appetite loss (AP)	1.5 (7.1)	17.4 (18.7)	4.8 (14.4	10.6 (21.7)	8.4 (21.2)	13.6 (23.0)	4.0 (12.5)	10.2 (21.3)
Constipation (CO)	1.5 (7.1)	9.7 (16.3)	3.9 (12.5)	10.6 (21.7)	4.4 (13.8)	12.9 (23.9)	4.9 (14.6)	12.3 (23.4)
Diarrhoea (DI)	6.2 (13.1)	14.3 (22.0)	8.8 (17.0)	19.2 (25.5)	7.6 (16.1)	15.7()23.2	8.1 (16.7)	18.6 (24.9)
Financial difficulties (FI)	10.8 (17.3)	28.7 (21.9)	1.8 (9.3)	6.9 (19.1)	8.2 (21.0)	11.5(25.6)*	4.9 (16.4)	11.8 (25.0)

Abbreviations: CI, Colostomy Impact; EORTC QLQ-C30, European Organization for the Research and Treatment of Cancer (EORTC) quality of life 30-item core questionnaire; QoL, quality of life.

The sizes of the clinical relevance of differences in emotional function (EF) as estimates of clinically relevance were not available. *Difference between minor and major CI groups is not significant.

RESULTS

A total of 2470 patients answered the questionnaire. Patient characteristics by country are presented in Table 1. Response rates were between 51% and 93%.

For the evaluation of construct validity, the CI scores in patients grouped according to the anchor questions are shown in Figure 1. For the primary anchor question and the question regarding restrictions in everyday activities, the differences in CI score were significant in all countries. Differences between groups were significant for all five anchor questions in Denmark, the Netherlands, China and Sweden.

Differences in scale scores and the level of clinical relevance of differences in the EORTC QLQ-C30 between the minor and major CI groups are presented in Table 2 for countries having the number of participants stated in the power calculation. Differences were significant across all scales and measures in all countries except for 'nausea and vomiting', 'insomnia' and 'financial difficulties' in Spain and 'social function' and 'nausea and vomiting' in China. The differences in scale scores for the EORTC QLQ-C30 between the CI groups were all clinically relevant. Figure 2 shows mean scale scores for the EORTC QLQ-C30 by CI group. The stoma-specific scale of the EORTC QLQ-CR29 showed significantly more symptoms in the major than the minor CI group in all countries except for Australia. Differences ranged from 7.4 (Spain) to 14.2 (China).

ROC analyses were performed, yielding sensitivity and specificity of the CI score. Table 3 shows results for the primary anchor question and the question regarding satisfaction with life with a colostomy, the latter showed the best sensitivity in all countries except for Spain. Sensitivity ranged between 42% and 78% for the anchor question concerning adaptation, between 66% and 82% for the anchor question concerning embarrassment and between 60% and 89% for the anchor question concerning restrictions in daily activities (data not shown).

For discriminative validity, the primary anchor question was used to describe differences in HRQoL between subgroups. For groups showing a difference in the rate of patients reporting impaired HRQoL, the CI score correspondingly showed significantly higher CI scores in the groups with inferior HRQoL except for age groups where the difference in CI score did not reach significance (p = 0.0974). Similarly, for subgroups that reported 'some/a lot' impact on HRQoL at equal rates no difference in CI scores were present, with the exception of patients with Clavien–Dindo ≥II complications who had significantly higher CI scores but did not report an impact on HRQoL more often Table 4.

A total of 359 patients answered the retest. Of these, 121 were excluded as they reported change in stoma function since answering the primary questionnaire, leaving 238 eligible patients (75 from Denmark, 75 from Sweden, 45 from Spain and 43 from the Netherlands). No differences were found in item-level score or sum-score between the test and the retest. ICC scores showed moderate reliability in Sweden and the Netherlands for both sum-score (0.663 and 0.701, respectively) and item-level scores (0.640 and 0.749,

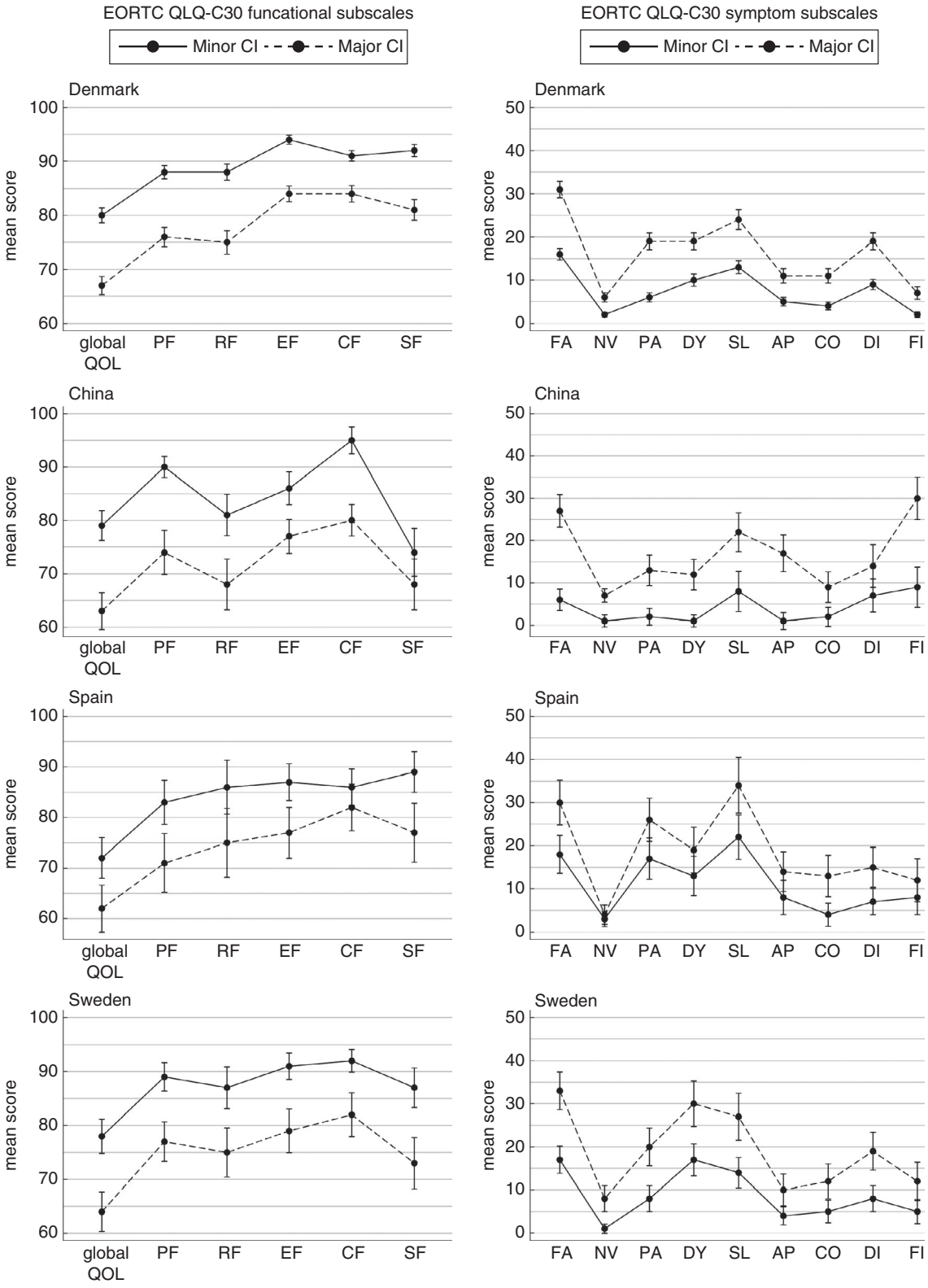

FIGURE 2 EORTC QLQ-C30 scale scores by country (mean +/- SD). Higher functional scores represent better functioning whereas higher symptom scores represent more symptoms. All differences are statistically significant except for NV and SF in China and NV, DY and FI in Spain. All significant differences are clinically relevant

respectively) and for sum-score in Denmark (0.705). Reliability was good on item level in Denmark (0.783) and excellent regarding both sum-score (0.919) and on item level (0.898) in Spain. Bland–Altman plots for assessing test–retest agreement are presented in Figure 3 for the sum-scores. Distributions were acceptable and the 95% LoA representing the individual difference between the test and the

TABLE 3 Sensitivity and specificity of the Colostomy Impact score across countries

	Sensitivity (%)	Specificity (%)	ROC AUC
Overall, do you think that the colostomy impairs your quality of life? Not at all/a little vs. some/a lot			
Australia	79.17	37.1	0.6542
China	85.1	59.0	0.8378
Denmark	68.50	61.50	0.7175
The Netherlands	57.89	60.27	0.6500
Portugal	60.00	50.00	0.6304
Spain	66.67	58.55	0.6600
Sweden	80.8	58.0	0.7403
How would you rate your overall satisfaction with your life with a colostomy? Good/acceptable vs. poor/very poor			
Australia	90.00	35.06	0.7617
China	100.00	43.00	0.9138
Denmark	80.70	55.68	0.7498
The Netherlands	100.00	56.60	0.9009
Portugal	80.00	46.51	0.7860
Spain	64.71	53.76	0.6216
Sweden	84.21	52.89	0.7480

Abbreviations: AUC, area under the curve; ROC, receiver operating characteristic.

retest ranged between −10.27 to 10.48 points (Sweden) and −4.89 to 6.04 points (Spain). LoA on item level was between −3.78 to 3.67 (Sweden) and −1.69 to 2.33 (Spain) (not shown).

DISCUSSION

The CI score is now available in Danish, English, Chinese, Dutch, Portuguese, Spanish and Swedish and this study reports high validity and reliability of the CI score in all included countries.

A thorough assessment of the validity and reliability of a PROM is crucial before a translated version is taken into use because cultural differences may affect the psychometric properties of the instrument. In all countries, the translated versions demonstrated convincing construct validity. CI scores were significantly higher in patients reporting that their colostomy reduced their overall QoL and in patients reporting that colostomy restricted their daily activities. Overall, our hypotheses regarding the anchor questions were satisfied in 30 of 35 tests, demonstrating that the CI score is a valid measure of stoma dysfunction affecting HRQoL. Likewise, differences in EORTC QLQ-C30 scales and measures, including the stoma-specific scale of the EORTC QLQ-CR29, between CI groups were found in 58 out of 64 scales (16 scales per country) and all

TABLE 4 Discriminative validity: the Colostomy Impact (CI) score was hypothesized to provide a significantly different CI score between groups if a difference was present measured by the anchor question and no difference in CI score if the groups reported the same rates of reduced health-related quality of life

	No/a little impact (%)	Some/a lot impact (%)	*P*-value	Median CI (IQR)	*P*-value
Gender					
Male	73.9	26.1		9 (6–15)	
Female	75.4	24.6	0.439	10 (5–15)	0.1298*
Age					
<65 years	70.2	29.8		10 (5–16)	
>65 years	75.4	24.5	0.0196	9 (5–14)	0.0974
BMI					
<25 kg/m^2	77.0	23.0		8 (5–13)	
>25 kg/m^2	73.3	26.7	0.0396	10 (6–15)	0.0000*
Household financial burden					
Unburdened	39.6	60.4		9 (5–14)	
Burdened	23.5	76.5	0.0000	12 (8–19)	0.0000*
Setting					
Acute	65	35		8.5 (3–15)	
Elective	75	25	0.2753	9 (5–14)	0.8942*
Complications					
<Clavien–Dindo III	75	25		9 (5–14)	
≥Clavien–Dindo III	69	31	0.1208	11 (7–17)	0.0002
Oncological treatment					
Any	72.3	26.2		10 (5–15)	
None	76.3	23.7	0.5314	10 (5–15)	0.7903*

Abbreviations: BMI, body mass index; IQR, interquartile range.
Agreement is indicated by an asterisk (*).

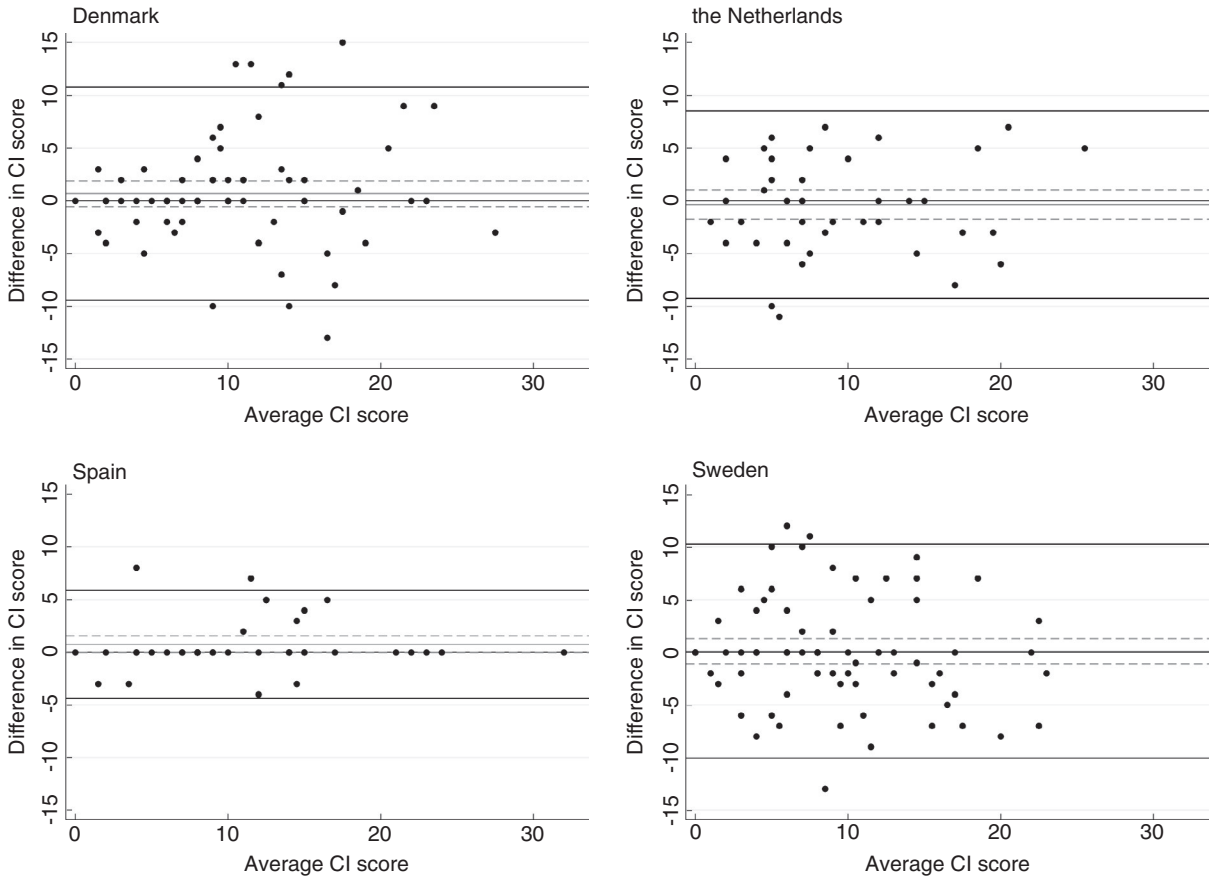

FIGURE 3 Bland Altman plots with difference in sum-score between the test and retest (light grey lines) with 95% CI (dashed light grey lines) and limits of agreement (dark grey lines) for each country

significant differences were clinically relevant, further supporting the construct validity of the CI score.

The EORTC QLQ-C30 is commonly used, and previous reports on colorectal cancer patients with a permanent colostomy correspond well to our results; mean scale scores found in three other cross-sectional studies are similar to the mean scores of the minor and major CI groups found in our study [19–21]. For the three included countries for which reference values of the background population were available, scale scores in the minor CI groups resemble scores of the reference population [22–24], indicating that the CI score categorized patients meaningfully into a minor group resembling the background population and a major group with impaired HRQoL.

The ROC analysis supported the hypothesis regarding sensitivity in one country and specificity in three countries for the primary anchor question. However, when looking at the additional anchor questions, sensitivity of the CI score detected dissatisfaction with life with a stoma with a sensitivity of 80% or more in all countries except Spain. Specificity was lower, which was also the case in the development of the CI score where a high sensitivity was intentionally prioritized. Thus, specificities in the ranges seen here were not surprising and patients with a sum-score in the major CI range (≥10) should have the possibility to decline further evaluation if they are satisfied with their colostomy and QoL. The anchor questions are not validated, but the primary anchor has repeatedly been used in a similar form for the development and validation of PROMs [6,25–27]. The broad phrasing may prompt different interpretations, especially when used in an international setting, which may affect sensitivity assessment. The addition of the other four anchor questions allowed us to further explore the construct measured by the CI score, and the score demonstrated highest sensitivity to dissatisfaction with life with a colostomy and to restrictions in daily activities.

The CI score was able to discriminate between groups that differed with regard to the proportion of patients reporting 'some/a lot' CI in the anchor question as five of our seven hypotheses were supported. Differences in median CI scores were modest (1–3 points) and whether this difference is clinically relevant remains unknown. However, differences in categorization according to the primary anchor question were similarly small.

This study demonstrates that the CI score is a reliable measure over time with no difference in scores between the test and the retest and ICCs in the moderate to excellent range.

We have demonstrated that the CI score is valid and reliable, and we highly encourage its use in the clinical setting where purposeful use of resources is crucial. Follow-up programmes after completed cancer treatment often focus on recurrence. However, concerns about late effects from the cancer treatment should be addressed [28–30]. Systematic use of the CI score will enable clinicians to screen colostomy patients in a quick and reliable way to identify

those suffering from stoma dysfunction so that targeted interventions can be initiated to improve patients' HRQoL.

A strength of this study is the meticulous professional translation of the CI score with adaptation from healthcare professionals ensuring semantically equivalent versions in all languages. Another strength is the use of validated EORTC questionnaires. The cross-sectional design and high number of patients included from each country reduces the risk of selection bias and increases generalizability. The large number of significance tests, however, increases the risk of type I error, which is a limitation to this study and was not taken into account when deciding on the significance level. Furthermore, the cross-sectional design hinders conclusions regarding causality when looking at differences in HRQoL between subgroups and evaluation of the responsiveness and minimal important change in the CI score. Numerous factors may affect HRQoL differently in different cultural settings, which may explain differences in scores between countries seen in this study. Financial concerns, cultural and religious norms and access to healthcare services vary between countries, and the impact of such differences on HRQOL should be further investigated. The CI score has proven to be valid and reliable in seven countries. However, validity cannot necessarily be generalized, and further validation studies are needed if the score is to be translated and used in new countries; future studies evaluating responsiveness and minimal important change of the CI score are encouraged.

ACKNOWLEDGEMENTS

We thank Nuno Vilela, Joaquim Costa Pereira, Edgar Amorim, Diana Parente, Rita Malaquias and António Oliveira for assistance in data acquisition for the Portuguese cohort. We also thank Meritxell Pera Ferreruela for assistance with the database at Vall D'Hebron hospital in Spain.

CONFLICT OF INTERESTS

None to disclose.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICAL STATEMENT

This is the authors' own original work, and has not been previously published elsewhere.

AUTHOR CONTRIBUTION

Conceptualisation: Peter Christensen, Katrine Jøssing Emmertsen, Anne Thyø, Helle Ø Kristensen. *Methodology*: Peter Christensen, Katrine Jøssing Emmertsen, Anne Thyø, Helle Ø Kristensen, Thomas Pinkney, Neil Smart. *Data curation*: Helle Ø Kristensen, Andrea M. Warwick, Dong Pang, Edgar J. B. Furnée, Sanne J. Verkuijl, Nuno José Rama, Hugo Domingos, João Maciel, Alejando Solis-Peña, Eloy Espín Basany, Marta Hidalgo-Pujol, Sebastian Biondo, Annika Sjövall; statistical analyses, Helle Ø Kristensen. *Writing-original draft preparation*: Helle Ø Kristensen. *Writing–review and editing*: Helle Ø Kristensen, Peter Christensen, Katrine Jøssing Emmertsen, Anne Thyø, Thomas Pinkney, Neil Smart, Andrea M. Warwick, Dong Pang, Edgar J. B. Furnée, Sanne J. Verkuijl, Nuno José Rama, Hugo Domingos, João Maciel, Alejando Solis-Peña, Eloy Espín Basany, Marta Hidalgo-Pujol, Sebastian Biondo, Annika Sjövall. *Supervision*: Peter Christensen, Katrine Jøssing Emmertsen, Anne Thyø. All authors have read and agreed to the published version of the manuscript.

ORCID

Helle Ø Kristensen https://orcid.org/0000-0003-0153-7689
João Maciel https://orcid.org/0000-0003-2496-1642
Sebastian Biondo https://orcid.org/0000-0002-7374-0371

REFERENCES

1. Battersby NJ, Juul T, Christensen P, Janjua AZ, Branagan G, Emmertsen KJ, et al. Predicting the risk of bowel-related quality-of-life impairment after restorative resection for rectal cancer: a Multicenter Cross-Sectional Study. Dis Colon Rectum. 2016;59(4):270–80.
2. Bretagnol F, Rullier E, Laurent C, Zerbib F, Gontier R, Saric J. Comparison of functional results and quality of life between intersphincteric resection and conventional coloanal anastomosis for low rectal cancer. Dis Colon Rectum. 2004;47(6):832–8.
3. Pachler J, Wille-Jørgensen P. Quality of life after rectal resection for cancer, with or without permanent colostomy. Cochrane Database of Syst Rev. 2012;12:CD004323.
4. Kasparek MS, Hassan I, Cima RR, Larson DR, Gullerud RE, Wolff BG. Quality of life after coloanal anastomosis and abdominoperineal resection for distal rectal cancers: sphincter preservation vs quality of life. Colorectal Dis. 2011;13(8):872–7.
5. de Campos-Lobato LF, Alves-Ferreira PC, Lavery IC, Kiran RP. Abdominoperineal resection does not decrease quality of life in patients with low rectal cancer. Clin Sci. 2011;66(6):1035–40.
6. Thyø A, Emmertsen KJ, Pinkney T, Christensen P, Laurberg S. The colostomy impact score: development and validation of a patient reported outcome measure for rectal cancer patients with a permanent colostomy. A population-based study. Colorectal Dis. 2017;19(1):25–33.
7. Kasparek MS, Hassan I, Cima RR, Larson DR, Gullerud RE, Wolff BG. Long-term quality of life and sexual and urinary function after abdominoperineal resection for distal rectal cancer. Dis Colon Rectum. 2012;55(2):147–54.
8. Prieto L, Thorsen H, Juul K. Development and validation of a quality of life questionnaire for patients with colostomy or ileostomy. Health Qual Life Outcomes. 2005;3(c):1–12.
9. Grant M, Ferrell B, Dean G, Uman G, Chu D, Krouse R. Revision and psychometric testing of the City of Hope Quality of Life-Ostomy Questionnaire. Qual Life Res. 2004;13(8):1445–57.
10. Nafees B, Rasmussen M, Lloyd A. The Ostomy-Q: development and psychometric validation of an instrument to evaluate outcomes associated with ostomy appliances. Ostomy Wound Manag. 2017;63(1):12–22.
11. Kristensen HØ, Krogsgaard M, Christensen P, Thomsen T. Validation of the Colostomy Impact score in patients ostomized for a benign condition. Colorectal Dis. 2020;22(12):2270–7.
12. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, et al. EORTC QLQ-C30 Scoring Manual. The EORTC QLQ-C30. Introduction. EORTC QLQ-C30 Scoring Man. 2001;30:1–67.

13. Cocks K, King MT, Velikova G, St-James MM, Fayers PM, Brown JM. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer quality of life questionnaire core 30. J Clin Oncol. 2011;29(1):89–96.
14. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737–45.
15. Aaronson NK, Aaronson NK, Ahmedzai S, Ahmedzai S, Bergman B, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365–76.
16. Schober P, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg. 2018;126(5):1763–8.
17. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform. 2019;95:103208.
18. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. NIH public access author manuscript. J Biomed Inform. 2009;42(2):377–81.
19. Engel J, Kerr J, Schlesinger-Raab A, Eckel R, Sauer H, Hölzel D. Quality of life in rectal cancer patients. Ann Surg. 2003;238:203–13.
20. Camilleri-Brennan J, Steele RJC. Objective assessment of morbidity and quality of life after surgery for low rectal cancer. Colorectal Dis. 2002;4(1):61–6.
21. Mols F, Lemmens V, Bosscha K, van den Broek W, Thong MSY. Living with the physical and mental consequences of an ostomy: a study among 1–10-year rectal cancer survivors from the population-based PROFILES registry. Psychooncology. 2014;23(9):998–1004.
22. Juul T, Petersen MA, Holzner B, Laurberg S, Christensen P, Grønvold M. Danish population-based reference data for the EORTC QLQ-C30: associations with gender, age and morbidity. Qual Life Res. 2014;23(8):2183–93.
23. Derogar M, Van Der Schaaf M, Lagergren P. Reference values for the EORTC QLQ-C30 quality of life questionnaire in a random sample of the Swedish population. Acta Oncol. 2012;51(1):10–6.
24. Nolte S, Liegl G, Petersen MA, Aaronson NK, Costantini A, Fayers PM, et al. General population normative data for the EORTC QLQ-C30 health-related quality of life questionnaire based on 15,386 persons across 13 European countries, Canada and the United States. Eur J Cancer. 2019;107:153–63.
25. Emmertsen KJ, Laurberg S. Low anterior resection syndrome score: development and validation of a symptom-based scoring system for bowel dysfunction after low anterior resection for rectal cancer. Ann Surg. 2012;255(5):922–8.
26. Emmertsen KJLSTA. The rectal cancer female sexuality score: development and validation of a scoring system for female sexual function after rectal cancer surgery. Dis Colon Rectum. 2018;61(6):656–66.
27. Mortensen AR, Thyø A, Emmertsen KJ, Laurberg S. Chronic pain after rectal cancer surgery – development and validation of a scoring system. Colorectal Dis. 2019;21(1):90–9.
28. El-Shami K, Oeffinger KC, Erb NL, Anne W, Bretsch J, Pratt-Chapman ML, et al. American Cancer Society colorectal cancer survivorship care guidelines. CA Cancer J Clin. 2015;65(6):428–55.
29. Greenfield DM, Absolom K, Eiser C, Walters SJ, Michel G, Hancock BW, et al. Follow-up care for cancer survivors: the views of clinicians. Br J Cancer. 2009;101(4):568–74.
30. Jacobs LA, Shulman LN. Follow-up care of cancer survivors: challenges and solutions. Lancet Oncol. 2017;18(1):e19–29.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Kristensen HØ, Thyø A, Jøssing Emmertsen K, et al. Translation and international validation of the Colostomy Impact score. *Colorectal Dis*. 2021;23:1866–1877. https://doi.org/10.1111/codi.15635