HHS Public Access

Author manuscript

Cancer. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

Cancer. 2015 September 1; 121(17): 3027–3035. doi:10.1002/cncr.29437.

Minimal Clinically Important Differences in the Edmonton Symptom Assessment Scale in Cancer Patients: A Prospective Multicenter Study

David Hui, MD, MSc[1], **Omar Shamieh, MD, MBA**[2], **Carlos Eduardo Paiva, MD, PhD**[3], **Pedro Emilio Perez-Cruz, MD**[4], **Jung Hye Kwon, MD, PhD**[5], **Mary Ann Muckaden, MD**[6], **Minjeong Park, MS**[7], **Sriram Yennu, MD**[1], **Jung Hun Kang, MD, PhD**[8], and **Eduardo Bruera, MD**[1]

[1] Department of Palliative Care and Rehabilitation Medicine, MD Anderson Cancer Center, Houston, USA

[2] Department of Palliative Care, King Hussein Cancer Center, Amman, Jordan

[3] Department of Medical Oncology, Barretos Cancer Hospital, Barretos, Brazil

[4] Department of Palliative Care, Pontificia Universidad Catolica de Chile, Santiago, Chile

[5] Department of Medical Oncology, Kangdong Sacred Heart hospital, Seoul, Republic of Korea

[6] Department of Palliative Care, Tata Memorial Center, Mumbai, India

[7] Department of Biostatistics, MD Anderson Cancer Center, Houston, USA

[8] Division of Hematology/Oncology Department of Internal Medicine, Gyeongsang University College of Medicine, Gyeongsang, Republic of Korea

Abstract

Background—The Edmonton Symptom Assessment Scale (ESAS) is widely used for symptom assessment in the clinical and research settings. We used the sensitivity-specificity approach to identify the minimal clinically important difference (MCID) for improvement and deterioration for each of the 10 ESAS symptoms.

Methods—This multicenter, prospective, longitudinal study enrolled advanced cancer patients. ESAS was measured at first clinic visit and a second visit 3 weeks later. For each symptom, we assessed Patient's Global Impression (“better”, “about the same”, or “worse”) at the second visit as the external criterion, and determined the MCID based on the optimal cutoff in receiver-operating characteristic (ROC) curve. We conducted sensitivity analysis by estimating MCIDs using other approaches.

Results—Among the 796 participants, the median duration between the 2 study visits was 21 days (interquartile range 18-28 days). The area under the ROC curve varied between 0.70-0.87, suggesting good responsiveness. For all 10 symptoms, the optimal cutoff was ≥1 point for

Corresponding author: David Hui, MD, MSc, Department of Palliative Care & Rehabilitation Medicine Unit 1414, University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX, USA 77030, Phone: 713-606-3376, Fax: 713-792-6092, dhui@mdanderson.org.

Disclosure: The authors have declared no conflicts of interest.

improvement and ≤−1 point for deterioration, with sensitivities of 59%-85% and specificities of 69%-85%. Using other approaches, the MCIDs varied between 0.8 and 2.2 for improvement and between −0.8 and −2.3 for deterioration in within-patient analysis, between 1.2 and 1.6 with the ½ standard deviation approach, and between 1.3 and 1.7 with the standard error of measurement approach.

Conclusions—ESAS was responsive to change. The optimal cutoffs were ≥1 point for improvement and ≤−1 point for deterioration for each of the 10 symptoms. Our findings have implications for sample size calculations and response determination.

Keywords

Neoplasms; outcome measures; pain; sample size; sensitivity and specificity; symptom assessment

INTRODUCTION

Patients with advanced cancer frequently experience significant symptom burden throughout the disease trajectory [1]. Routine standardized symptom assessment employing patient-reported outcomes represents the cornerstone of personalized symptom management [2]. A number of symptom batteries have been developed, with the Edmonton Symptom Assessment Scale (ESAS) being one of the most widely used questionnaires in clinical practice and research [3]. ESAS has been translated and adopted for symptom screening in many countries in North America, South America, Europe, Asia and Africa, and has been validated in different oncology and palliative care settings [4-11].

One critical aspect related to the symptom assessment using ESAS involves identifying what constitutes a minimal clinically important difference (MCID) [12]. In the acute pain setting, the MCID for a single 0-10 numeric rating scale has been studied [13]; however, the MCID of the 10 ESAS symptoms has not been systematically assessed in a prospective fashion [14]. A better understanding of the MCID of ESAS has important implications for symptom response determination and sample size calculation. In this multicenter prospective study, we determined the MCID for each of the 10 ESAS symptoms in patients with advanced cancer using the sensitivity-specificity anchor-based approach.

METHODS

Participants

This is an international longitudinal observational study. The inclusion criteria included the following: (1) diagnosis of advanced cancer, defined as locally advanced, recurrent or metastatic disease, (2) 18 years of age or greater, (3) seen at an outpatient clinic at one of the 6 participating centers, and (4) scheduled to return to clinic 14 to 34 days after the first study visit for a second ESAS questionnaire. Patients with delirium (Memorial Delirium Assessment Scale [MDAS] of 13 or greater) were excluded. The institution review boards at all participating centers approved the study. All participants provided written informed consent.

Participating centers included MD Anderson Cancer Center in Houston; United States, King Hussein Cancer Center in Amman, Jordan; Barretos Cancer Hospital in Barretos, Brazil; Pontificia Universidad Catolica de Chile in Santigo, Chile, Kangdong Sacred Heart hospital in Seoul, Republic of Korea, and Tata Memorial Center in Mumbai, India. All 6 institutions were tertiary care hospitals with access to comprehensive cancer treatments and supportive care. The centers in Houston, Jordan, Brazil and India are part of the Sister Institution Research Network, a multi-national cancer research cooperative. All participants were enrolled from the palliative care outpatient clinics at consultation with the following exceptions: a minority of Brazilian patients were enrolled at an outpatient palliative care follow up visit, US patients were consented during their first followup clinic visit because all assessments for the first study visit were routinely collected at consultation, and Korean patients were enrolled from oncology clinics. These minor variations in inception cohort provided us with a more diverse patient population to determine MCID and increased its generalizability.

Data collection

Data collection occurred between December 8, 2011 and April 30, 2014. We collected baseline patient characteristics, including age, sex, race, education level, cancer diagnosis, CAGE questionnaire [15] and MDAS [16] during the first study visit. For the purpose of this study, we considered all Brazilians and Chileans to be of Hispanic ethnicity. We also assessed ESAS and Karnofsky performance status during both the first and second visits, and Patient's Global Impression Scale (PGI) at the second visit. The site principal investigators all visited Houston to learn about the study procedures. To ensure data is collected in an accurate fashion, the study PI had regular teleconference with the research team at each site 1-2 times per month to provide training and longitudinal monitoring.

ESAS assesses the average intensity of 10 symptoms (pain, fatigue, nausea, depression, anxiety, drowsiness, shortness of breath, appetite, feelings of well-being and sleep) over the past 24 hours, each with an 11-point numerical rating scale that ranges from 0 (no symptom) to 10 (worst intensity) [3]. It has been translated into the languages in respective countries and by MAPI Research trust (i.e. English, Arabic, Portugese, Spanish, Korean and Hindi) and validated both linguistically and psychometrically [5, 8, 11, 17, 18].

PGI is a validated global rating of change scale used to evaluate subjective patients' response at the second visit [19, 20]. Patients were asked to answer the question for each of the 10 ESAS symptoms: "How is your symptom over the last 24 hours compared to your last visit?" for each of the 10 ESAS symptoms ("better", "about the same", "worse"). If the patient answered "better", they were asked "how much better?" ("much better", "better", "a little better"). Alternatively, if the patient answered "worse", they were asked "how much worse?" ("much worse", "worse", "a little worse"). PGI has been commonly used as a secondary outcome in a large number of pain studies and also used in several studies as an anchor for establishing clinical importance levels [21, 22].

Statistical analysis

Our sample size calculation revealed that it would require 777 patients to test the dual null hypotheses H0: true positive fraction <0.70 or false positive fraction >0.25 and H1: true positive fraction ≥0.84 and false positive fraction ≤0.10 with 80% power at 0.5% significance (5% divided by 10 symptoms to account for multiple testing), assuming at least 20% of patients will be better according to PGI. The final sample was 796 because this study required patients to complete both study visits, and that by the time we reached our target sample some remaining patients from each study site had already completed the first study visit and thus followed to completion.

We summarized the patient characteristics with descriptive statistics, including means, standard deviations, medians, interquartile ranges, and 95% confidence intervals. We compared the changes in ESAS scores between the first and second visits using paired t-tests.

Figure 1 outlines some commonly used approaches to identify MCID. For our primary analysis, we used the PGI as an external criterion against which ESAS changes were anchored and calibrated. We determined the MCID using sensitivity-specificity approach for both improvement (PGI “better” vs. PGI “about the same” and “worse”) and deterioration (PGI “worse” vs. PGI “better” and “about the same”). We plotted the receiver-operating characteristic (ROC) curves with true positive rate (sensitivity) on the y-axis and false positive rate (1 – specificity) on the x-axis. We then calculated the area under the curve (AUC), and determined the optimal cutoff for improvement and deterioration for each symptom based on the Youden J's index. The top left approach was also applied as a confirmatory measure [23].

We also conducted sensitivity analyses to estimate MCID using other commonly described anchor-based and distribution based approaches [24]. Specifically, we determined the within-patient changes by computing the average ESAS change for the PGI categories “a little better” and “a little worse” because these categories represented the smallest perceived change. 0.3 and 0.5 SD are often considered to be close approximates of MCID.[25, 26] We also examined the standard error of measurement (SEM), which represents the variation in the scores due to the unreliability of the scale using the following formula, $SEM = SD \times (1\text{-reliability})^{1/2}$ [27].

The Statistical Analysis Software 9.3 (SAS Institute Inc., Cary, NC, USA) software was used for statistical analysis. Statistically significance was declared when the P-value is <0.05.

RESULTS

Patient characteristics

The baseline demographics are shown in Table 1. The average age was 57 (range 19-85), 380 (48%) were female, and 229 (29%) were Caucasian, 190 (24%) had gastrointestinal cancers, and a 692 (87%) had metastatic disease. The median duration between the two study visits was 21 days (interquartile range 18-28 days).

ESAS Intensity

Table 2 shows the ESAS intensity at first and second clinic visit. The symptoms of highest average intensity were fatigue (4.9/10), pain (4.5/10) and poor well being (4.4/10). Pain, fatigue, depression, anxiety, poor well being, dyspnea and poor sleep all had statistically significant improvement in symptom intensity.

Patients reported if their perceived symptom change at the second visit relative to the first visit (Table 2). In PGI, 377 (47%), 293 (38%), and 293 (38%) felt that their pain, fatigue and poor well being improved, respectively. Dyspnea had the smallest proportion of patients reporting an improvement (19%).

We plotted the average change in ESAS intensity by PGI categories which demonstrates a gradient effect (Figure 2).

Determination of MCID

Figure 3 shows the ROC curves for each of the 10 ESAS symptom. Table 3 illustrates the ROC curve AUC and optimal cutoffs. The AUC varied between 0.70-0.87, suggesting good discrimination for ESAS. For all 10 symptoms, the optimal cutoffs were ≥1 point for improvement and ≤−1 point for deterioration based on both the Youden's J Index and top left methods. The sensitivities ranged between 59% and 85%, and specificities ranged between 69% and 85%.

We conducted sensitivity analyses by estimating MCID using other commonly applied approaches are shown in eTable 1. Based on the within-patient approach, the MCIDs were 0.8-2.2 for improvement and −0.8 to −2.3 for deterioration. Using the distribution approach, 0.5 standard deviation revealed MCIDs of 1.2-1.6 points, which was similar to one standard error of measurement (1.3-1.7 point).

Response determination

eTable 2 illustrates the proportion of patients in our cohort with symptom response in the followup visit (defined as change of ≥+1), which varied between 27% (nausea) and 48% (pain). In contrast, between 23% (dyspnea) and 37% (drowsiness) of patients experienced symptom deteriorate, defined as change of ≤−1.

DISCUSSION

This study is the largest study to date to identify MCID of ESAS, and is the only prospective study specifically powered to address this important question. We found that ESAS had moderate to high responsiveness to change. In sensitivity-specific analysis, a ≥1 point change was identified as the “universal” cutoff for both improvement and deterioration for all 10 symptoms. The other anchor based and distribution based approaches also showed an MCID between 1-2 for a majority of the symptoms. Our findings have implications for response determination and sample size calculations in symptom research.

Interestingly, we found that a 1 point cutoff was applicable to all 10 symptoms and for both improvement and deterioration. In a post-hoc analysis, Bush et al. reported that an average

change in ESAS Well being of 1.25 points corresponded with FACT-G change of ≥5 points, which is consistent with our findings.[28] In another secondary analysis, Reddy et al. assessed the MCID for ESAS-fatigue between baseline and day 8 in 194 cancer patients enrolled onto 2 double blind randomized controlled trials.[29] The anchor was the global benefit score, in which a score of at least 4/7 (moderately important, consistently beneficial) was considered a response. The optimal cutoff for improvement in ESAS-fatigue was identified as 4 points or more, which had a sensitivity of 66% and a specificity of 72%. The discrepancy between their findings and ours can be explained by different anchors and patient populations, and that fact that they were looking for what constituted a "≥moderate" instead of a "≥minimal" improvement.

More recently, Bedard et al. conducted a retrospective analysis of 276 cancer patients and identified the MCID using the between-patient change method for 8 ESAS symptoms with ESAS Wellbeing category change as an anchor.[14, 26] They reported that a decrease of 1.2 and 1.1 points in ESAS pain and depression, respectively, constituted clinically relevant improvement, and an increase of 1.4, 1.8, 1.1, 1.1, and 1.4 points in pain, tiredness, depression, anxiety, and appetite loss respectively, were the MCIDs for deterioration. Thus, the findings from this study were generally consistent with our prospective study despite the different methodologies.

A commonly cited MCID cutoff for pain 0-10 numeric rating scale is a ≥2/10 point or 33% decrease.[13] Importantly, this MCID was derived from data that assessed the change in pain intensity "now" measured 30 minutes from baseline, and the external criterion was the need for additional rescue opioid. This is in contrast to our study that assesses average ESAS symptom intensity over the past 24 hours between baseline and 3 weeks later. Thus, different timeframe anchors yielded different MCIDs.

In the research setting, MCID is an important determinant of sample size and whether we declare an intervention's effect clinically meaningful or not. A larger cutoff would set a higher bar for the intervention but require a smaller sample size, and vice versa. We calculated that to detect a ≥1 point improvement with 80% power and alpha 5% would require 284 patients using a two sample t-test, assuming a standard deviation of 3 points. In contrast, a 2 point cutoff would require only 72 patients. Given that the median sample size of supportive oncology randomized trials was only 70, they could potentially be underpowered if ESAS was the primary endpoint.

This study has a number of limitations. First, our MCIDs were derived from ambulatory cancer patients seen predominately in the palliative care setting with 2 visits approximately 3 weeks apart. Further prospective studies are required to determine if the MCID of ESAS varies with different patient populations, settings, and time intervals, and whether early symptom response is associated with other clinical outcomes such as quality of life improvement, cancer treatment response and survival.. Second, the sensitivity and specificity cutoffs for our MCIDs were less than 80%. A moderate sensitivity means that some patients who actually experienced a symptom change may have a lower ESAS change of <1 point (i.e. false negative), while a moderate specificity means that some patients who did not feel that their symptoms have changed in a meaningful manner may have ESAS

change of ≥1 point (i.e. false positive). Thus, MCID cutoffs are more appropriate for group averages than individual response determination. Third, we used PGI as the gold standard, which has face validity and was easy to understand by patients [30, 31]. However, other external criteria such as quality of life questionnaires and functional scales may also be useful.

In summary, ESAS was responsive to change and that the optimal cutoff for improvement/deterioration was ≥1 point for all 10 symptoms. Further studies should examine if this cutoff remains relevant in other patient populations, settings and different time frames. Findings from this study may facilitate the design and interpretation of symptom control studies employing ESAS as the primary outcome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank all the patients who participated in this study. We also thank Swati Bansal, Dr. Odai Khamash, Dr. Abdelrahman Alhawamdeh, Natalia Campacci, Camila Souza Crovador, Won Ji Yuen, Dr. Sarika Sane, and Dr. Mrunal Marathe for their assistance with data collection.

Funding: This research is supported by the Sister Institution Network Fund from the University of Texas MD Anderson Cancer Center, King Hussein Cancer Center, Barretos Cancer Hospital and Tata Memorial Center. EB is supported in part by National Institutes of Health grants R01NR010162-01A1, R01CA122292-01, and R01CA124481-01. DH is supported in part by an American Cancer Society Mentored Research Scholar Grant in Applied and Clinical Research, MRSG-14-1418-01-CCE.

References

1. Seow H, Barbera L, Sutradhar R, et al. Trajectory of performance status and symptom scores for patients with cancer during the last six months of life. J Clin Oncol. 2011; 29:1151–8. [PubMed: 21300920]
2. Hui D, Bruera E. A personalized approach to assessing and managing pain in patients with cancer. J Clin Oncol. 2014; 32:1640–6. [PubMed: 24799495]
3. Bruera E, Kuehn N, Miller MJ, et al. The Edmonton Symptom Assessment System (ESAS): a simple method for the assessment of palliative care patients. J Palliat Care. 1991; 7:6–9. [PubMed: 1714502]
4. Philip J, Smith WB, Craft P, et al. Concurrent validity of the modified Edmonton Symptom Assessment System with the Rotterdam Symptom Checklist and the Brief Pain Inventory. Support Care Cancer. 1998; 6:539–41. [PubMed: 9833305]
5. Chang VT, Hwang SS, Feuerman M. Validation of the Edmonton Symptom Assessment Scale. Cancer. 2000; 88:2164–71. [PubMed: 10813730]
6. Stromgren AS, Goldschmidt D, Groenvold M, et al. Self-assessment in cancer patients referred to palliative care: a study of feasibility and symptom epidemiology. Cancer. 2002; 94:512–20. [PubMed: 11900236]
7. Moro C, Brunelli C, Miccinesi G, et al. Edmonton symptom assessment scale: Italian validation in two palliative care settings. Support Care Cancer. 2006; 14:30–37. [PubMed: 15937688]
8. Nekolaichuk C, Watanabe S, Beaumont C. The Edmonton Symptom Assessment System: a 15-year retrospective review of validation studies (1991--2006). Palliat Med. 2008; 22:111–22. [PubMed: 18372376]
9. Barbera L, Seow H, Howell D, et al. Symptom burden and performance status in a population-based cohort of ambulatory cancer patients. Cancer. 2010; 116:5767–76. [PubMed: 21136579]

10. Dudgeon D, King S, Howell D, et al. Cancer Care Ontario's experience with implementation of routine physical and psychological symptom distress screening. Psychooncology. 2012; 21:357–64. [PubMed: 21308858]
11. Watanabe SM, Nekolaichuk CL, Beaumont C. The Edmonton Symptom Assessment System, a proposed tool for distress screening in cancer patients: development and refinement. Psychooncology. 2012; 21:977–85. [PubMed: 21671304]
12. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989; 10:407–15. [PubMed: 2691207]
13. Farrar JT, Portenoy RK, Berlin JA, et al. Defining the clinically important difference in pain outcome measures. Pain. 2000; 88:287–94. [PubMed: 11068116]
14. Hui D, Bruera E. Minimal clinically important differences in the edmonton symptom assessment system: the anchor is key. J Pain Symptom Manage. 2013; 45:e4–5.
15. Ewing JA. Detecting alcoholism. The CAGE questionnaire. JAMA. 1984; 252:1905–07. [PubMed: 6471323]
16. Breitbart W, Rosenfeld B, Roth A, et al. The Memorial Delirium Assessment Scale. J Pain Symptom Manage. 1997; 13:128–37. [PubMed: 9114631]
17. Carvajal A, Centeno C, Watson R, et al. A comprehensive study of psychometric properties of the Edmonton Symptom Assessment System (ESAS) in Spanish advanced cancer patients. Eur J Cancer. 2011; 47:1863–72. [PubMed: 21514818]
18. Kwon JH, Nam SH, Koh S, et al. Validation of the Edmonton Symptom Assessment System in Korean patients with cancer. J Pain Symptom Manage. 2013; 46:947–56. [PubMed: 23628516]
19. Fischer D, Stewart AL, Bloch DA, et al. Capturing the patient's view of change as a clinical outcome measure. JAMA. 1999; 282:1157–62. [PubMed: 10501119]
20. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. The Journal of manual & manipulative therapy. 2009; 17:163–70. [PubMed: 20046623]
21. Lauridsen HH, Hartvigsen J, Manniche C, et al. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. BMC Musculoskelet Disord. 2006; 7:82. [PubMed: 17064410]
22. Farrar JT, Pritchett YL, Robinson M, et al. The clinical importance of changes in the 0 to 10 numeric rating scale for worst, least, and average pain intensity: analyses of data from clinical trials of duloxetine in pain disorders. J Pain. 2010; 11:109–18. [PubMed: 19665938]
23. Akobeng AK. Understanding diagnostic tests 3: Receiver operating characteristic curves. Acta Paediatr. 2007; 96:644–7. [PubMed: 17376185]
24. Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J. 2007; 7:541–6. [PubMed: 17448732]
25. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003; 41:582–92. [PubMed: 12719681]
26. Bedard G, Zeng L, Zhang L, et al. Minimal Clinically Important Differences in the Edmonton Symptom Assessment System in Patients With Advanced Cancer. J Pain Symptom Manage. 2012
27. Wyrwich KW, Wolinsky FD. Identifying meaningful intra-individual change standards for health-related quality of life measures. J Eval Clin Pract. 2000; 6:39–49. [PubMed: 10807023]
28. Bush SH, Parsons HA, Palmer JL, et al. Single- vs. multiple-item instruments in the assessment of quality of life in patients with advanced cancer. J Pain Symptom Manage. 2010; 39:564–71. [PubMed: 20303030]
29. Reddy S, Bruera E, Pace E, et al. Clinically important improvement in the intensity of fatigue in patients with advanced cancer. J Palliat Med. 2007; 10:1068–75. [PubMed: 17985963]
30. Lydick E, Epstein RS. Interpretation of quality of life changes. Qual Life Res. 1993; 2:221–6. [PubMed: 8401458]
31. Juniper EF, Guyatt GH, Willan A, et al. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. J Clin Epidemiol. 1994; 47:81–7. [PubMed: 8283197]

Precis

When assessing symptom response, how much of a change in a 0-10 point numeric rating scale is considered clinically significant? In this multicenter, prospective, longitudinal study involving 796 patients with advanced cancer, we found that the minimal clinically important difference was universally a 1 point difference for both improvement and deterioration for each of the 10 symptoms in the Edmonton Symptom Assessment Scale, one of the most widely used symptom assessment batteries in oncology.

Minimal Clinically Important Difference (MCID) Determination

"There are many methods available to ascertaining an MCID. None are perfect, but all are useful."
J Sloan COPD 2005

Anchor-based methods—both the change in the patient reported outcome (PRO) of interest and an anchor (gold standard) are assessed. The MCID is determined based on the magnitude of change in the PRO that reflects the smallest clinically significant change as defined by the anchor.

- **Sensitivity-specificity approach**: different PRO change cutoffs yield different specificities and specificities when assessed against response as defined by the anchor in 2x2 tables. The MCID cutoff is based on the optimal balance between sensitivity and specificity in a receiver-operating characteristic curve.
- **Within patient changes approach**: MCID is the average PRO score among patients who reported a small change in the anchor.

Distribution-based methods—the MCID is defined by a magnitude of change in the PRO beyond what can be explained by random variations in the PRO due to patient sampling and/or the scale's reliability.

- **Standard deviation (SD) approach**: 0.3 SDs and 0.5 SDs of baseline PRO scores are often used as benchmarks for MCID.
- **Standard error of measurement (SEM) approach**: SEM is a function of SD and reliability of the PRO measure. 1 SEM is often used to estimate the MCID.

Figure 1.
Minimal Clinically Important Difference Determination.

Figure 2.

Average change in ESAS intensity between the first and second study visit by PGI category (n=796)

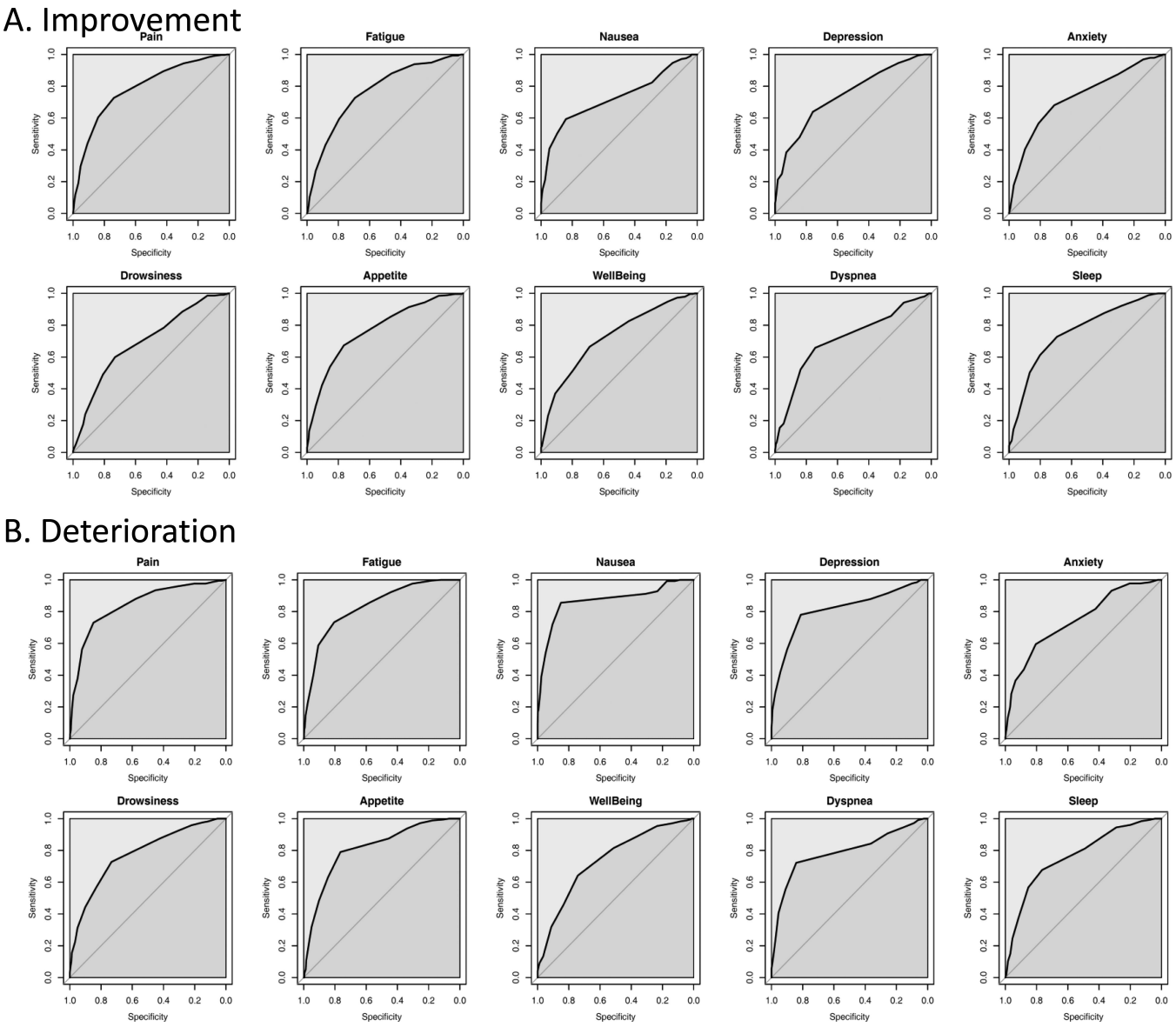

Figure 3.

Receiver-operating characteristic curves for (A) Improvement and (B) Deterioration for the 10 Edmonton Symptom Assessment Scale Symptoms. The area under the curve ranged between 0.7 to 0.86, suggesting good discrimination.

Table 1

Patient characteristics

	Brazil N=131 (%) *	Chile N=71 (%) *	India N=44 (%) *	Jordan N=182 (%) *	Korea N=68 (%) *	USA N=300(%) *	Total N=796 (%) *
Age, average (range)	58 (26-81)	60 (28-85)	51 (24-73)	55 (19-84)	59 (37-81)	58 (21-85)	57 (19-85)
Female	61 (47)	44 (62)	26 (59)	86 (47)	19 (28)	144 (48)	380 (48)
Race							
Caucasian	0	0	0	0	0	229(76)	229 (29)
Black	0	0	0	0	0	37 (12)	37 (5)
Hispanic	131 (100)	71 (100)	0	0	0	22 (7)	224 (28)
Asian	0	0	44 (100)	0	68 (100)	11 (4)	123 (15)
Other	0	0	0	182 (100)	0	1 (0)	183 (24)
Marital status							
Single	16 (12)	16 (23)	2 (5)	19 (10)	5 (7)	45 (15)	103 (13)
Married	85 (65)	40 (56)	38 (86)	139 (76)	54 (79)	200 (67)	556 (70)
Divorced	30 (23)	15 (21)	4 (9)	24 (13)	9 (13)	53 (18)	135 (17)
Education							
Illiterate	0	0	6 (14)	0	0	0	6 (1)
High school or less	114 (87)	21 (30)	32 (73)	111 (61)	52 (76)	77 (26)	407 (51)
Some college up to Bachelor's	16 (12)	47 (66)	5 (11)	58 (32)	15 (22)	173 (58)	314 (39)
Advanced degree	1 (1)	3 (4)	1 (2)	13 (7)	1 (2)	50 (17)	69 (9)
Cancer							
Breast	26 (20)	11 (15)	6 (14)	40 (22)	3 (4)	48 (16)	134 (17)
Gastrointestinal	25 (19)	26 (37)	7 (16)	31 (17)	33 (49)	68 (23)	190 (24)
Genitourinary	30 (23)	6 (8)	2 (5)	14 (8)	2 (3)	25 (8)	79 (10)
Gynecological	16 (12)	6 (8)	15 (34)	7 (4)	1 (1)	20 (7)	65 (8)
Head and neck	5 (4)	0	7 (16)	18 (10)	7 (10)	40 (13)	77 (10)
Hematological	4 (3)	8 (11)	1 (2)	10 (5)	6 (9)	8 (3)	37 (5)
Other	10 (8)	6 (8)	3 (7)	25 (14)	3 (4)	40 (13)	86 (11)

	Brazil N=131 (%)*	Chile N=71 (%)*	India N=44 (%)*	Jordan N=182 (%)*	Korea N=68 (%)*	USA N=300(%)*	Total N=796 (%)*
Respiratory	15 (11)	8 (11)	3 (7)	37 (20)	13 (19)	51 (17)	128 (16)
Stage							
Advanced	0	0	0	8 (4)	4 (6)	8 (3)	20 (3)
Locally advanced	23 (18)	8 (11)	6 (14)	9 (5)	4 (6)	34 (11)	84 (11)
Metastatic	108 (82)	63 (89)	38 (86)	165 (91)	60 (88)	258 (86)	692 (87)
CAGE positive	38 (29)	6 (8)	6 (14)	7 (4)	10 (15)	43 (14)	110 (14)
MDAS, average (SD)	2 (1)	2 (2)	2 (1)	3 (2)	2 (2)	1 (1)	2.0 (1.7)
KPS, average (SD)	71 (13)	78 (13)	77 (6)	68 (14)	77 (11)	NA	72 (13)
Duration between visits, median (Q1, Q3)	23 (21, 26)	25 (21, 32)	17.5 (14, 22)	21 (15, 28)	21 (14.5, 25)	22 (18, 28)	21 (18, 28)

*unless otherwise specified

Table 2

Changes in Edmonton Symptom Assessment Scale

	ESAS, average (SD)					Global symptom assessment		
	First clinic visit	Second clinic visit	Change*	Percentage change from first visit (%)*	P-value†	Better N (%)	Same N (%)	Worse N (%)
Pain	4.55 (3)	3.79 (3)	0.76 (3.01)	17	<0.0001	377 (47)	251 (32)	167 (21)
Fatigue	4.93 (2.82)	4.47 (2.92)	0.46 (2.95)	9	<0.0001	293 (37)	297 (37)	206 (26)
Nausea	1.72 (2.68)	1.59 (2.53)	0.14 (2.91)	8	0.10	209 (26)	460 (58)	125 (16)
Depression	2.55 (2.87)	2.38 (2.83)	0.18 (2.58)	7	0.02	169 (21)	494 (62)	132 (17)
Anxiety	3.18 (3.11)	2.82 (2.9)	0.36 (2.78)	11	0.0002	191 (24)	472 (59)	131 (16)
Drowsiness	3.31 (2.99)	3.26 (2.88)	0.05 (3.08)	2	0.31	217 (27)	410 (52)	169 (21)
Poor appetite	4.01 (3.03)	3.85 (3.06)	0.15 (3.16)	4	0.09	267 (34)	353 (44)	176 (22)
Poor well being	4.41 (2.76)	4.12 (2.74)	0.29 (2.92)	7	0.002	293 (37)	328 (41)	174 (22)
Dyspnea	2.51 (2.89)	2.23 (2.81)	0.27 (2.42)	11	0.0008	155 (19)	533 (67)	108 (14)
Poor sleep	4.08 (3.06)	3.46 (2.85)	0.62 (3.13)	15	<0.0001	258 (32)	410 (52)	128 (16)

*improvement is indicated by a positive value

†paired t-test was used to determine the difference in ESAS scores between first and second study visit

Table 3

Minimal Clinically Important Differences based on the Sensitivity-Specificity Approach

	Improvement						Deterioration					
Symptom	**Cutoff***	**Sensitivity**	**Specificity**	**Youden's J†**	**Top left‡**	**AUCψ**	**Cutoff***	**Sensitivity**	**Specificity**	**Youden's J†**	**Top left‡**	**AUCψ**
Pain	**≥+1**	**0.727**	**0.739**	**0.715**	**0.378**	0.788	**≤−1**	**0.731**	**0.849**	**0.579**	**0.309**	0.843
	≥+2	0.605	0.840	0.444	0.426		≤−2	0.563	0.922	0.485	0.444	
	≥+3	0.443	0.907	0.350	0.565		≤−3	0.377	0.951	0.328	0.625	
Fatigue	**≥+1**	**0.727**	**0.694**	**0.421**	**0.410**	0.766	**≤−1**	**0.733**	**0.805**	**0.538**	**0.331**	0.832
	≥+2	0.594	0.795	0.389	0.455		≤−2	0.587	0.907	0.494	0.423	
	≥+3	0.430	0.883	0.313	0.582		≤−3	0.398	0.939	0.337	0.605	
Nausea	**≥+1**	**0.593**	**0.841**	**0.434**	**0.437**	0.728	**≤−1**	**0.856**	**0.851**	**0.707**	**0.208**	0.868
	≥+2	0.502	0.897	0.400	0.508		≤−2	0.720	0.904	0.624	0.296	
	≥+3	0.407	0.947	0.354	0.596		≤−3	0.544	0.948	0.492	0.459	
Depression	**≥+1**	**0.639**	**0.758**	**0.397**	**0.434**	0.745	**≤−1**	**0.780**	**0.813**	**0.593**	**0.289**	**0.812**
	≥+2	0.479	0.843	0.322	0.544		≤−2	0.561	0.900	0.461	0.451	
	≥+3	0.385	0.928	0.313	0.620		≤−3	0.417	0.944	0.361	0.586	
Anxiety	**≥+1**	**0.681**	**0.711**	**0.392**	**0.431**	0.733	**≤−1**	**0.595**	**0.805**	**0.401**	**0.449**	0.746
	≥+2	0.565	0.812	0.378	0.473		≤−2	0.435	0.882	0.317	0.577	
	≥+3	0.403	0.899	0.302	0.605		≤−3	0.366	0.935	0.301	0.637	
Drowsiness	**≥+1**	**0.599**	**0.732**	**0.331**	**0.482**	0.696	**≤−1**	**0.728**	**0.733**	**0.461**	**0.381**	0.778
	≥+2	0.488	0.810	0.298	0.546		≤−2	0.562	0.834	0.396	0.468	
	≥+3	0.341	0.877	0.218	0.670		≤−3	0.444	0.901	0.345	0.565	
Poor appetite	**≥+1**	**0.673**	**0.765**	**0.438**	**0.403**	0.771	**≤−1**	**0.790**	**0.765**	**0.555**	**0.315**	0.812
	≥+2	0.538	0.854	0.391	0.485		≤−2	0.631	0.844	0.475	0.401	
	≥+3	0.421	0.905	0.326	0.587		≤−3	0.483	0.903	0.386	0.526	
Poor well being	**≥+1**	**0.664**	**0.689**	**0.354**	**0.457**	0.725	**≤−1**	**0.642**	**0.743**	**0.384**	**0.441**	0.738
	≥+2	0.510	0.800	0.310	0.529		≤−2	0.462	0.832	0.294	0.563	
	≥+3	0.370	0.910	0.280	0.637		≤−3	0.318	0.913	0.231	0.688	
Dyspnea	**≥+1**	**0.658**	**0.743**	**0.401**	**0.428**	0.712	**≤−1**	**0.722**	**0.842**	**0.564**	**0.320**	0.789

	Improvement						Deterioration					
Symptom	**Cutoff***	**Sensitivity**	**Specificity**	**Youden's J†**	**Top left‡**	**AUCψ**	**Cutoff***	**Sensitivity**	**Specificity**	**Youden's J†**	**Top left‡**	**AUCψ**
	≥+2	0.523	0.836	0.359	0.505		≤−2	0.556	0.910	0.465	0.453	
	≥+3	0.323	0.900	0.223	0.685		≤−3	0.407	0.955	0.362	0.594	
Poor sleep	**≥+1**	**0.728**	**0.693**	**0.420**	**0.411**	0.759	**≤−1**	**0.677**	**0.765**	**0.442**	**0.400**	0.765
	≥+2	0.611	0.803	0.414	0.436		≤−2	0.567	0.853	0.420	0.457	
	≥+3	0.502	0.868	0.370	0.515		≤−3	0.370	0.918	0.288	0.635	

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

*Both Youden's J method and top left method provided the same optimal cutoff. Improvement is indicated by a positive value, and deterioration is indicated by a negative value.

†The cutoff value is selected based on the largest Youden's J value, which represents the point on the ROC curve that represents the largest vertical distance from the ROC curve to the diagonal line of equality. The optimal cutoff is highlighted in bold.

‡The cutoff value is chosen based on the smallest top left value, which is on the point on the ROC curve that represents the shortest distance to the top left corner of the graph (where sensitivity = 100% and specificity = 100%). The optimal cutoff is highlighted in bold.

ψThe area under the curve for each receiver-operating characteristic curve is shown (i.e. not cutoff specific). AUC is an indicator of the ability of the scale to discriminate change. The AUCs were between 0.7 to 0.87 for the ROC curves, which suggest good discrimination.