Published in final edited form as:

J Clin Epidemiol. 2015 June ; 68(6): 655–661. doi:10.1016/j.jclinepi.2015.01.028.

DOMAIN-SPECIFIC TRANSITION QUESTIONS DEMONSTRATED HIGHER VALIDITY THAN GLOBAL TRANSITION QUESTIONS AS ANCHORS FOR CLINICALLY IMPORTANT IMPROVEMENT

Michael M. Ward, MD, MPH,

Intramural Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health.

Lori C. Guthrie, RN, BSN, and

Intramural Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health.

Maria Alba, MD

Intramural Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health.

Abstract

Objective—Estimates of minimal clinically important differences in health measures may be affected by the anchor used. We examined if domain-specific transition questions had higher construct validity than global health transition questions as anchors for measures in a given domain.

Study design and setting—In a prospective study of 249 patients with rheumatoid arthritis (RA), we examined changes in pain, physical function, joint swelling, stiffness, fatigue, and depression with treatment. We related these changes to a domain-specific transition question, global arthritis transition question, and the Short Form-36 health transition item.

Results—Changes in all six clinical measures were more highly correlated with the domain-specific transition questions than with the global arthritis question and SF-36 transition question. Discrimination between patients who improved or not was also better using domain-specific questions. Estimates of minimal clinically important improvement (MCII) differed with the anchor when these were based on mean changes. MCII estimates from receiver operating characteristic curve analysis were not influenced by the choice of anchor when anchors had high agreement.

© 2015 Published by Elsevier Inc.

Address for correspondence: Michael M. Ward, MD, NIAMS/NIH, Building 10 CRC, Room 4-1339, 10 Center Drive, Bethesda, MD 20892. Telephone (301) 496-7263; Facsimile (301) 480-2714; wardm1@mail.nih.gov..

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Financial Disclosures: M. M. Ward none; L. C. Guthrie none; M. Alba none

Conclusion—Domain-specific transition questions had higher construct validity as anchors for determining clinically important differences in health measures focused on a single domain than either global disease or general health transition questions.

Keywords

Clinically important difference; outcome measures; anchor-based methods; transition questions

INTRODUCTION

Interpretation of changes in health status measures and other patient-reported outcomes is aided by knowing the amount of change that is considered meaningful or important [1]. Distribution-based methods and anchor-based methods are the two approaches commonly used to estimate clinically important differences [2]. Anchor-based methods relate the change in the measure to an external standard (i.e. anchor) that indicates whether a recognizable change in status has occurred. These anchors may be clinical, such as reduced use of analgesics as the anchor for measures of pain severity, or the anchor may be physician or patient judgments. Patient judgments are not only convenient but may also be the most relevant. These anchors are often single-item transition questions that ask the patient to assess if his or her health has improved, worsened, or remained unchanged over some time, and usually also include an estimate of the importance of any change [3].

Diseases often impact several health domains. For example, benign prostatic hypertrophy affects not only urinary symptoms but also sleep and fatigue, and chronic arthritis commonly affects pain, physical functioning, stiffness, and work ability. To assess health among patients with these conditions comprehensively, several domain-specific measures are often used, because their specificity tends to make them more responsive than global health measures [4]. In choosing anchors to estimate clinically important differences in multiple domains, investigators must decide whether to use a series of domain-specific transition questions (e.g. Has your pain changed? Has your sleep changed?), or to apply the response to a single global transition question (e.g. Has your health changed?) to all domain measures. Domain-specific transition questions are more targeted, and would therefore presumably elicit judgments that were more closely linked to changes in the clinical measure [5-9]. However, using multiple domain-specific transition questions increases respondent burden, and a global transition question may adequately substitute, particularly if all domains are expected to respond similarly to treatment. Some investigators have used study-specific global transition questions [10-12], while others have used the health transition question of the Short Form-36 (SF-36) as the anchor, often when no other transition questions were included [13-15].

The relative validity of domain-specific transition questions and global health transition questions has not previously been examined. We compared the construct validity of six domain-specific transition questions, one global disease transition question, and the SF-36 health transition question as anchors in a study of clinically important changes in patients with active rheumatoid arthritis (RA).

METHODS

Subjects

We enrolled subjects with active RA in a prospective longitudinal study to determine estimates of minimal clinically important improvement for measures of RA activity [16]. Eligible subjects were adults with RA seen in our clinics who had at least 6 tender joints and who in the judgment of their rheumatologist required an escalation in treatment with either disease-modifying medications or prednisone. The study protocol was approved by the institutional review board and all subjects provided written informed consent.

Study assessments

We assessed subjects twice, at a baseline visit when active RA was present and treatment was escalated, and at a second visit after either one month (for those treated with prednisone) or four months (for those treated with escalation of disease-modifying medications). Each assessment included a joint count for tenderness (68 joints) and swelling (66 joints). We also collected patient-reported outcomes using written questionnaires, including a pain severity score by visual analog scale (0 = none; 100 = very severe), and morning stiffness severity score by visual analog scale (0 = none; 100 = very severe). We used the Health Assessment Questionnaire Disability Index (HAQ) as a measure of limitations in physical function [17]. The HAQ is a 20-item questionnaire that asks respondents to report the degree of difficulty they have performing tasks in eight functional areas (0 = no difficulty; 3 = unable to do). The HAQ score is the mean of the highest-rated item in each of the eight areas. We also included the SF-36, and used the vitality/energy scale as the measure of fatigue, where higher values indicate better energy and less fatigue [18]. Lastly, we used the Center for Epidemiologic Studies-Depression scale (CES-D), a 20-item questionnaire that asks the frequency of depressive symptoms over the past week (range 0 – 60, with higher scores indicating more symptoms) [19].

At the second visit, subjects also completed written transition questions about their arthritis overall and for six domains (pain, ability to do things, joint swelling, stiffness, fatigue, and depression) (Table 1). Each question asked them to judge whether the symptom or problem had improved, worsened, or stayed the same since their first visit, and to rate the importance of any change on a 7-point scale (hardly important at all; a little important; somewhat important; moderately important; a good deal important; very important; extremely important) [20]. We also used the health transition item of the SF-36 completed at the second visit as a global health transition question. The SF-36 transition item asks respondents to compare their current health to their health one year ago.

Statistical analysis

We used three approaches to compare domain-specific and global transition questions. First, to determine if responses to these questions differed, we examined agreement between each domain-specific question (improved, unchanged, or worsened) and the global arthritis transition question and the SF-36 health transition question, using weighted kappas. Second, we constructed a 15-point transition score for each transition question using the subject's rating of the importance of any change, with a possible range of -7 (extremely important

improvement) to +7 (extremely important worsening), with 0 reflecting no change. For the SF-36 transition question, the transition score ranged from -2 (much better than 1 year ago) to +2 (much worse than 1 year ago). We tested the correlation between each transition score and changes in the clinical measures using Spearman correlations. Positive correlations indicated that subjective judgments of improvement corresponded to decreases in RA activity, with correlations in the range of 0.30 – 0.50 or greater considered evidence of validity [21,22]. Third, we examined discrimination of each question using receiver operating characteristic (ROC) curves [23]. In this analysis, we used the transition question as the dependent variable to classify patients as “improved” or “not improved.” Higher ROC curve areas indicate better discrimination between subjects who are improved or not improved across all levels of change in the clinical measure.

We also conducted two analyses to examine the degree to which the choice of the anchor influenced the estimate of minimal clinically important improvements (MCII). If domain-specific transition questions demonstrated higher construct validity than global transition questions but the resultant MCII estimates were only trivially different, it may not matter which transition question was used. To determine the impact of the choice of transition question on the MCII, we first compared mean changes in clinical measures by the degree of importance of improvements for each type of transition question. For the domain-specific and global arthritis transition questions, we examined four groups: those who reported no change; those who reported improvement no more than “a little important”; those who reported improvement that was moderately or a good deal important; and those who reported improvement that was very or extremely important. For the SF-36 transition question, we examined three groups: those who reported no change; those who reported their health was somewhat better than 1 year ago; and those who reported being much better than 1 year ago. The change among those reporting the least noticeable degree of improvement has been used in previous studies as an estimate of MCII [20]. Second, we used the change in the clinical measure associated with a specificity for improvement of 0.80 on the ROC curve as the MCII estimate [16]. We used SAS programs, version 9.3 (SAS Institute, Cary, NC) for analysis.

RESULTS

Of 250 subjects enrolled, we included 249 subjects who completed the SF-36 health transition item at the second visit. Their mean ($\pm$ standard deviation) age was 50.9 ± 13.7 years, and median duration of RA was 6.2 years; 78% were women.

On the domain-specific transition questions, improvement was reported by 65% of subjects for pain, 61% for ability to do things, 63% for joint swelling, 58% for stiffness, 47% for fatigue, and 49% for depression. Improvement on the global arthritis transition question was reported by 67%, and on the SF-36 health transition question by 53%. Agreement between responses on the global arthritis transition question and the pain transition question was high (kappa 0.81), and was somewhat lower for the functional ability, joint swelling, and stiffness transition questions (Table 2). Agreement was much lower with the fatigue and depression transition questions. Agreement between the SF-36 health transition question and the domain-specific transition questions was limited, with all kappas less than 0.40.

Correlations between transition scores and changes in the clinical measures were highest for the domain-specific transition scores (Table 3). These correlations were somewhat higher than those with the global arthritis transition score for all measures but functional limitations, and substantially higher than those with the SF-36 transition score.

ROC curve areas were also highest for the domain-specific transition questions, but these were only minimally different from those of the global arthritis transition question (Table 4). ROC curve areas of the SF-36 health transition question were substantially lower.

We next examined the effect of different anchors on MCII estimates, using the pain score, swollen joint count, and the SF-36 vitality/fatigue scale as examples. These measures were chosen because their transition questions had differing degrees of agreement with the global arthritis transition question (as shown in Table 2). For the pain score, the mean change in the subgroup with the lowest level of improvement was -13.6 using the domain-specific transition question, -20.3 using the global arthritis transition question, and -21.0 using the SF-36 health transition question (Table 5). Differences in the MCII estimates were also found for the swollen joint count and SF-36 vitality/fatigue scale, indicating that there was no apparent relationship with the degree of agreement between transition questions. Of note, those who reported no change on the SF-36 health transition question had substantial improvements on each measure, in contrast to those who reported no improvement on the domain-specific or global arthritis transition questions.

Using the ROC curve method, MCII estimates were quite similar for the pain score and swollen joint count when either the domain-specific or global arthritis transition question was used as the anchor, but MCII estimates diverged for the SF-36 vitality/fatigue scale (Table 6). This pattern follows the degree of agreement between these transition questions. MCII estimates based on the SF-36 health transition question were substantially different, and had lower associated sensitivities.

DISCUSSION

Domain-specific transition questions in this study demonstrated higher correlations with changes in clinical measures, as well as better discrimination in ROC curve analysis, than either a global arthritis transition question or the health transition item of the SF-36. Although there were only marginal differences in ROC curve areas between the domain-specific transition questions and the global arthritis transition question, the correlations with clinical changes were notably higher for each domain-specific question save one. These results indicate that domain-specific transition questions have better construct validity as anchors for domain-focused measures.

Patients' judgments of improvement on the global arthritis transition question agreed with judgments on the pain, physical function, joint swelling and stiffness transition questions more than with the fatigue and depression transition questions, which likely reflect the place of pain, stiffness, swelling, and functioning as core features of RA. Changes in fatigue and depression may be more likely influenced by factors other than RA, and therefore not judged similarly as RA activity. Agreement between the domain-specific transition questions and

the SF-36 health transition item was even lower, possibly because this question relates to health in general rather than arthritis specifically, and because the time comparison extends to the prior year rather than to the start of an intervention. The poor correlation of the SF-36 health transition item with clinical changes likely also relates to the influence of factors other than RA on patients' judgments of changes in health, and to its coarser gradation. Similar findings have been reported previously. For example, correlations between the SF-36 health transition item and changes in pain measures after spinal surgery ranged from 0.26 to 0.37, and between physical and mental health transition items and the Veterans RAND-36 ranged from 0.18 to 0.21 [10,13].

The choice of the anchor affected estimates of the MCII when these were based on mean changes. Large changes in clinical measures were found among patients who reported no change on the SF-36 health transition question, highlighting the poor specificity of this item. Using ROC curves, the SF-36 transition item also had poor discrimination, and produced MCII estimates that differed greatly from those of the other anchors. Between domain-specific and global arthritis transition questions, the choice of the anchor for ROC curve analysis appeared to matter little for the two domains for which there was good agreement on changes (pain and joint swelling). This might be expected because the outcome in this analysis is collapsed to either improved or not improved. However, the choice of the anchor for vitality did appear to matter, reflecting the poorer agreement between the fatigue-specific transition question and the global arthritis transition question. These findings suggest that use of a disease-focused global transition question might be acceptable for domain-specific measures, but only if these domains are central features of the disease and an ROC curve analysis is planned. Demonstration of agreement between responses on domain-specific and global transition questions would be helpful before electing to use single global items.

Some limitations of our study should be noted. We only collected information on patient-reported transition questions, and were not able to compare these to physician judgments or clinical anchors. Transition questions (other than the SF-36 health transition question) were administered together, which may have inflated agreement among responses. Patients received different treatments, some of which may have improved some domains more than others. However, we were only interested in measuring health changes regardless of the cause of the change. Lastly, we only studied patients with RA, and we do not know if studies of other conditions would have similar results.

Although several studies have noted the limitations of subjective ratings of improvement as anchors, including their dependence not only on the amount of health change experienced but the final state achieved, these ratings remain among the most commonly used anchors [6,7,24]. Our findings indicate that studies of clinically important differences should generally use domain-specific transition questions as anchors for measures that are focused on a single health domain. The health transition item of the SF-36 had poor validity, and its use as the anchor for disease- and domain-specific measures is not recommended.

ACKNOWLEDGEMENTS

This study was supported by the Intramural Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, and U.S. Public Health Service grant AR45177.

REFERENCES

1. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR, Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc.* 2002; 77:371–83. [PubMed: 11936935]
2. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for examining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol.* 2008; 61:102–9. [PubMed: 18177782]
3. Wyrwich KW, Bullinger M, Aaronson N, et al. Estimating clinically significant differences in quality of life outcomes. *Qual Life Res.* 2005; 14:285–95. [PubMed: 15892420]
4. Wiebe S, Guyatt G, Weaver B, Matijevic S, Sidwell C. Comparative responsiveness of generic and specific quality-of-life instruments. *J Clin Epidemiol.* 2003; 56:52–60. [PubMed: 12589870]
5. Guyatt GH, Nogradi S, Halcrow S, Singer J, Sullivan MJJ, Fallen EL. Development and testing of a new measure of health status for clinical trials in heart failure. *J Gen Intern Med.* 1989; 4:101–7. [PubMed: 2709167]
6. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. *J Clin Epidemiol.* 2002; 55:900–8. [PubMed: 12393078]
7. Metz SM, Wywrich KW, Babu AN, Kroenke K, Tierney WM, Wolinsky FD. Validity of patient-reported health-related quality of life global ratings of change using structural equation modeling. *Qual Life Res.* 2007; 16:1193–202. [PubMed: 17551850]
8. Escobar A, Garcia Perez L, Herrera-Españeira C, et al. Total knee replacement: minimal clinically important differences and responders. *Osteoarthritis Cartilage.* 2013 dx.doi.org/10.1016/j.joca.2013.09.009.
9. Yost KJ, Sorensen MV, Hahn EA, Gelndenning GA, Gnanasakthy A, Cella D. Using multiple anchor- and distribution-based estimates to evaluate clinically meaningful change on the Functional Assessment of Cancer Therapy-Biologic Response Modifiers (FACT-BRM) instrument. *Value Health.* 2005; 8:117–27. [PubMed: 15804320]
10. Rose AJ, Sacks NC, Deshpande AP, Griffin SY, Cabral HJ, Kazis LE. Single-change items did not measure change in quality of life. *J Clin Epidemiol.* 2008; 61:603–8. [PubMed: 18471664]
11. Kvamme MK, Kristiansen IS, Lie E, Kvien TK. Identification of cutpoints for acceptable health status and important improvement in patient-reported outcomes, in rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. *J Rheumatol.* 2010; 37:26–31. [PubMed: 19955045]
12. Colangelo KJ, Pope JE, Peschken C. The minimally important difference for patient-reported outcomes in systemic lupus erythematosus including the HAQ-DI, pain, fatigue, and SF-36. *J Rheumatol.* 2009; 36:2231–7. [PubMed: 19723907]
13. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J.* 2008; 8:968–74. [PubMed: 18201937]
14. Perneger TV, Etter J-F, Rougemont A. Prospective versus retrospective measurement of change in health status: a community based study in Geneva, Switzerland. *J Epidemiol Community Health.* 1997; 51:320–5. [PubMed: 9229064]
15. Parker SL, Mendenhall SK, Shau D, Adogwa O, Cheng JS, Anderson WN, et al. Determination of minimum clinically important difference in pain, disability, and quality of life after extension of fusion for adjacent-segment disease. *J Neurosurg Spine.* 2012; 16:61–7. [PubMed: 21962034]
16. Ward MM, Guthrie LC, Alba MI. Clinically important changes in individual and composite measures of rheumatoid arthritis activity. Thresholds applicable in clinical trials. *Ann Rheum Dis.* 2014 doi: 10.1136/annrheumdis-2013-205079.
17. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcomes in arthritis. *Arthritis Rheum.* 1980; 23:137–45. [PubMed: 7362664]
18. Ware, JR., Jr; Kosinski, M.; Dewey, JE. How to score version 2 of the SF-36 health survey. QualityMetric Incorporated; Lincoln, RI: 2000.
19. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977; 1:385–401.

20. 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]
21. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Second Edition. Lawrence Erlbaum; Hillsdale, NJ: 1988.
22. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. *J Clin Epidemiol*. 2002; 55:900–8. [PubMed: 12393078]
23. Ward MM, Marx AS, Barry NN. Identification of clinically important changes in health status using receiver operating characteristic curves. *J Clin Epidemiol*. 2000; 53:279–84. [PubMed: 10760638]
24. Kamper SJ, Ostelo RWJG, Knol DL, Maher CG, de Vet HCW, Hancock MJ. Global perceived effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol*. 2010; 63:760–6. [PubMed: 20056385]

What is new?

- This is the first study to test if domain-specific transition questions have higher validity as anchors for estimating clinically important differences than global transition questions.
- Domain-specific transition questions were more highly correlated with changes in clinical measures than either a global disease transition question or the health transition item of the SF-36, and had better discrimination.
- Domain-specific transition questions should be used as anchors for clinical measures focused on a specific domain.

Table 1

Transition questions used in the study.

	Question	Response options	Grading of importance of changes (for those reporting improvement or worsening)
Global	Since the start of the study, OVERALL my arthritis has:	Improved Stayed the same Worsened	Almost none, hardly at all A little important Somewhat important Moderately important A good deal important Very important Extremely important
Domain-specific	Since the start of the study, my PAIN has: Since the start of the study, my ABILITY TO DO THINGS has: Since the start of the study, my JOINT SWELLING has: Since the start of the study, my STIFFNESS IN THE MORNING has: Since the start of the study, my FATIGUE has: Since the start of the study, my DEPRESSION has:	Improved Stayed the same Worsened	Almost none, hardly at all A little important Somewhat important Moderately important A good deal important Very important Extremely important
SF-36 Transition question	Compared to one year ago, how would you rate your health in general now?	Much better now than 1 year ago Somewhat better now than 1 year ago About the same as 1 year ago Somewhat worse now than 1 year ago Much worse now than 1 year ago	

Table 2

Agreement between responses on domain-specific transition questions and global arthritis transition question or the Short Form-36 health transition question.*

Comparison	Improved on both questions, N (%)	Unchanged on both questions, N (%)	Worse on both questions, N (%)	Discordant responses, N (%)	Weighted kappa (95% CI)
Pain/global arthritis	158 (63.4)	44 (17.7)	19 (7.6)	28 (11.3)	0.81 (0.75, 0.88)
Physical function/global arthritis	145 (58.2)	46 (18.5)	13 (5.2)	45 (18.1)	0.69 (0.61, 0.78)
Joint swelling/global arthritis	145 (58.2)	36 (14.5)	17 (6.8)	51 (20.5)	0.65 (0.56, 0.74)
Stiffness/global arthritis	138 (55.4)	46 (18.5)	13 (5.2)	52 (20.9)	0.65 (0.57, 0.74)
Fatigue/global arthritis	111 (44.6)	39 (15.7)	12 (4.8)	87 (34.9)	0.46 (0.37, 0.55)
Depression/global arthritis	111 (44.6)	37 (14.9)	8 (3.2)	93 (37.3)	0.42 (0.33, 0.51)
Pain/SF-36 health	113 (45.4)	28 (11.2)	19 (7.6)	89 (35.7)	0.35 (0.24, 0.45)
Physical function/SF-36 health	108 (43.4)	31 (12.5)	16 (6.4)	94 (37.7)	0.39 (0.30, 0.49)
Joint swelling/SF-36 health	104 (41.8)	25 (10.0)	20 (8.0)	100 (40.2)	0.33 (0.22, 0.43)
Stiffness/SF-36 health	99 (39.8)	28 (11.2)	14 (5.6)	108 (43.4)	0.32 (0.22, 0.42)
Fatigue/SF-36 health	82 (32.9)	33 (13.2)	18 (7.2)	116 (46.7)	0.29 (0.19, 0.39)
Depression/SF-36 health	83 (33.3)	31 (12.4)	14 (5.6)	121 (48.7)	0.27 (0.17, 0.37)

* CI = confidence interval

Table 3

Correlations between changes in clinical measures and transition scores.*

Clinical measure	Baseline value	Change during the study	Correlation		
			Domain-specific transition score	Global arthritis transition score	SF-36 health transition score
Pain (0 – 100)	60.7 ± 25.2	-21.0 ± 30.3	0.42	0.37	0.28
Health Assessment Questionnaire (0 – 3)	1.4 ± 0.7	-0.4 ± 0.6	0.47	0.46	0.38
Swollen joint count (0 – 66)	16.1 ± 8.9	-6.0 ± 8.6	0.38	0.30	0.24
Stiffness (0 – 100)	59.7 ± 27.1	-23.0 ± 30.0	0.48	0.40	0.24
Vitality/Fatigue (0 – 100)	39.2 ± 22.3	10.0 ± 21.7	0.40	0.28	0.22
Center for Epidemiologic Studies-Depression scale (0 – 60)	18.6 ± 11.7	-3.4 ± 10.0	0.39	0.30	0.23

* Values adjacent to clinical measures are possible ranges. Plus/minus values are mean ± standard deviation. Negative changes represent improvement for each clinical measure except Vitality/Fatigue, for which positive changes represent improvement. We reversed the signs for correlations with the Vitality/Fatigue scale for consistency with the other measures.

Table 4

Receiver operating characteristic curve areas for changes in six clinical measures of rheumatoid arthritis using either the domain-specific transition question, global arthritis transition question, or the Short Form-36 health transition questions as the anchor. Values are area (95% confidence interval).

Clinical Measure	Domain-specific transition question	Global arthritis transition question	Short Form-36 health transition question
Pain	0.76 (0.70, 0.83)	0.74 (0.67, 0.81)	0.64 (0.57, 0.71)
Health Assessment Questionnaire	0.79 (0.73, 0.85)	0.76 (0.70, 0.83)	0.68 (0.61, 0.75)
Swollen joint count	0.71 (0.63, 0.79)	0.69 (0.62, 0.77)	0.60 (0.53, 0.68)
Stiffness	0.79 (0.73, 0.85)	0.77 (0.70, 0.83)	0.63 (0.56, 0.70)
Vitality/Fatigue	0.66 (0.58, 0.74)	0.66 (0.59, 0.74)	0.58 (0.50, 0.65)
Depression	0.69 (0.62, 0.76)	0.68 (0.61, 0.76)	0.60 (0.53, 0.68)

Table 5

Mean (95% confidence interval) changes in pain score, swollen joint count, and Short Form-36 vitality scale by degree of importance of improvement on the domain-specific transition question, global arthritis transition question, and Short Form-36 health transition question. Patients who reported worsening are not included in this analysis.

Clinical Measure	Subgroup	Domain-specific transition question		Global arthritis transition question		Short Form-36 health transition question	
		N	Mean change	N	Mean change	N	Mean change
Pain (0 – 100)	No change	57	-8.1 (-13.4, -2.9)	58	-9.7 (-16.2, -3.3)	55	-18.0 (-10.2, -26.1)
	Improved, no more than a little important	10	-13.6 (-41.4, 14.3)	13	-20.3 (-44.4, 3.7)	64	-21.0 (-14.3, -27.9)*
	Improved, moderately or a good deal important	32	-30.9 (-40.0, -21.8)	40	-31.3 (-40.0, -22.5)	66	-33.5 (-25.6, -41.5)*
Swollen joint count (0 – 66)	Improved, very or extremely important	119	-30.9 (-36.6, -25.2)	113	-28.7 (-34.5, -22.8)		-
	No change	59	-2.3 (-4.2, -0.4)	58	-3.2 (-5.5, -1.0)	55	-4.8 (-2.7, -7.0)
	Improved, no more than a little important	10	-6.6 (-13.6, 0.4)	13	-8.4 (-13.4, -3.3)	64	-4.9 (-2.8, -7.0)*
Vitality/Fatigue (0 – 100)	Improved, moderately or a good deal important	36	-7.0 (-9.8, -4.0)	40	-7.9 (-10.4, -5.3)	66	-10.6 (-8.0, -12.1)*
	Improved, very or extremely important	111	-8.9 (-10.5, -7.3)	113	-7.7 (-9.3, -6.0)		-
	No change	97	1.7 (-7.2, 3.4)	58	2.5 (-2.6, 7.6)	55	9.1 (3.0, 15.0)
	Improved, no more than a little important	6	4.2 (-16.1, 24.5)	13	-1.3 (-14.4, 11.9)	64	5.7 (1.0, 10.4)*
	Improved, moderately or a good deal important	24	14.4 (7.4, 21.3)	40	14.7 (8.9, 20.6)	66	20.3 (14.5, 26.1)*
	Improved, very or extremely important	87	20.4 (15.1, 25.6)	113	15.4 (11.0, 19.7)		-

* These subgroups for the Short Form-36 transition question are health somewhat better than 1 year ago, and health much better than 1 year ago.

Table 6

Estimates of minimal clinically important improvement (MCII) for the pain score, swollen joint count, and Short Form-36 vitality/fatigue scale, based on receiver operating characteristic curves using either the domain-specific transition question, global arthritis transition question, or Short Form-36 health transition question as the anchor among the 249 patients.

	Domain-specific transition question		Global arthritis transition question		Short Form-36 health transition question	
	MCII (95% CI)	Sensitivity	MCII (95% CI)	Sensitivity	MCII (95% CI)	Sensitivity
Pain	-22.0 (-29.0, -15.6)	0.56	-21.0 (-28.5, -13.5)	0.55	-37.0 (-53.0, -27.0)	0.31
Swollen joint count	-8 (-12, -6)	0.46	-8 (-12, -6)	0.43	-11 (-14, -8)	0.30
Vitality/Fatigue	20.0 (15.1, 25.2)	0.50	15.0 (7.0, 21.8)	0.48	11.0 (3.1, 19.7)	0.30

Estimates were the change in the measure associated with a specificity for improvement of 0.80. CI = confidence interval.