

Establishing Minimal Important Differences for the VR-12 and SANE Scores in Patients Following Treatment of Rotator Cuff Tears

Lingjie Zhou,* MPH, Madhuri Natarajan,[†] MPH, Bruce S. Miller,[‡] MD, MS, and Joel J. Gagnier,^{§¶||} ND, MSc, PhD

Investigation performed at MedSport Clinic, Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, Michigan, USA

Background: Minimal important differences (MIDs) for the Veterans RAND 12-Item Health Survey (VR-12) and the Single Assessment Numeric Evaluation (SANE) have not been reported in patients following treatment for rotator cuff tears (RCTs).

Purpose: To determine the MIDs for the VR-12 and SANE among patients with RCT after treatment.

Study Design: Cohort study (diagnosis); Level of evidence, 2.

Method: A total of 222 patients diagnosed with RCT completed the VR-12 and SANE at baseline and then received surgical or nonsurgical treatment. After 64 weeks, 160 patients completed the VR-12, the SANE, and a global change questionnaire. We applied a distribution-based approach to estimate the MIDs for the SANE and for the physical component score (PCS) and mental component score (MCS) of the VR-12. We then used the global rating score as an anchor for 20 patients who perceived a minimal improvement, and we applied an anchor-based approach. One-half standard deviation of the baseline score was used in the distribution-based approach. Linear regression analyses and backward model selection were conducted to evaluate the associations between patients' characteristics and the anchor-based MIDs.

Results: The MIDs derived from distribution-based method estimates for the VR-12 PCS, MCS, and SANE scores were 4.94, 5.99, and 11.80, respectively. The MIDs estimated using the anchor-based method for the PCS, MCS, and SANE scores were 2.57 (90% CI, -1.62 to 6.76), 1.87 (90% CI, -2.07 to 5.80), and 27.25 (90% CI, 16.17 to 38.33), respectively. The final regression model for significant predictors of the MID on the PCS included baseline PCS ($P < .001$), body mass index ($P = .014$), symptom duration ($P = .011$), diabetes ($P = .009$), and surgery ($P = .089$). The final model for the MID on the MCS included baseline MCS ($P < .001$), patient sex ($P = .027$), and diabetes ($P = .083$). The final model for the MID on SANE included baseline SANE score ($P = .059$) and diabetes ($P = .050$).

Conclusion: This is the first study to assess the MIDs for the VR-12 and SANE scores in patients with rotator cuff disease. The estimates of MID will facilitate the interpretation and application of these outcome measures in clinical practice and research.

Keywords: minimal important difference; VR-12; SANE

^{||}Address correspondence to Joel J. Gagnier, ND, MSc, PhD, Med-Sport Domino's Farms, 24 Frank Lloyd Wright Drive, Ann Arbor, MI 48106-0391, USA (email: jgagnier@umich.edu).

*Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA.

[†]Department of Infectious Diseases and Travel Medicine, MedStar Georgetown University Hospital, Washington, DC, USA.

[‡]Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, Michigan, USA.

[§]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA.

The authors declared that they have no conflicts of interest in the authorship and publication of this contribution.

Ethical approval for this study was obtained from the University of Michigan Medical School Institutional Review Board.

The Orthopaedic Journal of Sports Medicine, 6(7), 2325967118782159

DOI: 10.1177/2325967118782159

© The Author(s) 2018

Rotator cuff tears (RCTs) are among the most common shoulder disorders, and they have a significant impact on patients. For example, RCTs substantially reduce patients' quality of life. In 2002, rotator cuff disorders accounted for more than 4.5 million physician visits and approximately 40,000 inpatient procedures in the United States.²⁰ The common symptoms of RCTs include intense pain, difficulty in getting to sleep and staying asleep, and restrictions in ability and mobility.¹⁷ These symptoms usually have a significant impact on patients' activities of daily living, leisure activities, occupation, and emotional state.¹⁷ Therefore, to assess the severity of RCTs and evaluate the effectiveness of different treatment options, we need valid and reliable instruments and appropriate criteria to measure the patient's perceived and self-reported symptoms.

The Veterans RAND 12-Item Health Survey (VR-12) and the Single Assessment Numeric Evaluation (SANE) are 2 patient-reported outcome measures (PROMs) used to evaluate the impact of RCTs on life quality and shoulder function.^{6,9,22} The VR-12 is developed from the SF-36 (36-Item Short Form Health Survey) and is used to assess health-related quality of life and estimate disease burden.⁹ The SANE score is a single numeric self-evaluation of shoulder function.⁶ These 2 PROMs enable clinicians and researchers to interpret and integrate patients' opinions in clinical decision making when treating patients with RCT.

One challenge in interpreting the results of the PROMs is to determine whether a detected change in the score is sufficient to warrant a change in treatment.²¹ A statistically significant difference does not necessarily represent a meaningful change for the patients and clinicians, who are usually more interested in the clinical relevance of the detected differences.²¹ To determine how much change is enough to be considered important to patients, the minimal important difference (MID) analysis was developed to assist in the interpretation of PROMs among other types of outcome measures. MID is defined as "the smallest difference in scores of a PROM that is perceived by patients as beneficial or harmful, which would lead the clinician to consider a change in treatment."²¹ It provides us with evidence on the effects of a treatment perceived by patients and assists health care professionals with improving the effectiveness of health care services. Previous studies have suggested that the MID may vary across populations and treatments.^{1,21,23} It is rarely appropriate to apply a single MID for all patient populations, and multivariate adjustment of baseline characteristics has been used to help reduce the bias.^{2,3,4,16} Therefore, to use MIDs in clinical and research settings to assess the perceived changes of patients with RCT, we need to establish the MID estimates in a representative study sample of patients with RCT.

Anchor-based and distribution-based methods are commonly used to estimate the MID.²¹ The anchor-based approaches apply patient-based or clinician-based variables as anchors to determine the change in health status in the target population and examine the relationship between the anchor and the PROM being investigated.²¹ To determine the MID for a PROM, the anchor-based approach uses an anchor to identify the patients who perceive a MID after receiving the treatment and then to calculate in the change in the PROM scores of those patients. For example, in our study, we used a global health questionnaire as the anchor to quantify the difference in VR-12 scores between baseline and 64 weeks. In the global health questionnaire, the patients were asked to evaluate their changes in overall health using a 5-point scale. We first specified that a score of 2 or 3 on the global health questionnaire represented a minimal important change of health for the patients. Then, using this external criterion (anchor), we could identify the patients who perceived a minimal important change in their overall health after the treatments. Therefore, the mean change in the VR-12 score of those patients was the MID for the VR-12 score. The distribution-based approaches use statistical measures of the distribution of the PROM scores to estimate the MID

and do not directly indicate the clinical relevance of the change.^{11,21} To achieve a better estimation, studies have suggested using multiple approaches to estimate the MID for a PROM.^{11,21}

To our knowledge, although VR-12 and SANE are widely used in patients with RCT, no study exists that investigates the MIDs for these measures in patients following treatment for RCT. Therefore, the primary objective of this study was to establish MIDs for the VR-12 score and SANE rating using both anchor- and distribution-based techniques in patients with RCT. Our secondary objective was to examine factors that can predict variations in the MIDs.

METHODS

Study Design

After approval by our internal review board, we used data from a prospective cohort study including 222 patients. In the distribution-based approach, the VR-12 and SANE scores of all the patients were used to determine the final MIDs. In MID analysis using the anchor-based approach, according to the results of the global health questionnaire, we included 21 patients who indicated only a minimal important improvement (patients with no, moderate, or large improvements were excluded from this analysis).

Population

The participants in our study were adults seeking treatment for RCTs. The criteria for study inclusion were as follows: (1) adult patient (≥ 18 years old); (2) unilateral RCT of any size diagnosed by use of magnetic resonance imaging or diagnostic musculoskeletal ultrasonography; (3) first-time RCT in the affected shoulder; (4) signed informed consent; and (5) availability for 12-month follow-up. Exclusion criteria were history of surgery, infection, fracture, dislocation, and open repair of the affected shoulder.

To recruit participants, clinicians identified patients experiencing RCT. Included patients needed to sign an informed consent document and fill out multiple questionnaires.

Intervention

The study patients were treated by 3 shoulder surgeons. Operative or nonoperative treatments were assigned to a patient based on the agreement of the patient and his or her physician.

Data Collection

Initially, each participant received a paper copy of the questionnaire to complete. Our clinical research coordinator was available to assist patients who had questions. The following information was collected: demographic information, smoking status, duration of symptoms, tear size, tear location, time since injury, cause of injury, medical history, physical examination results, surgical or nonsurgical treatment, and the Charlson Comorbidity Index (CCI). The CCI is

commonly used to categorize comorbidities of patients based on the International Classification of Diseases diagnostic codes.¹⁸ Patients with a higher CCI score have a higher risk of mortality or more use of health care resources.¹⁸ Outcome measure follow-ups were conducted at 64 weeks postoperatively. The follow-up questionnaire contained VR-12, SANE, and a global rating of change question.

Outcome Measures

Our outcome measures included the VR-12, the SANE rating, and a global rating of change questionnaire.

The VR-12 is one of the most commonly used health-related quality of life instruments. It consists of 12 items corresponding to 8 principal domains, including general health perception, physical functioning, role limitations due to physical and emotional problems, bodily pain, energy-fatigue, social functioning, and mental health.⁹ A physical component score (PCS) and a mental component score (MCS) are generated through summarizing the 12 items that measure physical and mental health status.⁹

The SANE rating is commonly used by North American surgeons to evaluate patients' satisfaction with their shoulder function and is easy to administer.⁶ It has only 1 question: "How would you rate your shoulder today as a percentage of normal?"⁶ The patients need to provide their responses as a numeric percentage from 0 to 100.

The global change questionnaire was given to the participants at the 64-week follow-up to identify those who perceived a minimal important change. The patients were first asked whether after the treatment, they felt better, worse, or the same; then they were asked to quantify their change on a 5-point ordinal rating scale, with 5 representing a "great deal different." We only considered the patients giving a rating of 2 or 3 points as having a MID.¹⁰

Distribution-Based Approach

A distribution-based approach is used to estimate MID based on the distribution statistics of observed PROM scores from study samples.²¹ Various distribution-based methods have been used, such as SD and effect size. Our study applied the 1/2 standard deviation (SD) of the baseline PROM score to approximate the MID, as it was suggested to be meaningful to the patients.^{19,21} For each PROM score, we divided the SDs of their baseline scores by 2 and used the results as our distribution-based MID estimations. Because only the baseline PROM scores are needed during this process, all the study participants with a baseline VR-12 or SANE score were included in the analysis.

Anchor-Based Approach

Estimation of the anchor-based MIDs requires the baseline scores, global health ratings (anchor), and final PROM scores. In addition, since the anchor-based MIDs are the score changes for only the patients with a minimal important improvement, we conducted the analysis based on the information of the patients from the minimal important change group. This group included the patients who indicated a

better overall health condition and gave a rating of 2 or 3 on the ordinal scale of the global health questionnaire.

For each PROM, we first identified all the patients in the minimal important group and subtracted their baseline score from the 64-week score to compute the individual MIDs, which were the differences between baseline and final scores. We then calculated the means of the individual MIDs. In the regression analysis, we used the individual MID as an outcome to study the associations between patients' baseline characteristics and the MID.

Statistical Analysis

We examined the demographic, treatment, and comorbidity characteristics of included patients. Paired *t* test was used to examine whether the mean changes in the VR-12 PCS, the VR-12 MCS, and the SANE from baseline to 64 weeks among the minimal improvement group were statistically significant. Then, we used simple linear regression to investigate the association between each patient's demographic, treatment, and comorbidity characteristics and the anchor-based MIDs. The response variable was the MID for each participant, and the predictor variables included age, sex, body mass index (BMI), CCI scores, baseline PROM scores, diabetes, smoking status, side of injury, symptom duration, treatment type, and tear size. In addition, to build a model that could adequately predict the MID of the VR-12 and SANE with few variables, we used backward model selection. The first step in the backward model selection process was a multivariable analysis with independent variables that were significantly associated with the outcome in simple linear regressions. From the results of simple linear regressions for each of the MIDs, the 6 variables with the lowest *P* value were chosen as candidate variables in this step. We then eliminated the variables that were not significantly associated with the outcome after controlling for other variables. Finally, we reached a regression model with only variables that were associated with the outcome after adjusting for the effects of other covariates. The criterion used for the model selection was *P* value less than .10. All the analyses were performed using SAS 9.4 software (SAS Institute).

RESULTS

The demographic and clinical characteristics of the participants are presented in Table 1. A total of 222 patients with full-thickness RCTs were recruited and completed the baseline assessments. Among those patients, 133 (59.9%) were males and 89 (40.1%) were females. Their average age was 60.5 years (90% CI, 59.4-61.5 years). The mean BMI was 30.25 kg/m² (90% CI, 29.6-30.9 kg/m²). A total of 127 patients (57.2%) received surgical treatment and 95 (42.8%) received nonsurgical treatments. Mean baseline VR-12 PCS score was 38.0 (90% CI, 36.9-39.1), and mean baseline MCS score was 50.0 (90% CI, 48.7-51.3). Mean baseline SANE score was 28.3 (90% CI, 25.7-31.0).

Of the 222 patients, 160 (72.1%) completed the 64-week assessment and evaluated their improvement in health status using a global rating of change questionnaire. Twenty-one

TABLE 1
Demographic and Baseline Characteristics^a

Characteristic	Distribution-Based Analysis ^b			Anchor-Based Analysis ^c		
	n	%	Mean (SD)	n	%	Mean (SD)
Total	222	-	-	21	-	-
Age	222	-	60.45 (9.62)	21	-	63.05 (9.49)
BMI	222	-	30.25 (6.15)	21	-	31.23 (5.47)
CCI	217	-	0.23 (0.82)	20	-	0.15 (0.49)
Baseline PCS	222	-	38.01 (9.89)	21	-	39.24 (11.26)
Baseline MCS	222	-	49.99 (11.97)	21	-	50.73 (10.86)
Baseline SANE	213	-	28.34 (23.60)	21	-	24.62 (16.08)
Sex						
Male	133	59.91	-	14	66.67	-
Female	89	40.09	-	7	33.33	-
Smoke						
Yes	28	12.61	-	4	19.05	-
No	194	87.39	-	17	80.95	-
Diabetes						
Yes	30	13.51	-	3	14.29	-
No	192	86.49	-	18	85.71	-
Injection						
Yes	87	39.19	-	8	38.10	-
No	135	60.81	-	13	61.90	-
Surgery						
Yes	127	57.21	-	6	28.57	-
No	95	42.79	-	15	71.43	-
Tear size						
Missing	2	0.90	-	-	-	-
Small/medium	159	71.62	-	12	57.14	-
Large/massive	61	27.48	-	9	42.86	-
Side						
Right	135	60.81	-	13	61.90	-
Left	87	39.19	-	8	38.10	-
Symptom duration						
≤1 year	154	69.37	-	12	57.14	-
>1 year	68	30.63	-	9	42.86	-

^a BMI, body mass index; CCI, Charlson Comorbidity Index; MCS, mental component score; PCS, physical component score; SANE, Single Assessment Numeric Evaluation.

^b All 222 patients who participated in the study were included in estimating the distribution-based minimal important difference (MID).

^c The anchor-based MIDs were estimated using information from 21 patients who indicated a minimal important improvement.

patients indicated a minimal improvement according to the global rating questionnaire. Among the participants in the minimal improvement group, 20 participants finished the 64-week VR-12 and SANE. For the participants who did not complete the 64-week questionnaire, the reasons were death, dropout, and unknown.

Distribution-Based Analysis

The MIDs estimated using the distribution-based method can be found in the Table 2. The 1/2 SD-based MID estimates for the PCS, MCS, and SANE score were 4.94, 5.99, and 11.80, respectively.

Anchor-Based Analysis

The mean MIDs between baseline and 64 weeks were 2.57 for the PCS, 1.87 for the MCS, and 27.25 for the SANE score (Table 2). The results of the paired *t* tests showed no statistically significant difference in the MCS and PCS scores between baseline and 64 weeks posttreatment. However, the difference between the baseline and 64-week SANE scores was statistically significant ($P < .001$).

In the simple linear regressions with the MID of PCS as the dependent/response variable, baseline PCS and symptom duration showed negative significant associations with the MID (Table 3), indicating that the MID decreased as baseline score or symptom duration increased. The results of the simple linear regressions for the MID of MCS indicated 3 significant negative associations: baseline MCS, female sex, and no diabetes (Table 4). In the simple linear regression analyses for the MID of SANE ratings, only diabetes and the baseline SANE score were significantly negatively associated with the MID.

From the results of simple linear regressions on the PCS, we started the backward selection modeling with the following variables: baseline PCS, diabetes, symptom duration, BMI, CCI, and surgery. The final model predicting the MID for PCS consisted of 5 significant predictors: baseline PCS, BMI, symptom duration, diabetes, and surgery (Table 5). The results of regression analyses showed that when adjusted for all of the other covariates, the MID of PCS decreased by β for every unit increase in BMI, baseline PCS, presence of diabetes, and symptoms longer than 1 year and increased for patients receiving surgery.

The backward selection modeling for the MID of MCS was begun with the following variables: baseline MCS score, sex, diabetes, surgery, age, and smoking. The final model for the MID of MCS included baseline MCS score, sex, and diabetes (Table 6). In the final multivariable model, the MID of MCS in the cohort of patients with RCT decreased with higher baseline MCS and was increased in male patients and in patients with diabetes.

For the MID of the SANE score, after backward selection was performed, the MID decreased with higher baseline SANE scores and in patients with diabetes (Table 7).

DISCUSSION

PROM questionnaires are widely used in evidence-based medicine to bring patients' opinions into the decision-making process. However, without a proper interpretation of the numbers, the results of PROMs are of little value in making decisions about treatments.¹² To determine how much of a change in a PROM matters to patients, we need to explore the MIDs of PROMs for different populations.

The current study aimed to calculate the MID for the VR-12 score and the SANE rating in patients with RCT and to identify important predictors of the MIDs. We had several reasons for selecting the VR-12 and SANE for the current study. First, both of them are simplified, validated, and widely used PROMs. The SANE score is

TABLE 2
VR-12 (MCS and PCS) Minimal Important Difference (MID)^a

	n	Mean	SD	90% CI (lower bound)	90% CI (upper bound)	P Value ^b (t test)
Anchor-based MID						
PCS	20.00	2.57	10.84	-1.62	6.76	.303
MCS	20.00	1.87	10.18	-2.07	5.80	.422
SANE	20.00	27.25	28.65	16.17	38.33	<.001
Distribution-based MID (1/2 SD)						
PCS	222.00	4.94		-	-	-
MCS	222.00	5.99		-	-	-
SANE	213.00	11.80		-	-	-

^aMCS, mental component score; PCS, physical component score; SANE, Single Assessment Numeric Evaluation; VR-12, Veterans RAND 12-Item Health Survey.

^bP values were from the t test comparing the 64-week score to the baseline score.

TABLE 3
Simple Linear Regression Analyses for VR-12 (MCS and PCS) Anchor-Based MID^a

Covariate	PCS			MCS		
	β	90% CI	P Value	β	90% CI	P Value
Age	<0.00	-0.46 to 0.45	.995	-0.20	-0.62 to 0.22	.411
BMI	-0.53	-1.31 to 0.25	.252	-0.06	-0.81 to 0.70	.895
Baseline score	-0.55	-0.86 to -0.23	.007 ^b	-0.58	-0.87 to -0.30	.003 ^b
CCI	-4.62	-13.47 to 4.24	.377	-1.27	-10.03 to 7.48	.803
Sex (ref = female)	-0.35	-9.77 to 9.07	.950	9.97	2.11 to 17.82	.041 ^b
Injection (ref = no injection)	4.14	-4.51 to 12.79	.417	-2.28	-10.50 to 5.94	.647
Surgery (ref = no surgery)	4.66	-4.56 to 13.89	.392	4.64	-4.00 to 13.28	.364
Tear size (large/massive) (ref = small/medium)	2.01	-6.63 to 10.65	.691	-3.66	-11.67 to 4.35	.439
Side (ref = left)	-2.06	-10.83 to 6.71	.688	0.18	-8.09 to 8.46	.970
Diabetes (ref = no diabetes)	-9.37	-20.84 to 2.10	.174	11.73	1.44 to 22.02	.064 ^c
Smoking (ref = no smoking)	4.45	-6.19 to 15.09	.477	-4.83	-14.77 to 5.11	.410
Symptom duration (>1 year) (ref = <1 year)	-9.71	-17.42 to -1.99	.043 ^c	1.59	-6.53 to 9.71	.738

^aBMI, body mass index; CCI, Charlson Comorbidity Index; MCS, mental component score; MID, minimal clinical difference; PCS, physical component score; ref, reference; VR-12, Veterans RAND 12-Item Health Survey.

^bP < .05.

^cP < .1.

single-item and the VR-12 has only 12 items. Using these PROMs helps relieve the burden for both the respondent and the clinician, considering the total amount of information needed in clinical research. Second, since VR-12 summarizes different health domains into physical and mental component scores, it allows us to make comparisons between the changes in patients' physical and mental health status. At the same time, the SANE score assesses the functional change of the patients. By combining the results of those 2 PROMs, we were able to provide an

assessment of patients' perceived changes on both overall health status and shoulder function.

From the results of our study, we found that the MIDs generated from the anchor-based approach were smaller than those generated from the distribution-based approach, with 1 exception: the anchor-based MID for the SANE rating was much higher than the distribution-based MID. This latter finding could be due to the relatively high amount of variability of the SANE ratings in our study sample. Because the SANE measures shoulder function,

TABLE 4
Simple Linear Regression Analyses
for SANE Anchor-Based MID^a

Covariate	β	90% CI	P Value
Age	0.56	-0.63 to 1.74	.425
BMI	0.66	-1.45 to 2.77	.595
Baseline SANE	-0.71	-1.36 to -0.06	.073 ^b
CCI	16.95	-6.95 to 40.66	.230
Sex (ref = female)	-3.93	-28.78 to 20.93	.787
Injection (ref = no injection)	9.17	-13.83 to 32.16	.498
Surgery (ref = no surgery)	5.12	-19.70 to 29.94	.725
Tear size (large/massive) (ref = small/medium)	17.12	-4.73 to 38.97	.191
Side (ref = left)	3.13	-20.14 to 26.39	.818
Diabetes (ref = no diabetes)	-33.24	-62.17 to -4.30	.062 ^b
Smoking (ref = no smoking)	-3.44	-31.94 to 25.06	.837
Symptom duration (>1 year) (ref = <1 year)	-15.00	-37.11 to 7.11	.255

^aBMI, body mass index; CCI, Charlson Comorbidity Index; MID, minimal clinical difference; ref, reference; SANE, Single Assessment Numeric Evaluation.

^b $P < .1$.

TABLE 5
Final Model for VR-12 PCS Anchor-Based MID
($R^2 = 0.80$)^a

Covariate	β	90% CI	P Value
BMI	-0.73	-1.18 to -0.27	.014
Baseline PCS	-0.59	-0.80 to -0.38	<.001
Diabetes (ref = no diabetes)	-10.96	-17.32 to -4.60	.009
Symptom duration (>1 year) (ref = <1 year)	-7.81	-12.49 to -3.13	.011
Surgery (ref = no surgery)	5.22	0.19 to 10.25	.089

^aBMI, body mass index; MID, minimal clinical difference; PCS, physical component score; ref, reference; VR-12, Veterans RAND 12-Item Health Survey.

TABLE 6
Final Model for VR-12 MCS Anchor-Based MID
($R^2 = 0.67$)^a

Covariate	β	90% CI	P Value
Baseline MCS	-0.55	-0.78 to -0.32	<.001
Sex (ref = female)	7.89	2.25 to 13.53	.027
Diabetes (ref = no diabetes)	7.68	0.43 to 14.93	.083

^aMCS, mental component score; MID, minimal clinical difference; ref, reference; VR-12, Veterans RAND 12-Item Health Survey.

TABLE 7
Final Model for SANE Anchor-Based MID ($R^2 = 0.34$)^a

Covariate	β	90% CI	P Value
Baseline SANE	-0.69	-1.29 to -0.10	.059
Diabetes (ref = no diabetes)	-32.51	-59.33 to -5.70	.050

^aMID, minimal clinical difference; ref, reference; SANE, Single Assessment Numeric Evaluation.

it may be more sensitive to the differences in the severity of RCT among patients. However, the VR-12 is a measure of health quality of life, which may not be as strongly related to RCT severity and symptoms as the SANE rating.

As far as we know, there are few published studies about the MID of the VR-12 or the SANE score.^{13,24} None of the available studies used a population of patients with RCT or investigated the associations between patients' baseline characteristics and the MID.

Kronzer et al¹³ estimated the VR-12 MID for patients undergoing elective surgery. Their anchor-based MID values for PCS and MCS were 4 and 6.5, respectively, which are much higher than the anchor-based MID estimations of our study. However, the methods used in the study by Kronzer et al¹³ are questionable. In the global assessment, they asked patients to rate their quality of life as "better," "same," or "worse" without specifying the magnitude of change. Their anchor-based MIDs were computed using the VR-12 score of all the patients who reported a "better" quality of life. The results were the mean change in scores for all the patients who perceived an improvement. Thus, the MID estimations were not an indicator of minimal important improvement and are likely to be much higher than the true value.

In our study, we were able to identify a group of patients with minimal important improvement from all patients who perceived an improvement. Therefore, our MID estimations can represent the minimal change of the PROM score for those patients with improved health. In other words, those estimates can be used as a threshold of patients' perceived health improvement.

The MIDs for the VR-12 were similar to the MID of the SF-36 reported in previous studies,^{5,25} in which the MID for PCS ranged from 1.6 to 7.0 and the MID for MCS ranged from 2.3 to 8.7. The small differences between our study and other studies could be caused by the differences in the characteristics of study samples and the approaches toward estimating MID.

In terms of the SANE score, Winterstein et al²⁴ reported MIDs of 7 for a 6-month follow-up and 19 for a 12-month follow-up, which are lower than our estimate. However, since our study has a longer follow-up period, 64 weeks (16 months), it is reasonable for the patients to have a higher MID.

Using the same patient cohort as our study, Gagnier et al⁸ established the MIDs for the American Shoulder and Elbow Surgeons (ASES) score and the Western Ontario Rotator Cuff Index (WORC). That study used similar distribution- and anchor-based approaches to the current study and also found differences in the MIDs estimated from the 2 types of approaches. However, in the regression analysis, Gagnier

et al⁸ included only age, sex, and CCI score as independent variables and did not find any statistically significant associations between these variables and the MID. In our study, we also found that age or CCI score did not predict the MID, except that sex was associated with the MID of the MCS. However, the WORC and ASES mostly measure physical health, and only 3 among the 21 items in the WORC are about emotional function. Therefore, in terms of mental health, WORC and ASES scores are not comparable with the results of the MCS. Gagnier et al⁸ reported similar results on MIDs for physical health outcome measures compared with our study. Because they did not include other demographic variables, such as diabetes and BMI, which were found in our study to be significant predictors of MID, we were unable to compare the results and will require future studies to confirm our findings.

According to the results of the paired *t* tests, for the patients in the minimal change group, the 64-week MCS and PCS scores were not statistically significantly different from the baseline scores. This indicates that even if a change in VR-12 score is not statistically significant, it might still be perceived by the patients and might be of clinical importance for the evaluation of treatment. In addition, because statistical significance is largely influenced by sample size, the nonsignificant *P* value might be related to the small sample size of the minimal improvement group. It is possible that future studies with a much larger sample size would detect a statistically significant MID for the VR-12. In terms of the paired *t* test on the SANE score, the *P* value was less than .001, which indicates that the anchor-based MID for the SANE rating is statistically significant. Given that the primary objectives of treatment for RCT are to relieve pain, restore shoulder function, and improve quality of life, which are mostly subjective, perceived changes are of greater direct importance to patients than statistically significant changes. Therefore, the MID estimates can aid in the interpretation of clinical relevance of the PROMs, and more studies are needed to validate the current results.

In the final model obtained after backward model selections, the baseline PCS, MCS, and SANE scores were negatively associated with the corresponding MIDs. Previous studies also observed negative associations between baseline scores and MIDs.^{2,7} Therefore, the baseline score can be used as a reliable predictor for the VR-12 and SANE MIDs at 64 weeks in the treatment and rehabilitation of patients with RCT. The negative associations indicate that patients with a higher baseline score perceived a smaller change in their quality of life and shoulder function as meaningful. The reason for these negative associations might be related to the higher levels of sensitivity to change in physical and mental health status and shoulder functioning in patients with a better general condition at baseline.

The final model also suggested that a higher BMI was associated with a lower PCS MID at 64 weeks, which indicates that the patients with RCT who had a higher BMI considered a smaller change in their PCS as meaningful after receiving treatments. In addition, patients who had surgery had a higher MID for PCS compared with those who did not. This suggests that patients who have surgery might have a higher expectation for improvement. Therefore, they

consider a larger change in their physical health as minimally important. A sex-based difference in MCS MID was observed, and male patients had a much higher MID compared with female patients when we controlled for baseline score and diabetes. This result suggests that female patients with RCT considered a smaller change in their mental status as significant compared with male patients after receiving treatment. However, similar results were not observed in MID for PCS. This might be due to the different underlying factors and mechanisms that influence mental health status and physical health status.

A surprising finding of the multivariable analyses was that patients with diabetes had lower PCS MID and SANE MID but a higher MCS MID, compared with those without diabetes, on average. This means that patients with diabetes perceived a smaller change in their physical health, indicated by the PCS and SANE, as meaningful and a larger change in MCS as meaningful. Based on the β coefficients, the magnitudes of the negative associations between diabetes and the MIDs of the PCS and SANE were stronger than the positive association between diabetes and the MCS. A previous study suggested that self-rated mental health was not related to chronic physical conditions, while self-rated physical health was strongly related to chronic physical conditions.¹⁴ This could explain the difference in the magnitudes of the associations.

The current study provides a range of possible MIDs for the VR-12 and SANE, estimated by use of anchor-based and distribution-based approaches. Although the anchor-based MIDs are considered to be more clinically relevant, the small sample size used in the analysis limited the precision of our estimates. Therefore, future studies with a larger sample size are needed to systematically evaluate and determine the MID of the VR-12 and SANE for patients with RCT.

Despite limitations posed by the small sample size, we established models to predict the MID for the VR-12 and SANE based on a patient's baseline characteristics. Using these models, clinicians can predict which patients have a smaller MID and are therefore more likely to achieve a clinically important improvement in their mental and physical health after receiving treatments.

This study has several strengths. First, to our knowledge, it is the first to estimate MID for the VR-12 and SANE in patients with shoulder disorders. Second, we used both distribution-based and anchor-based approaches to compute the MID, which is a recommended method to estimate MID.²¹ Third, we conducted regression analysis on MID using patients' characteristics and the treatments as predictors and built models for prediction of MID on the VR-12 and the SANE. Our results can be used to facilitate future studies and inform decision making in clinical practices. For example, at the individual level, the MID can be used as a threshold to inform the clinician regarding how much change in the VR-12 or the SANE score is sufficient to warrant a change in the treatment. At the group level, the MID estimates can be used to decide whether the difference between the 2 treatments is clinically significant.

Our study also has some limitations. First, the results of the global change questionnaire are susceptible to recall bias, which could have biased our study findings.¹⁵ Second,

the minimal important change group had only 20 patients, which may limit our statistical power, but we were careful to choose only a small number of variables in our regression modeling to minimize type I errors. Third, we used a global self-perceived health rating as the anchor for clinically significant improvements, which is of course, highly subjective.

CONCLUSION

This is the first study to establish the MID of VR-12 and SANE scores in patients with RCTs. We found that several variables, including baseline scores, BMI, sex, diabetes, symptom duration, and surgery, predicted variations in the MID. Future studies with larger sample sizes, stratified by these variables, are needed to further explore the results of the current study.

REFERENCES

- Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. *J Clin Epidemiol*. 2017;82:128-136.
- Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. *J Rheumatol*. 2002;29(1):131-138.
- Angst F. MCID - The minimal clinically important difference assigns significance to outcome effects. *J Rheumatol*. 2016;43(2):258-259.
- Bellamy N, Hochberg M, Tubach F, et al. Development of multinational definitions of minimal clinically important improvement and patient acceptable symptomatic state in osteoarthritis. *Arthritis Care Res (Hoboken)*. 2015;67(7):972-980.
- Coteur G, Feagan B, Keininger DL, Kosinski M. Evaluation of the meaningfulness of health-related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients with active Crohn's disease. *Aliment Pharmacol Ther*. 2009;29(9):1032-1041.
- Cunningham G, Lädermann A, Denard PJ, Kherad O, Burkhart SS. Correlation between American Shoulder and Elbow Surgeons and Single Assessment Numerical Evaluation score after rotator cuff or SLAP repair. *Arthroscopy*. 2015;31(9):1688-1692.
- Escobar A, García Pérez L, Herrera-Espiñeira C, et al. Total knee replacement; minimal clinically important differences and responders. *Osteoarthritis Cartilage*. 2013;21(12):2006-2012.
- Gagnier JJ, Robbins C, Bedi A, Carpenter JE, Miller BS. Establishing minimally important differences for the American Shoulder and Elbow Surgeons score and the Western Ontario Rotator Cuff Index in patients with full-thickness rotator cuff tears. *J Shoulder Elbow Surg*. 2018;27(5):e160-e166.
- Iqbal S, Rogers W, Selim A, et al. The Veterans Rand 12 Item Health Survey (VR-12): what it is and how it is used. http://www.hosonline.org/globalassets/hos-online/publications/veterans_rand_12_item_health_survey_vr-12_2007.pdf. Published 2009. Accessed January 13, 2017.
- Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415.
- Jayadevappa R, Malkowicz SB, Wittink M, Wein AJ, Chhatre S. Comparison of distribution- and anchor-based approaches to infer changes in health-related quality of life of prostate cancer survivors. *Health Serv Res*. 2012;47(5):1902-1925.
- King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res*. 2011;11(2):171-184.
- Kronzer VL, Jerry MR, Ben Abdallah A, et al. Changes in quality of life after elective surgery: an observational study comparing two measures. *Qual Life Res*. 2017;26(8):2093-2102.
- Levinson D, Kaplan G. What does self rated mental health represent. *J Public Health Res*. 2014;3(3):287.
- Metz SM, Wyrwich 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(7):1193-1202.
- Mills KA, Naylor JM, Eyles JP, Roos EM, Hunter DJ. Examining the minimal important difference of patient-reported outcome measures for individuals with knee osteoarthritis: a model using the Knee injury and Osteoarthritis Outcome Score. *J Rheumatol*. 2016;43(2):395-404.
- Minns Lowe CJ, Moser J, Barker K. Living with a symptomatic rotator cuff tear "bad days, bad nights": a qualitative study. *BMC Musculoskelet Disord*. 2014;15:228.
- Needham DM, Scales DC, Laupacis A, Pronovost PJ. A systematic review of the Charlson Comorbidity Index using Canadian administrative databases: a perspective on risk adjustment in critical care research. *J Crit Care*. 2005;20(1):12-19.
- 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(5):582-592.
- Oh LS, Wolf BR, Hall MP, Levy BA, Marx RG. Indications for rotator cuff repair: a systematic review. *Clin Orthop Relat Res*. 2007;455:52-63.
- Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102-109.
- Shelbourne KD, Barnes AF, Gray T. Correlation of a single assessment numeric evaluation (SANE) rating with modified Cincinnati knee rating system and IKDC subjective total scores for patients after ACL reconstruction or knee arthroscopy. *Am J Sports Med*. 2012;40(11):2487-2491.
- Terwee CB, Roorda LD, Dekker J, et al. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol*. 2010;63(5):524-534.
- Winterstein AP, McGuine TA, Carr KE, Hetzel SJ. Comparison of IKDC and SANE outcome measures following knee injury in active female patients. *Sports Health*. 2013;5(6):523-529.
- Zhou F, Zhang Y, Sun Y, Zhang F, Pan S, Liu Z. Assessment of the minimum clinically important difference in neurological function and quality of life after surgery in cervical spondylotic myelopathy patients: a prospective cohort study. *Eur Spine J*. 2015;24(12):2918-2923.