What Improvement in Function and Pain Intensity is Meaningful to Patients Recovering from Low-Risk Arm Fractures?

Teun Teunis, MD
Abigail Finger, BA
David Ring, MD, PhD

ABSTRACT

BACKGROUND Small, statistically significant differences in patient-reported outcomes may be of statistical importance but of questionable clinical relevance. The substantial clinical benefit (SCB) measures meaningful change to the patient, but what defines the SCB during recovery from common arm fractures is unknown. The purpose of this investigation is to assess which changes in QuickDASH and pain intensity indicate a SCB during recovery from an adequately aligned metacarpal, distal radius, or radial head fracture. Furthermore, we test the null hypothesis that there are no factors independently associated with experiencing a SCB in both measures 2 to 6 months after such fractures.

METHODS We enrolled all adult patients (n = 120) with adequately aligned metacarpal fractures, non-or minimally displaced distal radius fractures, and isolated non- or minimally displaced radial head fractures. Eighty-two patients (68%) were available at 2 to 6 months after enrollment. We measured patient demographics, symptoms of depression, self-efficacy (i.e. the belief in one's ability to successfully achieve a desired outcome), QuickDASH, numerical pain score and satisfaction with upper extremity function at enrollment and after 2 to 6 months. We used change in satisfaction with upper extremity function as the primary endpoint to identify patients who experienced a substantial clinical improvement.

RESULTS The SCB was 28 points for QuickDASH and 2 points for pain intensity. Accounting for potential cofounding variables using multivariable analysis, there were no variables associated with achieving a SCB in QuickDASH. Greater self-efficacy at enrollment was independently associated with a lower chance of experiencing a SCB in pain intensity.

CONCLUSION Less adaptive people – with lower self-efficacy – were more likely to achieve a SCB. Measures such as clinically relevant differences may aid to determine effective treatments and select treatments that help patients with less self-efficacy, whether or not those treatments effectively treat the underlying pathophysiology.

AUTHOR AFFILIATIONS

Orthopaedic Hand & Upper Extremity Service, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

CORRESPONDING AUTHOR

Teun Teunis, MD
Orthopaedic Hand & Upper Extremity Service
Massachusetts General Hospital
Harvard Medical School
55 Fruit Street
Boston, MA 02114
Phone: (617) 726-4700
Fax: (617) 724-8532
teunteunis@gmail.com

Small differences in patient-reported outcomes before and after treatment might be statistically significant but clinically irrelevant to patients.[1] The minimal clinically important or meaningful difference (MCID) was devised to address this issue.[2] The MCID is the smallest difference in a score that a patient would perceive as important. This measure is helpful for sample size calculations. Trials should aim to detect the smallest difference that a patient would find meaningful.

The MCID is calculated as the difference between people who don't improve and those

who improve minimally. It's not possible to measure a MCID when all patients are expected to improve, as with low-risk arm fractures (e.g. a nondisplaced distal radius or radial head fracture).[3] For this setting the appropriate measure of a meaningful change to patients is termed "substantial clinical benefit" (SCB). The SCB is the difference in clinical outcome scores between patients with marginal improvement as compared to those with significant improvement.[4]

The SCB for pain intensity and QuickDASH (Disabilities of the Arm, Shoulder and Hand) during recovery after a low-risk arm fracture is unknown. Also, while patient-reported outcomes after musculoskeletal trauma are associated with psychological factors,[5] it is unknown what factors are associated with experiencing a SCB. This study aims to assess which changes in QuickDASH score and pain intensity score indicate a SCB during recovery from an adequately-aligned metacarpal, distal radius, or radial head fracture. This study tests the hypotheses that there are no factors independently associated with experiencing a SCB in (1) QuickDASH and (2) pain score 2 to 6 months after such fractures.

MATERIALS & METHODS

After institutional review board approval was obtained for secondary use of previously collected data, we reviewed 120 adult (18 years or older) patients between May 2012 and September 2014. Patients were enrolled if they had an adequately aligned metacarpal fracture (n=63), non or minimally displaced distal radius fracture (n=39) and isolated non or minimally-displaced radial head fracture (n=18). Patients who were non-English speaking or unable to complete enrollment forms were excluded. Patients were enrolled at the end of their clinical visit with a hand surgeon after diagnosis was determined. Informed consent for the original study was obtained prior to enrollment. During the study period, four invited patients declined participation.

Measured variables

At the end of their clinical visit, patients chose whether or not to schedule a follow-up appointment. Subsequently, a research assistant recorded (1) patients' demographics: age, sex, marital status, employment, education, and tobacco use; (2) fracture type: metacarpal, distal radius, radial head; and (3) QuickDASH, Patient Health Questionnaire-2 (PHQ-2), Pain Self-Efficacy Questionnaires (PSEQ), pain score, overall satisfaction with their hand/arm, satisfaction with care delivered, satisfaction with the physician, overall evaluation of the visit, and if the patient felt that the initial visit could have been performed by phone or email without the need for physical examination.

Patients were contacted between two and six months after enrollment (mean 93 days ±46 days) by phone or email. At this time we recorded QuickDASH, pain score, satisfaction with upper extremity function, satisfaction with care, satisfaction with the physician, and current employment status.

We used the anchor-based approach to define the SCB.[4] The anchor-based approach uses an external criterion (the anchor) to identify substantial benefit. We used change in satisfaction with upper extremity function to identify patients who sustained a

SCB. The biggest decrease in satisfaction was eight points, while the largest increase was nine points. We selected the middle value (median) of 1.4 points as our cut-off to define clinical benefit (i.e., a median split). If satisfaction changed less than or equal to 1.4 points, we regarded this as no or limited clinical benefit. If satisfaction increased more than 1.4 points this indicated SCB.[4] Because the anchor-based method does not take into account an instrument's measurement error, we calculated the minimum detectable change in QuickDASH and pain intensity. The minimum detectable change establishes what change in outcome can reliably be measured by the instrument. If the SCB is below the minimum detectable change, it cannot be reliably measured.

Questionnaires

Disability was measured using the, QuickDASH.[6,7] This 11-question survey is scored from 0-100 with a higher score indicating greater disability.

The Patient Health Questionnaire-2 (PHQ-2)[8] measures symptoms of depression and consists of 2 questions. Scores range from 0 to 6, with higher scores indicating more depressive symptoms.

The Pain Self-Efficacy Questionnaire (PSEQ)[9] is a 10-measure questionnaire, scored from 0-60 with higher scores indicative of having more confidence performing activities while having pain.

Subjects rated pain intensity, satisfaction with upper extremity function, and overall evaluation of their visit on 11-point ordinal scales ranging from 0 (no pain, dissatisfaction, and useless) to 10 (worst pain, satisfaction, and worthwhile). Care delivered at the office and satisfaction with the physician was rated on a 5-point ordinal scales ranging from 1 (poor) to 5 (excellent).

Statistical analysis

Missing variables were estimated by median imputation for questionnaire values as described below: 13 items for PSEQ at enrollment and 5 items for QuickDASH at follow-up; and mean imputation for 1 complete QuickDASH at enrollment, 1 complete PHQ-2 at enrollment, 1 complete PSEQ at enrollment, 5 education, 4 hand and arm satisfaction, 1 satisfaction with personal manner of the physician, 2 evaluation of the visit, and 2 pain at enrollment.

We used frequencies to describe discrete variables; continuous variables are reported as means and standard deviations.

We calculated the minimum detectable change using a one-sided confidence level of 90%, similar to previous literature:

Minimum detectable change = 1.65 x $\sqrt{2}$ x standard error of measurement.[10]

The standard error of measurement is the error estimate for a single measurement. It is calculated using the following formula:

Standard error of measurement = standard deviation x $\sqrt{(1-\text{alpha})}$.[10]

We used the standard deviation of the variable at enrollment and alpha is the test-retest reliability coefficient. Enrollment QuickDASH standard deviation was 18.5 and 2.3 for pain score. The intraclass correlation coefficient for test-retest reliability of the QuickDASH is 0.96[11] and 0.88 for pain score.[12] This results in

a minimum detectable change of 8.6 for QuickDASH and 1.8 for pain score **(Table 2)**.

We determined the SCB by Receiver Operator Characteristics (ROC) curve. We considered changes in QuickDASH and pain score as diagnostic and changes in satisfaction with upper extremity function (the anchor) the gold standard. The ROC curve plots sensitivity against 1-specificity for all possible cutoff points of the change in QuickDASH and pain score. The C statistic (area under the ROC curve) ranges from 0.5 to 1, with a higher score indicating better discrimination. The optimal ROC cutoff point, representing the SCB, is the value for which the sum of percentages of false positives and false negatives is smallest. The C statistic can be used to assess how accurately the model identifies patients achieving a SCB. A C statistic of 0.5 indicates that the measure is no better than chance at predicting a result. A C statistic of greater that 0.7 is considered a reasonable predictive model and if greater than 0.8, the model is considered excellent.

Fischer's Exact test was conducted to determine the differences between categorical variables. Unpaired Student's t-test was used to determine the differences between continuous and dichotomous variables. All variables with p < 0.10 on explanatory bivariate analysis were entered into a multivariable logistic regression model **(Table 3)**. We regarded a two-tailed p value < 0.05 to be significant.

As there is no consensus regarding the optimal sample size to determine SCB a power analysis was not performed.

RESULTS

Patient Characteristics

A total of 120 patients were enrolled of whom 82 (68%) were available for followup two to six months after enrollment. Mean age was 45 (±20) years and 48 (59%) were women **(Table 1)**.

TABLE 1	Patient demographics
Variable	**Measurement**
Patients	82
Optional follow-up	46 (56%)
Age	45 (±20)
Female	48 (59%)
Marital Status	
Single	32 (39%)
Married / Living with partner	40 (49%)
Separated / Widowed	10 (12%)
Employed	63 (77%)
Education (yrs)	16 (±2.9)
Tobacco use	7 (8%)
Fracture type	
Metacarpal	41 (50%)
Distal radius	26 (32%)
Radial head	15 (18%)
Overall evaluation of visit	9.3 (±1.5)
Decisions at visit	
Change cast or splint	52 (63%)
Change protocol	5 (6.1%)
Psychosocial assistance	3 (3.7%)
None	22 (27%)
Patient health questionnaire-2	0.46 (±1.0)
Pain self-efficacy questionnaire	48 (±11)

Continuous variables as mean (±standard deviation); discrete variables as proportion (number).

TABLE 2	Change in patient reported outcomes	
Variable	**Enrollment**	**2-6 months after injury**
QuickDASH	46 (±19)	12 (±15)
Pain score	3.6 (±2.3)	1.6 (±2.4)
Satisfaction with upper extremity function	6.3 (±2.8)	8.2 (±1.8)
Satisfaction with care delivered	4.9 (±0.48)	4.6 (±0.80)
Satisfaction with personal manner of physician	4.9 (±0.32)	4.7 (±0.80)
Overall Evaluation of the visit	9.3 (±1.5)	
QuickDASH minimum detectable change	8.6	
QuickDASH substantial clinical benefit	28.3	
QuickDASH substantial clinical benefit C statistic	0.59	
Pain score substantial detectable change	1.8	
Pain score substantial clinical benefit	2.0	
Pain score substantial clinical benefit C statistic	0.65	

Variables as mean (±standard deviation). The C statistic is a measure of model fit and is the Area Under the Receiver Operating Characteristics Curve.

TABLE 3	Factors associated with achieving the substantial clinical benefit (SCB) in QuickDASH and pain score					
Variables	No SCB QuickDASH	SCB QuickDASH	P Value	No SCB Pain Score	SCB Pain Score	P Value
Patients	30	52		37	45	
Age	47 (±22)	44 (±18)	0.48	48 (±21)	43 (±18)	0.27
Sex			0.82			0.50
Male	13 (43%)	21 (40%)		17 (46%)	17 (38%)	
Female	17 (57%)	31 (60%)		20 (54%)	28 (62%)	
Marital status			0.25			0.61
Single	15 (50%)	17 (33%)		13 (35%)	19 (42%)	
Married / Living with partner	11 (37%)	29 (56%)		18 (49%)	22 (49%)	
Separated / Widowed	4 (13%)	6 (11%)		6 (16%)	4 (8.9%)	
Employed at enrollment			1.0			1.0
Yes	23 (77%)	40 (77%)		28 (76%)	35 (78%)	
No	7 (13%)	12 (13%)		9 (24%)	10 (22%)	
Education (yrs)	16 (±2.7)	16 (±3.0)	0.46	16 (±2.6)	16 (±3.2)	0.79
Smoking						0.45
Yes	7 (13%)	10 (16%)		2 (5.4%)	5 (11%)	
No	49 (87%)	54 (84%)		35 (95%)	40 (89%)	
Fracture type						
Metacarpal	15 (50%)	26 (50%)	0.95	20 (54%)	21 (47%)	0.29
Distal radius	9 (30%)	17 (33%)		13 (35%)	13 (29%)	
Radial head	6 (20%)	9 (17%)		4 (11%)	11 (24%)	
Decisions at visit			0.085			0.92
Change protocol	1 (3.3%)	4 (7.7%)		3 (8.1%)	2 (4.4%)	
Change cast or splint	15 (50%)	37 (71%)		23 (62%)	29 (64%)	
Psychosocial assistance	2 (6.7%)	1 (1.9%)		1 (2.7%)	2 (4.4%)	
None	12 (40%)	10 (19%)		10 (27%)	12 (27%)	
Optional follow-up			0.25			0.38
Yes	20 (67%)	26 (50%)		23 (62%)	23 (51%)	
No	10 (33%)	26 (50%)		14 (38%)	22 (49%)	
Patient reported outcomes						
Patient health questionnaire-2	0.47 (±1.1)	0.46 (±1.0)	0.96	0.24 (±0.60)	0.64 (±1.3)	0.085
Pain self-efficacy questionnaire	50 (±14)	47 (±8.5)	0.27	52 (±10)	45 (±10)	**0.0039**

Continuous variables as mean (±standard deviation); discrete variables as proportion (number); **bold** indicates statistically significant difference.

QuickDASH scores were 46 (±19) at enrollment and 12 (±15) after two to six months. Pain score was 3.6 (±2.3) and changed to 1.6 (±2.4). Satisfaction with upper extremity function was 6.3 (±2.8) after fracture and 8.2 (±1.8) two to six months later **(Table 2)**. Men (responder 41% [34/82] male vs. non-responder 76% [29/38] male, p< 0.001) and patients with greater symptoms of depression (mean responders PHQ-2 0.46 ± 1.0 vs. non-responders 1.2 ± 1.7, p = 0.0033) were less likely to be available for follow-up **(Appendix 1*)**.

Thresholds for Substantial Clinical Benefit

The SCB threshold was 28 points for QuickDASH and 2 points for pain intensity. Both are higher than the minimum detectable change. The C statistic of 0.59 indicated that our model for QuickDASH differentiated poorly between patients achieving a SCB and those who did not. Our model for pain score performed somewhat better, C statistic 0.65 **(Table 2)**.

Accounting for potential interaction of variables using multivariable analysis, there were no variables associated with achieving a SCB in QuickDASH **(Table 4)**.

TABLE 4	Multivariable analysis of independent variables			
	Odds ratio (95% CI)	Standard Error	P Value	C statistic
Achieving a substantial clinical benefit in QuickDASH after 2-6 mos				
Decision at visit				0.64
Change protocol	*reference value*			
Change cast or splint	0.62 (0.064 - 6.0)	0.71	0.68	
Psychosocial assistance	0.13 (0.0048 - 3.2)	0.21	0.21	
None	0.21 (0.020 to 2.2)	0.25	0.19	
Achieving a substantial clinical benefit in pain after 2-6 mos				0.74
Patient health questionnaire-2	1.2 (0.59 to 2.3)	0.22	0.7	
Patient self-efficacy questionnaire	0.93 (0.88 to 0.99)	0.029	**0.029**	

CI = confidence interval; **bold** indicates statistacally significant difference. The C statistic is a measure of model fit and is the Area Under the Receiver Operating Characteristics Curve.

Greater self-efficacy at enrollment was independently associated with a lower chance of experiencing a SCB in pain intensity (odds ratio 0.93, 95% confidence interval 0.88 to 0.99, standard error 0.029, P = 0.029, C statistic 0.74) during recovery. In other words, each point increase in self-efficacy decreased the odds of achieving a SCB in pain intensity by 0.07 **(Table 4)**. Additional analysis showed that less self-efficacy was associated with more pain at enrollment (r = -0.43, P < 0.001), but not 2 to 6 months after fracture (r = -0.16, P = 0.15).

DISCUSSION

Small statistically significant differences in patient-reported outcomes before and after treatment may not be clinically relevant. We calculated the amount of improvement in disability and pain intensity after low-risk arm fractures that patients would perceive as a substantial clinical benefit (SCB). The ability of the SCB to differentiate between people satisfied and unsatisfied with their arm function is limited. Less adaptive people (i.e. with lower self-efficacy) were more likely to achieve a SCB in pain intensity. This complicates the interpretation of minimum clinical improvement measures.

This study has some limitations. First, only 62% (82 of 120) of patients were available for a follow-up via phone or email 2 to 6 months after enrollment; however, this is not unusual for this type of prospective research studying a trauma cohort.[13] Secondly, median and mean imputations were used to calculate missing questionnaire values. While different values may have effected our results missing data was minimal and use of median and mean imputations was utilized only for a small amount of the total study data.

To our knowledge, no studies to date have assessed the SCB for QuickDASH and only one study has assessed the MCID for QuickDASH. Our SCB score was much higher (28 points on QuickDASH) than the MCID in patients with shoulder pain completing physical therapy (8 points).[14] This study found the MCID for numerical pain score to be 1.1 and found a lower C statistic for pain intensity (0.67) than QuickDASH (0.82). The SCB in back and leg pain after surgery for lumbar spinal stenosis was 2.5 points.[4]

This study found a C statistic of 0.84 for back pain and 0.82 for leg pain. The difference in SCB and C statistic for QuickDASH and numerical pain score show that clinically relevant improvement measures are likely affected by the condition being evaluated, the type of treatment, and the timing of evaluation. In other words, there may not be one clinical improvement threshold for a measurement instrument that applies to all contexts. Also the low C statistic in our study demonstrates that for the specific conditions evaluated calculation of a clinical benefit score by ROC curves cannot reliably differentiate between satisfied and unsatisfied patients.

We found no factors to be independently associated with experiencing a SCB in QuickDASH. Lower self efficacy – less belief in one's ability to successfully achieve a desired outcome – was associated with a higher change of achieving a SCB in pain intensity. We know of no study that previously assessed factors associated with achieving a clinical relevant difference in QuickDASH and pain intensity in this context. However, previous studies have found that functional outcome and pain intensity after musculoskeletal trauma is associated with stress, distress, and self-efficacy.[5]

Our results suggest that, in minor upper extremity fractures with uniformly good recovery, the biggest changes occur in the least adaptive people. This is may be due to the fact that less adaptive people have lower baseline scores. The fact that patients with more baseline limitations show greater improvement is consistent with prior research.[15] Potential reasons for this are regression to the mean, floor and ceiling effects of a questionnaire, and non-interval scales (e.g. on a 0 to 10 scale the experienced difference between 1 and 2 is potentially greater than the difference between 9 and 10). If payers, policy makers, doctors and patients utilize measures or thresholds such as clinically-relevant differences to determine effective treatments (i.e. treatments that exceed the MCID or SCB), we might select treatments that help patients with less self-efficacy, whether or not those treatments effectively treat the underlying pathophysiology. Future studies should assess the SCB for other conditions in which the majority of patients are expected to improve, and assess if psychological factors influence experiencing a SCB.

*APPENDIX

Supplementary material available from:

http://www.orthojournalhms.org/17/supplementaryMaterial.html

REFERENCES

1. Hinman RS, McCrory P, Pirotta M, et al. Acupuncture for chronic knee pain: a randomized clinical trial. JAMA. JAMA. 2014 Oct 1;312(13):1313-22. doi: 10.1001/jama.2014.12660.

2. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989 Dec;10(4):407-15.

3. Shauver MJ, Chung KC. The minimal clinically important difference of the Michigan hand outcomes questionnaire. J Hand Surg Am. 2009 Mar;34(3):509-14. doi: 10.1016/j.jhsa.2008.11.001.

4. Glassman SD, Copay AG, Berven SH, Polly DW, Subach BR, Carreon LY. Defining substantial clinical benefit following lumbar spine arthrodesis. J Bone Joint Surg Am. 2008 Sep;90(9):1839-47. doi: 10.2106/JBJS.G.01095.

5. Vranceanu AM, Bachoura A, Weening A, Vrahas M, Smith RM, Ring D. Psychological factors predict disability and pain intensity after skeletal trauma. J Bone Joint Surg Am. 2014 Feb 5;96(3):e20. doi: 10.2106/JBJS.L.00479.

6. Beaton DE, Wright JG, Katz JN, Upper Extremity Collaborative G. Development of the QuickDASH: comparison of three item-reduction approaches. J Bone Joint Surg Am. 2005 May;87(5):1038-46.

7. Niekel MC, Lindenhovius AL, Watson JB, Vranceanu AM, Ring D. Correlation of DASH and QuickDASH with measures of psychological distress. J Hand Surg Am. 2009 Oct;34(8):1499-505. doi: 10.1016/j.jhsa.2009.05.016. Epub 2009 Aug 22.

8. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care. 2003 Nov;41(11):1284-92.

9. Nicholas MK. The pain self-efficacy questionnaire: Taking pain into account. Eur J Pain. 2007 Feb;11(2):153-63. Epub 2006 Jan 30.

10. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes. 2006 Aug 22;4:54.

11. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. J Hand Ther. 2001 Apr-Jun;14(2):128-46.

12. Jackson D, Horn S, Kersten P, Turner-Stokes L. Development of a pictorial scale of pain intensity for patients with communication impairments: initial validation in a general population. Clin Med (Lond). 2006 Nov-Dec;6(6):580-5.

13. Ootes D, Buijze GA, Ring D. Predictors of missed appointments in prospective hand surgery research. Hand (N Y). 2012 Jun;7(2):177-80. doi: 10.1007/s11552-012-9411-7.

14. Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and Numeric Pain Rating Scale in patients with shoulder pain. J Shoulder Elbow Surg. 2009 Nov-Dec;18(6):920-6. doi: 10.1016/j.jse.2008.12.015. Epub 2009 Mar 17.

15. Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J. 2007 Sep-Oct;7(5):541-6. Epub 2007 Apr 2.