

University of Kentucky

UKnowledge

Physical Therapy Faculty Publications

Physical Therapy

1-2016

Specificity of the Minimal Clinically Important Difference of the Quick Disabilities of the Arm Shoulder and Hand (QDASH) for Distal Upper Extremity Conditions

Enrique V. Smith-Forbes
Fort Sam Houston Clinic, e_vsf12@hotmail.com

Dana M. Howell
University of Kentucky

Jason Willoughby
Drayer Physical Therapy Institute

Donald G. Pitts
Drayer Physical Therapy Institute

Timothy L. Uhl
University of Kentucky, tluhl2@uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/rehabsci_facpub



Part of the [Rehabilitation and Therapy Commons](#)

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Repository Citation

Smith-Forbes, Enrique V.; Howell, Dana M.; Willoughby, Jason; Pitts, Donald G.; and Uhl, Timothy L., "Specificity of the Minimal Clinically Important Difference of the Quick Disabilities of the Arm Shoulder and Hand (QDASH) for Distal Upper Extremity Conditions" (2016). *Physical Therapy Faculty Publications*. 53.

https://uknowledge.uky.edu/rehabsci_facpub/53

This Article is brought to you for free and open access by the Physical Therapy at UKnowledge. It has been accepted for inclusion in Physical Therapy Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

Specificity of the Minimal Clinically Important Difference of the Quick Disabilities of the Arm Shoulder and Hand (QDASH) for Distal Upper Extremity Conditions

Digital Object Identifier (DOI)

<http://dx.doi.org/10.1016/j.jht.2015.09.003>

Notes/Citation Information

Published in *Journal of Hand Therapy*, v. 29, no. 1, p. 81-88.

Published by Elsevier Inc. on behalf of Hanley & Belfus, an imprint of Elsevier Inc.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>.

The document available for download is the authors' post-peer-review final draft of the article.

Published by Elsevier Inc. on behalf of Hanley & Belfus, an imprint of Elsevier Inc.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Accepted Manuscript



Specificity of the Minimal Clinically Important Difference of the Quick Disabilities of the Arm Shoulder and Hand (QDASH) for Distal Upper Extremity Conditions

Major(P) Enrique, V. Smith-Forbes, PhD, OTR/L, CHT, Dana, M. Howell, PhD, OTD, OTR/L, Jason Willoughby, MHS, OTR/L, CHT, Donald, G. Pitts, MS, OTR/L, CHT, Tim, L. Uhl, PT, ATC, PhD, FNATA

PII: S0894-1130(15)00166-0

DOI: [10.1016/j.jht.2015.09.003](https://doi.org/10.1016/j.jht.2015.09.003)

Reference: HANTHE 931

To appear in: *Journal of Hand Therapy*

Received Date: 26 November 2014

Revised Date: 25 August 2015

Accepted Date: 21 September 2015

Please cite this article as: Smith-Forbes E,V, Howell D,M, Willoughby J, Pitts D,G, Uhl T,L, Specificity of the Minimal Clinically Important Difference of the Quick Disabilities of the Arm Shoulder and Hand (QDASH) for Distal Upper Extremity Conditions, *Journal of Hand Therapy* (2015), doi: 10.1016/j.jht.2015.09.003.

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 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.

1 **Specificity of the Minimal Clinically Important Difference of the Quick Disabilities of the**
2 **Arm Shoulder and Hand (QDASH) for Distal Upper Extremity Conditions**

3
4
5
6 *Authors:*

7 *Major(P) Smith-Forbes, Enrique, V., PhD, OTR/L, CHT^{a,*}*

8 *Howell, Dana, M., PhD, OTD, OTR/L^{bc}*

9 *Willoughby, Jason, MHS, OTR/L, CHT^d*

10 *Pitts, Donald, G., MS, OTR/L, CHT^d*

11 *Uhl, Tim, L., PT, ATC, PhD, FNATA^b*

12 ^a *Graduate Medical Education, Fort Sam Houston Clinic, Building 1179, Room 1A43, 3100*
13 *Schofield Road, Fort Sam Houston, TX 78234, USA*

14 ^b *Department of Rehabilitation Sciences, College of Health Sciences, University of Kentucky,*
15 *Lexington, KY, USA*

16 ^c *Department of Occupational Therapy, Eastern Kentucky University, Richmond, KY, USA*

17 ^d *Kentucky Hand & Physical Therapy/ Drayer Physical Therapy Institute, Lexington, KY, USA*

18 *This study was presented in part at the 37th American Society of Hand Therapy 2014 Meeting in*
19 *Boston, MA, and earned the ‘The First Time Scientific Session Presenter Award.’ This study*
20 *fulfilled in part the degree requirements for the first author. At the time of the study MAJ(P)*
21 *Smith-Forbes was affiliated with University of Kentucky.*

22 ** Corresponding author. Tel.: +1 832 971 7757; fax: +1 859 323 6003.*

23 *E-mail address: enrique.smith-forbes@uky.edu; e_vsf12@hotmail.com (E.V. Smith-Forbes).*

ABSTRACT: Retrospective cohort design. The minimal clinically important difference (MCID) for the quick Disabilities of the Arm, Shoulder and Hand (QDASH) has been established using a pool of multiple conditions, and only exclusively for the shoulder. Understanding diagnoses-specific threshold change values can enhance the clinical decision-making process. Before and after QDASH scores for 406 participants with conditions of surgical distal radius fracture, non-surgical lateral epicondylitis, and surgical carpal tunnel release were obtained. The external anchor administered at each fourth visit was a 15-point global rating of change scale. The test-retest reliability of the QDASH was moderate for all diagnoses: intraclass correlation coefficient model 2,1, for surgical distal radius = 0.71; non-surgical lateral epicondylitis = 0.69; and surgical carpal tunnel = 0.69. The minimum detectable change at the 90% confidence level was 25.28; 22.49; and 27.63 points respectively; and the MCID values were 25.8; 15.8 and 18.7, respectively. For these three distal upper extremity conditions, a QDASH MCID of 16-26 points could represent the estimate of change in score that is important to the patient and guide clinicians through the decision making process.

KEY WORDS: *disability evaluation, musculoskeletal diseases, outcome assessment, psychometrics, rehabilitation, upper extremity*

Level of Evidence: 2c

1 INTRODUCTION

2 The Minimal Clinically Important Difference (MCID) represents a change in score on a
3 standardized assessment that is perceived to be beneficial or harmful by the patient.¹ The MCID
4 may be calculated for patients with upper extremity (UE) deficits using two common UE
5 assessments, the quick Disabilities of the Arm, Shoulder and Hand (QDASH)² and The Global
6 Rating of Change (GROC).³ The MCID can be clinically used to interpret patient change scores
7 to guide clinical decision-making.

8 The QDASH, a region specific outcome measure, is a shortened version of the
9 Disabilities of the Arm Shoulder and Hand (DASH).⁴ Both instruments are widely used in
10 rehabilitation.^{5,6} The GROC, a generic global change scale, allow patients to decide how much
11 they have changed during recovery. The QDASH's MCID has been determined using the GROC
12 to identify those patients who have improved and comparing them to those who have not
13 improved with UE diagnoses.⁷ However, the results of these studies have generated a wide range
14 of MCID (8-20),⁷⁻¹¹ which represents 10-20% of the 100-point scale and suggests the instrument
15 may have poor responsiveness. One potential explanation for this variance may be because a
16 single diagnosis was not used in most of the previous studies.⁷ The MCID may differ among
17 diagnoses, and this may help explain the varying results in the literature.¹² This is the primary
18 rationale for examining MCID among separate diagnoses.

19 The QDASH's psychometric and clinimetric properties have been investigated. Rasch
20 analysis¹³ and classical theory¹⁴⁻¹⁶ have been used to investigate the strength and weaknesses of
21 the QDASH measures. A recent systematic review found the QDASH English version tool to
22 perform well with strong positive evidence for reliability and validity (hypothesis testing) and
23 moderate positive evidence for structural validity testing. Strong negative evidence was found

24 for responsiveness due to lower correlations with global estimates of change.¹⁷

25 Multiple approaches have been used to calculate the responsiveness of these measures.
26 The MCID current and previous values become critical in assisting providers in making clinical
27 decisions. Several authors have suggested clinicians and researchers work with a range of MCID
28 values instead of a fixed value,^{18,19} another has questioned the validity of a single overall MCID.⁸
29 Distribution-based and anchor-based methods have been the two general approaches used to
30 interpret changes. The strategy for distribution-based approaches lies in identifying the Minimal
31 Detectable Change (MDC), which is the smallest change in score that can be distinguished
32 beyond random error.²⁰ Distribution-based approaches do not give a good indication of the
33 importance of the observed change and therefore cannot provide the MCID.¹⁸ In contrast, with
34 anchor-based methods the choice of the anchor among other things will determine the precision
35 of the MCID.

36 Recent studies recommend the MCID be based primarily on anchor-based procedures,²¹
37 not be based on one study¹ and should be higher than the MDC values (the typical boundary of
38 stable patients),^{20,21} and not be based on a single study.¹ Nevertheless, there are limited studies
39 calculating the MCID through anchor-based approaches for the QDASH.⁷⁻¹⁰ Furthermore, it
40 seems the best option to determine MCID is to select a small range of threshold estimates from
41 the same sample and compare and interpret multiple reference standards.^{1,21,22} This approach has
42 been applied in a few studies on the DASH and QDASH.^{11,16} Some of the approaches to
43 calculate the MCID utilized in the literature are: 0.2 x standard deviation at baseline, 0.5 x
44 standard deviation at baseline, and one standard error of measurement (test-retest), among many
45 others.¹⁶

46 The main aim of this study was to use both anchor-based and distribution methods to

47 triangulate on MCID values for the QDASH. We used a retrospective large sample of patients
48 with UE musculoskeletal disorders who had undergone hand therapy. The objective was to
49 determine condition specific thresholds for the MCID in order to enhance confidence in
50 interpreting patient change scores for clinical decision-making.

51 **METHODS**

52 **Subjects**

53 This retrospective study population consisted of patients in a database seen at an
54 outpatient UE orthopedic condition rehabilitation multi-center, over the last 4 years. There were
55 approximately 5,000 patients in the existing database treated for multiple orthopedic conditions.
56 All data in the database was de-identified and transferred to a data sheet for study purposes and
57 then provided to the primary investigator (PI) for use by the database manager. The University of
58 Kentucky's Institutional Review Boards approved this exempt category study prior to data
59 analysis.

60 *Inclusion and Exclusion Criteria*

61 Subjects age 18-89, were included if they were not missing QDASH scores at initial visit
62 and visit 4, not missing last visit score determined per diagnoses at either visit 8 or visit 12, and
63 not missing associated GROG scores for the QDASH. Diagnoses not totaling at least 100
64 records, based on the above criterion were excluded. Surgical distal radius fracture, non-surgical
65 lateral epicondylitis, and carpal tunnel release were included as the three most common
66 conditions treated by hand therapists at these facilities.

67 **Assessment**

68 The QDASH uses 11 items to measure the degree of difficulty in performing various
69 physical activities due to a shoulder, arm, or hand problem. It utilizes a 5-point Likert scale for

70 seven functional items and three symptom items. Ten of the 11 items need to be completed for
71 the scores to be valid. The score is calculated on a 0-to-100 point scale. A higher score reflects
72 greater disability. The 2 optional scales of the QDASH (work and sport/music) are not
73 commonly collected in this clinical practice and therefore were not part of this study.

74 In contrast, the GROC scale²³ asks that a person assess his or her current health status in
75 relation to when they start their treatment and rate their level of change on a 15-point scale (-7 =
76 a very great deal worse, 0 = same, +7 = a very great deal better).²⁴ Both instruments have been
77 reported to be valid and reliable.^{2,25,26}

78 **Procedure**

79 The database was reviewed to identify the most commonly treated diagnoses. It is known
80 from review of the database that the typical number of visits for all diagnoses ranged from 8 to
81 12 visits. A screening process was used to identify that adequate scores were present at the time
82 point of interest at initial, 4th, 8th, and 12th visit (Figure 1). In addition, the range of days treated
83 was explored to determine a cutoff point for the last visit.

84 **Statistical Analysis**

85 *Descriptive statistics*

86 All statistical analyses were performed using Stata/ IC Version 13.1 (StataCorp LP,
87 College Station, TX). Baseline characteristics per diagnoses between improved and not improved
88 patients were determined for patient demographics of age, initial QDASH, and length of days in
89 care using a *t*-test for parametric data and a Wilcoxon Mann-Whitney test for nonparametric
90 data. A Chi-square test was used to calculate baseline gender differences (Table 1).⁹ Patients
91 were sub-divided per diagnoses into two groups each, stable and improved, in order to analyze
92 baseline characteristics. Stable patients were categorized from -2 to +3. Improved patients were

93 determined as reported scores on the GROC of ($\geq +4$),⁹ at visit 12 or visit eight for carpal tunnel
94 release.

95 *Validity and Reliability*

96 1) We examined *Convergent Validity* to determine the correlation between the QDASH and
97 the GROC using Pearson correlation coefficient (r). This was performed because the GROC was
98 the reference standard, or external criteria by which we judged that a real patient improvement
99 had occurred. We expected an at least a fair association ($r > 0.30$) between their final QDASH
100 score (visit eight or twelve), and their final GROC score (visit eight or twelve).

101 2) *Test-retest reliability* was calculated for the QDASH using an ANOVA (ICC_{2,2,1}) using a
102 group of stable patients on GROC (-2 to +2).⁹ In order to assess reliability, the fourth visit of the
103 QDASH was compared to the initial visit scores, as they were the earliest available repeated
104 QDASH scores.

105

106 *Responsiveness*

107 Responsiveness was determined by distribution-based and anchor-based methods.

108 a) *Distribution-based methods* determine the ability to detect change in general, and are
109 based on the statistical characteristics of the sample. We calculated the Standard Error of
110 Measurement (SEM), which links the reliability of a measurement tool to the standard deviation
111 of the population. This was obtained from an ANOVA using the entire population for the
112 diagnosis. We calculated the Minimal Detectable Change (MDC), which represents the smallest
113 change in score likely to reflect a true change, free from measurement error, ($MDC = SEM * z$ -
114 value $\sqrt{2}$.) We established a 90% confidence level (MDC₉₀) corresponding to a z -value of 1.65.
115 Meaning: If the patient has a change score greater or equal to the MDC₉₀ threshold it is possible

116 to state with 90% confidence that this change is real and not due to measurement error.

117 b) *Anchor-based methods* utilize an external patient criterion (an anchor) to determine if
118 changes in outcome are clinically meaningful. Two approaches were used; the mean change and
119 receiver-operating-characteristic (ROC) curve approaches. The GROC assessment was used as
120 the external reference in evaluating responsiveness.

121 c) *The Mean Change Approach*: Was calculated as the mean change score in the different
122 subgroups of patients who respectively reported themselves as not improved (-7 to 0), minimally
123 improved (+1 to +3), moderately improved (+4 to +5) and large changes (+6 to +7). We used
124 changes in those minimally improved to triangulate the MCID values.

125 d) *The ROC Curve Approach*: We determined the optimal cutoff score and the area under the
126 curve (AUC) considering the subjects improved with a GROC of +4 or greater.

127 A ROC curve plots sensitivity (y-axis) against 1 – specificity (x-axis). Following this rationale,
128 sensitivity was calculated as the number of patients correctly identified as improved based on the
129 cutoff value divided by all patients identified as having had a meaningful change (GROC +4 or
130 greater), whereas specificity refers to the number of patients who were correctly identified as not
131 improved based on the cutoff value divided by all patients who truly did not have a meaningful
132 change (GROC, less than +4). The optimal cutoff was chosen as the point that jointly maximized
133 sensitivity and specificity (was associated with the least amount of misclassification).

134 The AUC can be interpreted as the probability that a given diagnostic tool will correctly
135 assign a patient to the appropriate diagnostic category. In general, AUC values between 0.7 and
136 0.8 are judged as acceptable, and an AUC value greater than 0.8 is considered to have good to
137 excellent discrimination.²⁷ The greater the AUC, the greater a measure's ability to distinguish
138 patients who have improved from those who have not improved. In accordance with Turner et

139 al,²⁸ our ROC analysis will use the entire cohort, rather than just those subjects with ratings
140 adjacent to the dichotomization point to increase accuracy and obtain more reasonable estimates
141 of the MCID. We used the ICC test-retest from the product of our ANOVA that utilized a GROC
142 of (-2 to +2).⁹

143 To obtain CIs for the ROC-derived parameters, we drew 50 bootstrap samples and
144 calculated both the cutoff value and the AUC in each bootstrap replication. The mean of the 50
145 bootstrap AUC values was taken as the best estimate, with the 95% CI calculated as 1.96 . SD
146 (as an estimate of the standard error) of the bootstrap values.¹ This was done because the AUC
147 does not provide a CI, which in turn provides an estimate of how acceptable are our findings (.50
148 not good .70 acceptable, .80 good).

149 The MCID was set at the best triangulation of the results coming from both anchor-based
150 (mean change and the ROC curve) and distribution-based (the MDC₉₀ threshold) methods. This
151 is considering that the MCID should be based primarily on anchor-based procedures²¹ and be
152 higher than the MDC value. In this regard, the MDC should be interpreted as another piece in the
153 puzzle toward establishing the MCID, by benchmarking it to the boundaries of error.¹¹

154 According to Turner et al,²⁰ “if the two anchor-based methods calculated on the same
155 population yield different MCID values, then the knowledge that one value is below the MDC
156 could aid in the decision to select the other.” In addition, the ROC-curve approach was preferred
157 as the first choice as it successfully addresses most limitations of the mean change
158 approach.^{1,21,28} Furthermore, our calculation of the 95% CIs gave a useful indication of the sam-
159 pling variation.¹⁸

160 **RESULTS**

161 **Descriptive Statistics and Validity of the Measures**

162 After excluding for missing data, 406 patients met inclusion criteria for three diagnoses;
163 surgical distal radius fracture (n = 151), non-surgical lateral epicondylitis (n = 137), and carpal
164 tunnel release (n = 118). Most demographical data yielded no significant differences between
165 improved and not improved groups with exception of lower initial QDASH scores for the
166 improved group for surgical distal radius fracture, $P = .006$ and gender for carpal tunnel release,
167 $P = .04$, see Table 1. Scores for the QDASH (initial and last visit), last visit GROC, as well as
168 cutoff treatment sessions and duration of treatment days are presented in Table 2. Based on a
169 previous study consisting of multiple diagnoses, with an average duration of 10 visits /22 days,¹¹
170 a cutoff of 12 visits was chosen for surgical distal radius fracture and non-surgical lateral
171 epicondylitis. A cutoff of 8 visits for carpal tunnel release occurred due to a shorter duration, see
172 Table 2. Mean score changes for the QDASH questionnaire according to each GROC grade are
173 shown in Table 3.

174 The correlation between GROC and the score changes of the QDASH was significant for
175 all three diagnoses with a fair relationship for surgical distal radius fracture ($r = 0.39$, $P < 0.001$)
176 and for non-surgical lateral epicondylitis ($r = 0.39$, $P < 0.001$), and a weak, but significant
177 relationship for carpal tunnel release ($r = 0.22$, $P = 0.029$.) The test-retest reliability using a
178 group of stable patients on GROC (-2 to +2), had moderate agreement for all three diagnoses
179 surgical distal radius fracture: $ICC_{2,1} = 0.71$, (95% CI: 0.51, 0.83)- non-surgical lateral
180 epicondylitis: 0.69, (95% CI: 0.56, 0.79)- and carpal tunnel release: 0.69, (95% CI: 0.43, 0.84).

181 **Responsiveness**

182 *Distribution-based methods*

183 For the surgical distal radius fracture the SEM was 10.83 and the MDC_{90} corresponded to
184 25.28, for the non-surgical lateral epicondylitis the SEM was 9.63, and the MDC_{90} was 22.49;

185 and for the carpal tunnel release the SEM was 11.84, and the MDC_{90} was 27.63.

186 *Anchor-based methods*

187 The mean changes for the QDASH, per diagnoses, are reported in Table 3. In particular
188 those patients who were rated as having a small improvement (GROC, +1 to +3) had a mean
189 change improvement for surgical distal radius fracture of 25.8 points (95% CI: 14.4, 35.6) for the
190 QDASH; for non-surgical lateral epicondylitis of 15.3 points (95% CI: 11.4, 19.1); and for carpal
191 tunnel release of 18.7 points (95% CI: 8.5,25.2). Splitting the data according to a presence of
192 moderate or larger improvement ($\geq +4$) versus the remainder of the entire cohort, the AUC for
193 the QDASH for surgical distal radius fracture was 0.66 (95% CI: 0.56, 0.77), (Figure 2); 0.64,
194 (95% CI: 0.55, 0.73), (Figure 3); and for carpal tunnel release 0.66, (95% CI: 0.55, 0.77), (Figure
195 4). The ROC-curve cutoff scores that best identified meaningful improvement in clinical status
196 (as measured by GROC values of +4 or greater) for surgical distal radius fracture 15.8 points
197 (95% CI: -5.3, 36.9); for non-surgical lateral epicondylitis 15.8 points (95% CI: 1.0, 30.6) points;
198 and for carpal tunnel release 13.3points (-1.7, 28.3) for the QDASH.

199 *Surgical distal radius fracture triangulation*

200 We took into account the following data (a) an MDC_{90} of 25.28 points for the QDASH,
201 (b) a mean change for small improvement of 25.8 points for the QDASH, and (c) an ROC cutoff
202 score that best identified meaningful improvement in clinical status of 15.8 points (sensitivity
203 86%, specificity 37%, correctly classified 74%), for the QDASH. Analyzing the overall results
204 we had two competing anchor-based methods, the mean change = 25.8 and the ROC = 15.8.
205 Based on Turner et al,²⁰ recommendations, the MCID = 25.8, was selected since it was just right
206 over the MDC_{90} = 25.28 points.

207 *Non-surgical lateral epicondylitis triangulation*

208 We took into account the following data (a) an $MDC_{90} = 22.49$ points for the QDASH,
209 (b) a mean change for small improvement of 15.3 points for the QDASH, and (c) an ROC cutoff
210 score that best identified meaningful improvement in clinical status of 15.8 points (sensitivity
211 65%, specificity 59%, correctly classified 63%) for the QDASH. Analyzing the overall results
212 our two anchor-based methods yielded similar results, the mean change = 15.3 and the ROC =
213 15.8. However, both values were lower than the MDC_{90} of 22.49 points. Therefore, we selected a
214 MCID = 15.8 points from the AUC since it was the closest to the MDC_{90} .

215 *Carpal tunnel release triangulation*

216 We took into account the following data: (a) an MDC_{90} of 27.63 points for the QDASH,
217 (b) a mean change for small improvement of 18.7 points for the QDASH, and (c) an ROC cutoff
218 score that best identified meaningful improvement in clinical status of 13.3 points (sensitivity
219 76%, specificity 50%, correctly classified 69%) for the QDASH. Analyzing the overall results
220 we had competing values of mean change = 18.7, and an ROC = 13.3 points. However, again
221 both values were lower than the MDC_{90} of 27.63 points. Therefore, we selected a MCID = 18.7
222 points from the mean change approach, since it was the closest to the MDC_{90} .

223 **DISCUSSION**

224 In this era of evidence-based medicine, patients, clinicians and third-party payers demand
225 to know the effectiveness of therapeutic interventions. This study contributes to the body of
226 knowledge on the psychometric properties of the QDASH by examining the MCID for three
227 distal upper extremity conditions: surgical distal radius fracture, non-surgical lateral
228 epicondylitis, and carpal tunnel release.

229 In order to assess reliability, the fourth visit of the QDASH was compared to the initial
230 visit scores, as they were the earliest available repeated QDASH scores. The average time from

231 the initial to fourth QDASH visit were 9 ± 3 days for surgical distal radius fracture, 10 ± 6 days
232 for non-surgical lateral epicondylitis, and 11 ± 7 days for the carpal tunnel release. The test-retest
233 reliability for all three diagnoses ranged from 0.69 to 0.71, indicating moderate agreement.
234 Mintken et al., found a higher reliability of 0.90 examining a cohort of shoulder patients.⁹
235 Although, in our study the average length of days between tests was 10 days, which may have
236 contributed to recall bias. In Mintken et al's., study the average length of follow-up time was
237 even larger at 27 days.

238 This study used anchor-based and distribution-based methods to triangulate and assess
239 the MCID for the QDASH on three diagnoses: surgical distal radius fracture, non-surgical lateral
240 epicondylitis, and carpal tunnel release. During the triangulation of our results we considered
241 that the MCID should be based primarily on anchor-based procedures, and in the first instance on
242 the ROC curve,^{11,21,29} and if possible, to be higher than the MDC value.³⁰

243 Regarding the distribution-based approach, in our sample the MDC_{90} for all three
244 diagnoses was larger than the ROC calculated values. This is not uncommon^{9,31} as distributional
245 approaches are complicated by competing suggestions for the "beyond error" thresholds (e.g., 1,
246 1.96, or 2.77 SEM).^{32,33} Some authors have recommended a more reliable method to estimate the
247 MDC is to calculate 0.5 of the SD or 1 SEM.²⁰ Applying this method, all our MDC_{90} 's would fall
248 below the ROC calculated values. For the three diagnoses, the MDC_{90} values obtained were
249 above 20 points, and were larger than what is commonly reported in the literature. One reason
250 may be due to the retrospective nature of the data as higher quality control could have been
251 provided in a prospective study design. Nevertheless, one strength of this study was that all data
252 were collected on patients being treated in the course of normal hand therapy. The retrospective
253 nature is a limitation, but it is more indicative of a real and typical clinical result as this is exactly

254 what it is. Patients may or may not participate in a study due to time limitation. However, these
255 data were collected as a standard operation procedure and were extracted after the fact. This data
256 has strong external validity due to the manner in which it was originally collected.

257 The MCID measures important change because it uses a patient generated anchor for
258 comparison. In contrast, the MDC measures statistical distribution of margins of error.²⁰
259 Following Turner et al's recommendation, the MDC₉₀ was regarded as a benchmark to establish
260 margins of error for the MCID, and in our sample it represented the higher bound.^{11,20} Regarding
261 the anchor-based method, the first concern about the appropriateness of the cutoff values is the
262 selection of the anchor. We used a 15-point anchor (-7 = a very great deal worse, 0 = same, +7 =
263 a very great deal better) and considered patients +4 to +7 as significantly improved and others as
264 not significantly improved, to utilize the entire cohort.²⁸ There is no agreement in the literature
265 on what type of GROCs to use, which groups to include in the analysis, or the level at which to
266 dichotomize.^{11,28} Furthermore, different standards have been used to determine and select the
267 cutoff values for the QDASH.^{2,9-11} In addition, it is difficult to make any direct comparisons to
268 MCID's due to the methods employed including the choice of anchor, decision rules and types of
269 calculation procedures.^{11,20} In our sample, we found the ROC yielded values that were smaller
270 than the mean change approach within each category of small, moderate, and large changes, with
271 one exception (small changes for non-surgical lateral epicondylitis) which is similar to the MCID
272 review findings by Turner et al.²⁰ See Table 3.

273 We found the ROC values to fall within previously established MCID estimates for the
274 QDASH ranging from 8 to 20 points.¹¹ In particular, two of our ROC values of 15.8 points for
275 the surgical distal radius fracture and non-surgical lateral epicondylitis were similar to recent
276 estimates by the Franchignoni group at 15.91 points.¹¹ However, based on the recommended

277 methods of triangulation in the literature, the ROC value was only selected for non-surgical
278 lateral epicondylitis. After triangulation, only one of our MCID values (post-surgical distal
279 radius fracture, 25.8 points) fell outside the upper limit of 20 points reported in the literature.
280 Overall, one benefit of this sample is that it is one of the largest groups of patients to examine the
281 responsiveness of the QDASH.

282 In a recent review measuring clinical outcomes for distal radius fractures, pain and
283 function were regarded as the primary domains out of seven core areas of recommendations.³⁴
284 Considering this, in our study one explanation for a larger MCID for the two post-surgical
285 diagnoses, may be the perceived initial pain and edema restrictions from the surgical
286 intervention. Patients can be limited by the anticipation of pain and expectations of decreased
287 function following surgery.³⁵ Therefore, patients may perceive the need to regain greater ROM
288 and decrease pain before they can report a minimal improvement in their status. This reasoning is
289 supported by another study that examined patient satisfaction with outcomes after surgical distal
290 radius fractures.³⁶ That study concluded patients need to regain greater wrist range of motion
291 than what is necessary to perform activities of daily living, to be satisfied with treatment
292 outcomes.³⁶

293 **Limitations**

294 Patient baseline status and patient demographics can significantly affect MCID scores.³⁷
295 In our study there were significant baseline QDASH differences for surgical distal radius
296 fracture, $P = .006$; and gender for carpal tunnel release, $P = .04$. Therefore, the MCID should be
297 interpreted with caution. It is important to note the MCID will fluctuate based on what is
298 important to the patient, as it is not a fixed value,³⁷ and will vary based on the method chosen to
299 determine the MCID, as well as the type of population.²⁹ For this reason, the results of this study

300 can only be generalized to those groups of patients and individuals with similar characteristics to
301 this sample.³¹ In addition, the use of the GROC may have introduced recall bias and the use of a
302 retrospective sample, without pre-existing controls, may explain the large MDC₉₀ obtained for
303 each diagnosis as above indicated.

304 **CONCLUSION**

305 This study proposes the specific MCID values for the surgical distal radius fracture, non-
306 surgical lateral epicondylitis, and carpal tunnel release diagnoses, based on a comprehensive
307 triangulation of anchor-based and distribution-based approaches.¹¹ Based on triangulation
308 rules,^{1,16,20,21} we selected MCID values of 25.8 points for surgical distal radius fracture, 15.8
309 points for non-surgical lateral epicondylitis, and 18.7 points for carpal tunnel release. The
310 respective MDC₉₀ values can serve as margins of error²⁰ for surgical distal radius fracture
311 (25.28), non-surgical lateral epicondylitis (22.49) and carpal tunnel release (27.63) points for the
312 QDASH. We agree with other studies noting a need of the standardization of the MCID
313 methodology.^{11,20,29}

314 **Clinical Implications**

315 Clinicians can use these MCID scores for the surgical distal radius fracture, non-surgical
316 lateral epicondylitis and carpal tunnel release to understand how much change represents a
317 meaningful change to a patient with these specific diagnoses. Previously reported QDASH
318 MCID values ranged from 8- 20 points.⁷⁻¹¹ The results from this study indicate a MCID range of
319 16 to 26 points represents the minimal clinical change meaningful to patients presenting with
320 three specific elbow and wrist conditions. Specifically, post-surgical distal radius fracture
321 patients may need to have a larger improvement (25.8 points) than previously reported using a
322 pool of conditions (up to 20 points). These diagnoses specific MCID's can help guide decision-

323 making during the course of treatment. The selected MCID's serve as a gauge on how much
324 change a patient may need to undergo to experience a true change during the course of treatment,
325 while the MDC_{90} 's serve as error margins to the MCID's.

326

ACCEPTED MANUSCRIPT

327 **References**

328

- 329 1. Terwee CB, Roorda LD, Dekker J, et al. Mind the MIC: large variation among
330 populations and methods. *Journal of clinical epidemiology*. May 2010;63(5):524-534.
- 331 2. Beaton DE, Wright JG, Katz JN. Development of the QuickDASH: comparison of three
332 item-reduction approaches. *The Journal of bone and joint surgery. American volume*.
333 May 2005;87(5):1038-1046.
- 334 3. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths
335 and weaknesses and considerations for design. *The Journal of manual &*
336 *manipulative therapy*. 2009;17(3):163-170.
- 337 4. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome
338 measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The
339 Upper Extremity Collaborative Group (UECG). *American journal of industrial*
340 *medicine*. Jun 1996;29(6):602-608.
- 341 5. Bot AG, Becker SJ, van Dijk CN, Ring D, Vranceanu AM. Abbreviated Psychologic
342 Questionnaires Are Valid in Patients With Hand Conditions. *Clinical orthopaedics*
343 *and related research*. Aug 3 2013.
- 344 6. Roy JS, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic
345 review of four questionnaires. *Arthritis and rheumatism*. May 15 2009;61(5):623-
346 632.
- 347 7. Polson K, Reid D, McNair PJ, Larmer P. Responsiveness, minimal importance
348 difference and minimal detectable change scores of the shortened disability arm
349 shoulder hand (QuickDASH) questionnaire. *Manual therapy*. Aug 2010;15(4):404-
350 407.
- 351 8. Kennedy CA, Beaton DE, Solway S, McConnell S, Bombardier C. *The DASH and*
352 *QuickDASH Outcome Measure User's Manual*. 3rd ed. Toronto, Ontario, Canada:
353 Institute for Work and Health; 2011.
- 354 9. Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened
355 disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and
356 Numeric Pain Rating Scale in patients with shoulder pain. *Journal of shoulder and*
357 *elbow surgery / American Shoulder and Elbow Surgeons ... [et al.]*. Nov-Dec
358 2009;18(6):920-926.
- 359 10. Sorensen AA, Howard D, Tan WH, Ketchersid J, Calfee RP. Minimal clinically
360 important differences of 3 patient-rated outcomes instruments. *The Journal of hand*
361 *surgery*. Apr 2013;38(4):641-649.
- 362 11. Franchignoni F, Vercelli S, Giordano A, Sartorio F, Bravini E, Ferriero G. Minimal
363 Clinically Important Difference of the Disabilities of the Arm, Shoulder and Hand
364 Outcome Measure (DASH) and Its Shortened Version (QuickDASH). *The Journal of*
365 *orthopaedic and sports physical therapy*. Jan 2014;44(1):30-39.
- 366 12. Van Vliet MM, Maradey JA, Homa KA, Kerrigan CL. The usefulness of patient-
367 reported measures for clinical practice. *Plastic and reconstructive surgery*. Jul
368 2013;132(1):105-112.
- 369 13. Franchignoni F, Ferriero G, Giordano A, Sartorio F, Vercelli S, Brigatti E.
370 Psychometric properties of QuickDASH – A classical test theory and Rasch analysis
371 study. *Manual therapy*. 2011;16(2):177-182.

- 372 14. Angst F, Schwyzer HK, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult
373 shoulder function: Disabilities of the Arm, Shoulder, and Hand Questionnaire
374 (DASH) and its short version (QuickDASH), Shoulder Pain and Disability Index
375 (SPADI), American Shoulder and Elbow Surgeons (ASES) Society standardized
376 shoulder assessment form, Constant (Murley) Score (CS), Simple Shoulder Test
377 (SST), Oxford Shoulder Score (OSS), Shoulder Disability Questionnaire (SDQ), and
378 Western Ontario Shoulder Instability Index (WOSI). *Arthritis care & research*. Nov
379 2011;63 Suppl 11:S174-188.
- 380 15. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the
381 whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the
382 Arm, Shoulder and Hand outcome measure in different regions of the upper
383 extremity. *Journal of hand therapy : official journal of the American Society of Hand
384 Therapists*. Apr-Jun 2001;14(2):128-146.
- 385 16. Beaton DE, van Eerd D, Smith P, et al. Minimal change is sensitive, less specific to
386 recovery: a diagnostic testing approach to interpretability. *Journal of clinical
387 epidemiology*. May 2011;64(5):487-496.
- 388 17. Kennedy CA, Beaton DE, Smith P, et al. Measurement properties of the QuickDASH
389 (Disabilities of the Arm, Shoulder and Hand) outcome measure and cross-cultural
390 adaptations of the QuickDASH: a systematic review. *Quality of life research : an
391 international journal of quality of life aspects of treatment, care and rehabilitation*.
392 Mar 12 2013.
- 393 18. de Vet HC, Terluin B, Knol DL, et al. Three ways to quantify uncertainty in
394 individually applied "minimally important change" values. *Journal of clinical
395 epidemiology*. Jan 2010;63(1):37-45.
- 396 19. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal
397 changes in health status questionnaires: distinction between minimally detectable
398 change and minimally important change. *Health and quality of life outcomes*.
399 2006;4:54-54.
- 400 20. Turner D, Schunemann HJ, Griffith LE, et al. The minimal detectable change cannot
401 reliably replace the minimal important difference. *Journal of clinical epidemiology*.
402 Jan 2010;63(1):28-36.
- 403 21. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining
404 responsiveness and minimally important differences for patient-reported outcomes.
405 *Journal of clinical epidemiology*. Feb 2008;61(2):102-109.
- 406 22. Stratford PW, Riddle DL. When Minimal Detectable Change Exceeds a Diagnostic
407 Test-Based Threshold Change Value for an Outcome Measure: Resolving the Conflict.
408 *Physical therapy*. 2012;92(10):1338-1347.
- 409 23. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the
410 minimal clinically important difference. *Controlled clinical trials*. Dec
411 1989;10(4):407-415.
- 412 24. Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective
413 computation of responsiveness to change: the lesson of Cronbach. *Journal of clinical
414 epidemiology*. Aug 1997;50(8):869-879.
- 415 25. Gabel CP, Michener LA, Melloh M, Burkett B. Modification of the upper limb
416 functional index to a three-point response improves clinimetric properties. *Journal*

- 417 *of hand therapy : official journal of the American Society of Hand Therapists*. Jan-Mar
418 2010;23(1):41-51; quiz 52.
- 419 26. Wu A, Edgar DW, Wood FM. The QuickDASH is an appropriate tool for measuring
420 the quality of recovery after upper limb burn injury. *Burns : journal of the*
421 *International Society for Burn Injuries*. Nov 2007;33(7):843-849.
- 422 27. Wright AA, Cook CE, Baxter GD, Dockerty JD, Abbott JH. A comparison of 3
423 methodological approaches to defining major clinically important improvement of 4
424 performance measures in patients with hip osteoarthritis. *The Journal of orthopaedic*
425 *and sports physical therapy*. May 2011;41(5):319-327.
- 426 28. Turner D, Schunemann HJ, Griffith LE, et al. Using the entire cohort in the receiver
427 operating characteristic analysis maximizes precision of the minimal important
428 difference. *Journal of clinical epidemiology*. Apr 2009;62(4):374-379.
- 429 29. Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: a closer look at
430 the minimal clinically important difference (MCID). *The Journal of manual &*
431 *manipulative therapy*. Aug 2012;20(3):160-166.
- 432 30. Dawson J, Doll H, Boller I, et al. Comparative responsiveness and minimal change for
433 the Oxford Elbow Score following surgery. *Quality of life research : an international*
434 *journal of quality of life aspects of treatment, care and rehabilitation*. Dec
435 2008;17(10):1257-1267.
- 436 31. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons
437 Standardized Shoulder Assessment Form, patient self-report section: reliability,
438 validity, and responsiveness. *Journal of shoulder and elbow surgery / American*
439 *Shoulder and Elbow Surgeons ... [et al.]*. Nov-Dec 2002;11(6):587-594.
- 440 32. Bruynesteyn K, van der Heijde D, Boers M, et al. Minimal clinically important
441 difference in radiological progression of joint damage over 1 year in rheumatoid
442 arthritis: preliminary results of a validation study with clinical experts. *The Journal*
443 *of rheumatology*. Apr 2001;28(4):904-910.
- 444 33. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important
445 difference (MCID): a literature review and directions for future research. *Current*
446 *opinion in rheumatology*. Mar 2002;14(2):109-114.
- 447 34. Goldhahn J, Beaton D, Ladd A, Macdermid J, Hoang-Kim A. Recommendation for
448 measuring clinical outcome in distal radius fractures: a core set of domains for
449 standardized reporting in clinical practice and research. *Archives of orthopaedic and*
450 *trauma surgery*. 2014;134(2):197-205.
- 451 35. O'Brien L, Presnell S. Patient experience of distraction splinting for complex finger
452 fracture dislocations. *Journal of hand therapy : official journal of the American Society*
453 *of Hand Therapists*. Jul-Sep 2010;23(3):249-249; quiz 260.
- 454 36. Chung KC, Haas A. Relationship between patient satisfaction and objective
455 functional outcome after surgical treatment for distal radius fractures. *Journal of*
456 *hand therapy : official journal of the American Society of Hand Therapists*. Oct-Dec
457 2009;22(4):302-307; quiz 308.
- 458 37. Wang YC, Hart DL, Stratford PW, Mioduski JE. Baseline dependency of minimal
459 clinically important improvement. *Physical therapy*. May 2011;91(5):675-688.
460
461

Table 1

Baseline Statistics for improved patients and the not improved (scores represent means and standard deviations unless otherwise indicated)

Descriptor	Surgical			Nonsurgical			Carpal		
	Radius	Distal Fracture	<i>P</i>	Epicondylitis	Lateral	<i>P</i>	Tunnel	Release	<i>P</i>
	IP (n = 114)	NP (n = 37)		IP (n = 69)	NP (n = 68)		IP (n = 84)	NP (n = 34)	
Age	56(14.1)	52(12.6)	.16 ^a	47(9.2)	46(8.0)	.47 ^c	53(12.0)	53(11.9)	.93 ^c
Gender, (% male)	31(27%)	9(24%)	.73 ^b	35(51%)	31(46%)	.55 ^b	23(27%)	16(47%)	.04 ^b
Initial QDASH	60(19.8)	70(22.3)	.006 ^c	39(17.8)	43(19.8)	.23 ^c	56(23.6)	55(22.7)	.93 ^c
Duration of treatment (days of care)	35(12.3)	35(13.4)	.73 ^a	41(12.6)	38(10.1)	.22 ^a	26(10.0)	26(10.1)	.77 ^a

IP: Improved Patients; NP: Not-improved Patients; *P*: Significance

a: Wilcoxon (Mann Whitney-U); b: Chi-square tests; c: *t-test*

QDASH: The Quick Disabilities of The Arm Shoulder and Hand

Table 2
Scores of the QDASH and GROC

Descriptor	Surgical Distal Radius Fracture	Nonsurgical Lateral Epicondylitis	Surgical Carpal Tunnel Syndrome
Initial QDASH	63 ± 20.7	41 ± 18.8	56 ± 23.3
Last visit QDASH	29 ± 20.5	24 ± 15.6	30 ± 17.6
Last visit GROC	3.4 ± 2.0	3.4 ± 2.1	4.8 ± 1.7
Cutoff treatment sessions	12	12	8
Duration of treatment, d*	35 ± 13(21-97)	39 ± 11(24-92)	25 ± 9(14-56)

d*: Days of care, values are mean ± *SD* (range).

QDASH: The Quick Disabilities of The Arm Shoulder and Hand.

GROC: Global Rate of Change Scale.

Table 3

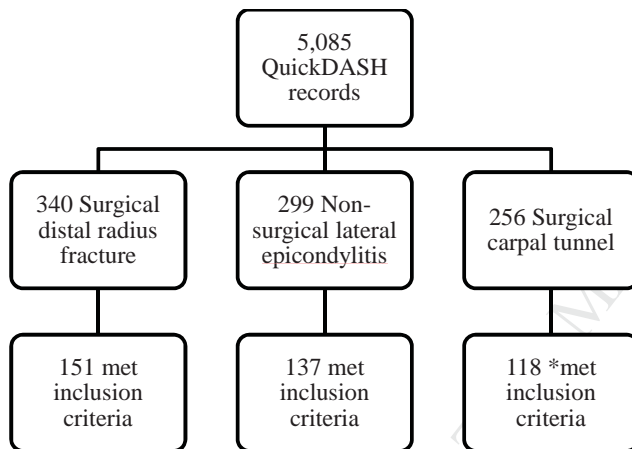
Mean score changes for the QDASH questionnaire according to each GROC scale grade

	Surgical Distal Radius Fracture		Nonsurgical Lateral Epicondylitis		Surgical Carpal Tunnel Syndrome	
	n(%)	QDASH	n(%)	QDASH	n(%)	QDASH
0 or less	4(3%)	9.7	11(8%)	2.6	7(6%)	15.9
+1 to +3	33(22%)	25.8	57(42%)	15.3	27(23%)	18.7
+4 to +5	54(36%)	29.6	52(38%)	17.6	57(48%)	26.6
+6 to +7	60(39%)	44.3	17(12%)	33.5	27(23%)	34.3

QDASH: The Quick Disabilities of The Arm Shoulder and Hand.

GROC: Global Rate of Change Scale.

Figure 1
Flow of charts meeting inclusion criteria



Inclusion criteria: have values for QDASH initial, visits 4,8 and 12, and GROG visit 12.
* = last visit for QDASH and GROG is visit 8 instead of 12.

Figure 2
QDASH Area Under The Curve (AUC) for surgical distal radius fracture.

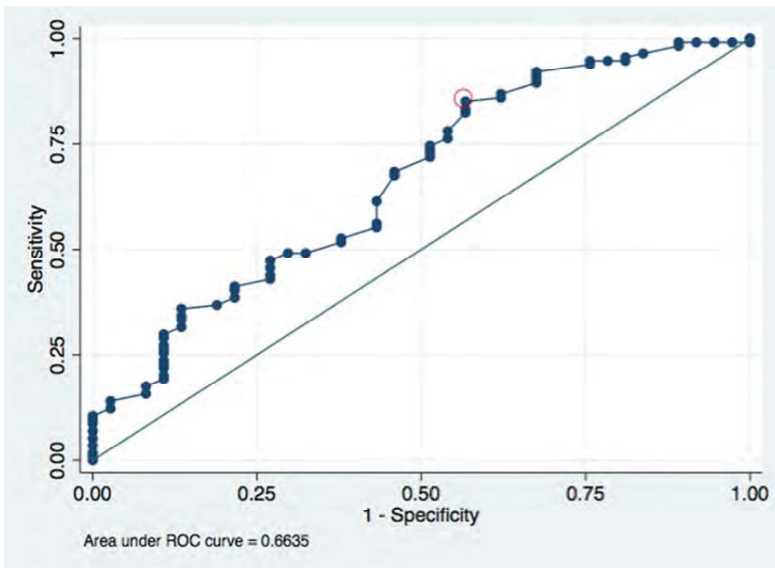


Figure 3
QDASH Area Under The Curve (AUC) for nonsurgical lateral epicondylitis.

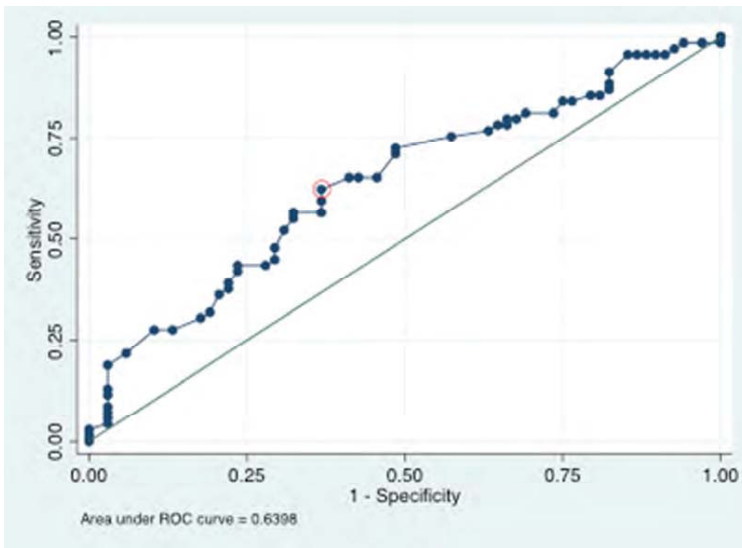
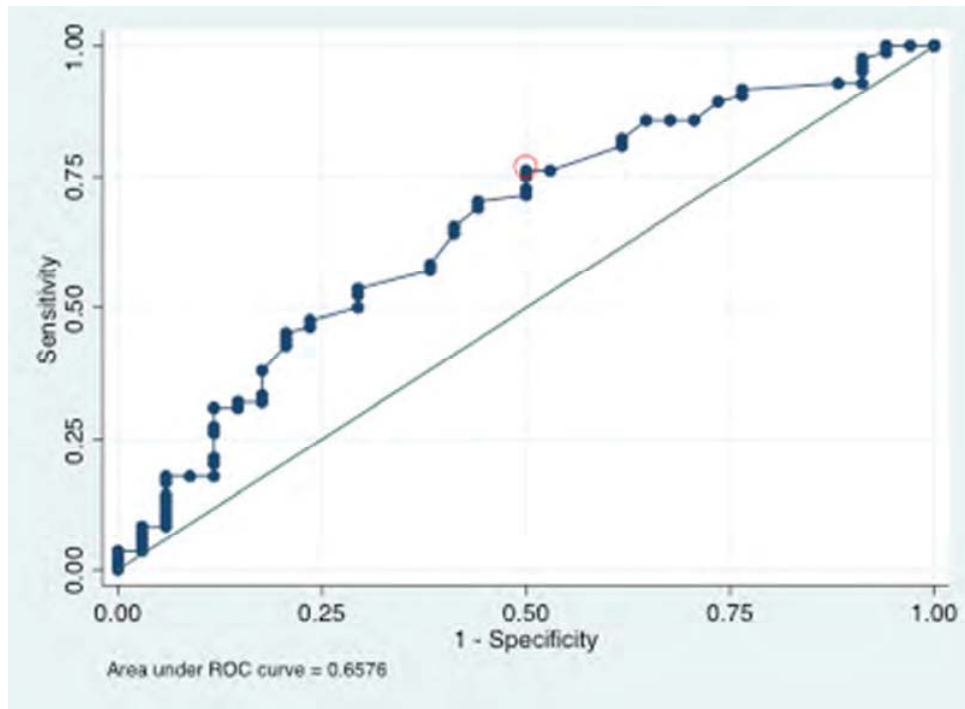


Figure 4
QDASH Area Under The Curve (AUC) for surgical carpal tunnel syndrome.



Highlights

- 406 participants, three diagnoses were assessed using triangulation methods
- For surgical distal radius fracture the MCID=25.8, MDC₉₀=25.28, and AUC=.66
- For non-surgical lateral epicondylitis the MCID=15.8, MDC₉₀=22.49, and AUC=.64
- For carpal tunnel release the MCID=18.7, MDC₉₀=27.63, and AUC=.66