The Usefulness of Patient Reported Measures for Clinical Practice.

Michael M. Van Vliet, MD

Johann A. Maradey

Karen A. Homa, PhD

Carolyn L. Kerrigan, MD

All authors affiliated with Dartmouth-Hitchcock Medical Center

Corresponding Author:

Michael M. Van Vliet, MD

Email Address: mikemvanvliet@gmail.com

Mailing Address: Re: Carolyn Kerrigan; Division of Plastic Surgery; 1 Medical Center Drive; Lebanon, NH 03756

Background: To assess the diagnostic potential of commonly used patient reported measures (PRMs), namely the Boston Carpal Tunnel Questionnaire (BCTQ function and BCTQ symptom severity), quickDASH, and SF-8.

Methods: PRM scores were retrospectively extracted from the records of 262 patients (397 hands) who visited the senior surgeon. Scores were compared using analysis of variance to determine statistical differences among diagnoses assigned by the same surgeon at the time of visit. Based on results of the post-hoc analysis, patients were grouped into one of two diagnostic groups: those with Dupuytren's disease and those with COT conditions (Carpal tunnel, Osteoarthritis, and Tenosynovitis). Logistic regression analysis determined whether higher scores were associated with COT conditions. A receiver operating characteristic curve was used to determine the most accurate PRM score for predicting a diagnostic group.

Results: Analysis of variance showed statistical differences among the five diagnoses for each PRM. Post-hoc analysis showed that Dupuytren's disease was significantly different from the other diagnoses. Logistic regression showed that COT conditions were statistically associated with higher BCTQ function, BCTQ symptom severity, and quickDASH scores compared to Dupuytren's. Lower physical and mental summary SF-8 scores were associated with the COT conditions. quickDASH scores ≥25, BCTQ symptom severity scores ≥2.5, and BCTQ function scores ≥2 are the best PRM threshold values for distinguishing between the two diagnostic groups.

Conclusion: The quickDASH and BCTQ PRMs have diagnostic potential. Establishing threshold values for predicting a diagnostic group may prove to be a useful tool for referring providers.

Level of Evidence III. Diagnostic

Introduction

The common complaint of "hand pain" and "loss of function" carries a large differential diagnosis. The hand surgeon and community providers are often called upon to guide treatment to the millions of people who seek relief from these common complaints every year. To add to the complexity, studies that have evaluated the diagnostic accuracy of both basic physical exam findings and nerve conduction velocity (NCV) modalities have produced contrasting results [1-5]. With the increased focus on cost-effective and cost-conscious care, identifying reliable and cost efficient diagnostic tools is paramount. The goal of this paper is to unveil the diagnostic potential of patient reported measures (PRM), which have become a vital part of assessments used in clinical trials and health services research.

Patient reported measures have become an integrated step in the care of patients at our institution. At the time of initial consultation, patients are invited to complete two commonly used PRMs, the quickDASH, and the SF-8. The Boston Carpal Tunnel Questionnaire is then administered to patients who either achieve a certain score on the quickDASH or to patients referred with a diagnosis of carpal tunnel syndrome. Prior to entering the exam room, the surgeon reviews the results of the PRMs. A hand surgeon at our institution has observed certain patterns in PRM scores among patients with similar diagnoses. In fact, she was recently consulting on a patient referred with the diagnosis of Dupuytren's disease. However, on the basis of elevated PRM scores, the surgeon knew that this diagnosis was unlikely the main disorder before even examining or meeting the patient.

On the basis of anecdotal observations, we hypothesize that patients' scores on the BCTQ functional and symptom severity measures and the quickDASH will differ amongst various hand diagnoses. In addition, we predict no difference in SF-8 health survey scores between the various diagnoses given the generic nature of this survey. By identifying a statistical difference between BCTQ and quickDASH values for certain conditions, we hope to identify a threshold which will be predictive of a certain diagnostic group of common hand complaints. The ultimate goal is to define the PRM as a diagnostic tool that will add predictive power in identifying the correct diagnosis.

Methods

Background

The BCTQ was reported in 1993 by Levine et. al. and comprises two scales, a functional status scale and a symptom severity scale [6]. The functional status scale has 8 items rated from 1 (no difficulty with activity) to 5 (cannot perform the activity at all). The symptom severity scale has 11 items rated on a scale of 1 through 5, 1 the mildest and 5 the most severe. The final scores for BCTQ are reported as an average of the ratings. The BCTQ is highly reliable, reproducible, and has been validated for assessing patient reported carpal tunnel outcomes [6-9]. It has limited reported use for comparison of non carpal tunnel ailments.

The DASH is a region specific scale of 30 items, self-administered, to assess the physical and social components of health related to the upper extremity. The DASH score is reported from 1 (no disability) to 100 [8]. A large number of literature reports have proven the validity, reliability, and responsiveness of the DASH for proximal and distal arm disorders [8-16]. In this study, we utilized the quickDASH [17], which has 11 items with the same scoring range. While studies have shown that the full-length DASH provides more specific and accurate results [18], numerous reports also indicate that the quickDASH can be used instead of the DASH with similar precision in upper extremity disorders [16, 19-

21]. Additional studies have shown the independent responsiveness, validity and reliability of the quickDASH [22-26] for various upper extremity disorders.

The SF-8 is an eight item, practical measurement of eight health domains that provides both physical and mental component summary scores. The scale is calibrated such that fifty is the norm-based average with a standard deviation of ten. The generic survey can be applied to patients across all ages, diseases, and treatment groups. It obtains information about the functional health and well being from a patient's point of view[27].

Subjects

After approval by the IRB committee, a query was created to extract from Dartmouth Hitchcock's data warehouse all patients (total 262) who had visited the senior surgeon in the past 5 years with ICD-9 codes of 354.0 (carpal tunnel), 715.14 (osteoarthritis), 727.03 (trigger finger), 727.05 (tenosynovitis of the hand), and 728.6 (Dupuytren's contracture). The query was built to return the following information for each patient: age, medical record number, date of visit, CPT codes in a given time frame, other recorded patient diagnoses, and results of the BCTQ, quickDASH, and SF-8. Patients were not included if PRM scores were missing from the record. All electronic records were manually reviewed to ensure accuracy of the query and to extract handedness, affected hand, and NCV motor latencies when applicable.

All recorded diagnoses were based on the surgeon's opinion after review of the history, physical exam, nerve conduction velocity (NCV) distal motor latency (in the case of suspected carpal tunnel) and PRM scores.

Statistical Analysis

Analysis of variance was used to compare PRM scores among all diagnoses. Post-hoc multiple-comparison test (Bonferroni) was used to determine which diagnoses had significantly different PRM scores. This established two diagnostic groups: Dupuytren's disease and COT conditions (C for Carpal tunnel, O for Osteoarthritis, and T for Tenosynovitis/Trigger finger). Logistic regression was used to determine whether the two diagnostic groups were associated with the PRMs, in which COT conditions were assigned 1 and Dupuytren's disease was assigned 0. Both a crude model and a model adjusted for age and SF-8 physical summary score to eliminate potential confounding factors were done, in which there were no major differences between models and the adjusted results were reported. A p-value less than 0.05 was considered significant. To determine the ability of BCTQ function and symptom severity and the quickDASH to predict a diagnosis, various cut-off scores were analyzed using the receiver operating characteristics (ROC). The best cut-off score was chosen by which one maximized the area under the curve and by examining specificity, sensitivity and accuracy values. The software used for statistical analysis was STAT 8.0 (College Station, TX)

RESULTS

A total of 397 hands from 262 patients were included in the study. Of the 397 hands included in the study, 169 were found to have carpal tunnel syndrome, 118 had Dupuytren's contracture, 10 had tenosynovitis, 7 had osteoarthritis, and 93 had multiple diagnoses, at least one of which was an inflammatory condition. The average age of the patients was 59.6 years (standard deviation of 13.4 years) and 89% or 232 of the patients were right-handed. Fifty-seven percent of the hand complaints from patients were for the right hand. Not all patients completed the 3 different PRM questionnaires. Figure 1 is a venn diagram that depicts the number of PRM scores available in the database in which some patients completed more than one PRM. Patients in our institution are given PRMs according to a

branching logic. All patients are given SF and quickDASH. The BCTQ is given only for carpal tunnel referrals and if a certain score was achieved on the quickDASH PRM.

Average PRM scores for the five different diagnostic groups originally extracted from Dartmouth-Hitchcock patient database are displayed in Table 1. The results of the analysis of variance for each of the PRMs showed that there were significant differences among the groups (all p-values <0.001). Table 2 lists the p-values for the Bonferroni post-hoc analysis of the PRM scores among all diagnosis, except the multiple diagnoses group. Dupuytren's disease was significantly different from the other diagnoses for most PRM scores. The multiple diagnosis group, not included in Table 2 due to its limited clinical applicability, was significantly different from Dupuytren's only.

On the basis of the post-hoc analysis (Table 2), patients were divided into two diagnosis groups: COT disorders and Dupuytren's disease. Table 3 shows the results of the logistic regression, in which higher BCTQ function, BCTQ symptom severity, and quickDASH scores were associated with COT conditions compared to Dupuytren's disease. Lower SF-8 mental and physical summary scores were also associated with COT conditions. Figures 2 through 4 are histograms showing the distributions of three of the PRMs for the two diagnosis groups.

Cut-off scores for the PRMs were chosen and assessed to determine the best score for predicting a COT disorder. Table 4 lists sensitivity, specificity and the AUC of a range of scores for the given PRM. Threshold values with the greatest AUC are highlighted. For example, a score greater or equal to 2 for BCTQ function was the best cut-off score to predict a COT disorder.

DISCUSSION

PRMs have widespread applicability in research and in determining effectiveness and satisfaction in individual health care. These measures, according to the data in this study, may have additional diagnostic potential. While some studies have shown gross distinctions between PRM scores for various diagnoses [16, 28], no studies to our knowledge have focused on this important observation as the focus of the study.

Other authors, however, have suggested utilizing PRMs for more than measuring patient centered outcomes. Jester et al. report that the DASH has potential in the development of patient-centered treatment programs which are tailored to the individual patients' requirements. The authors retrospectively grouped surgically treated patients and compared DASH scores among them. The treatment groups included over ten surgical modalities including burn operations, ray amputations, fusions, nerve releases, and arthroplasties. The DASH scores differentiated well among the groups, adding further evidence that patients with different upper extremity ailments score differently on this PRM. The authors did not, however, report predictive value, sensitivity, or specificity of a threshold number on the PRM which could be used to guide a clinical diagnosis [28].

This study suggests potential diagnostic ability of the BCTQ and quickDASH. It is the first report using the PRMs as a diagnostic tool with reported sensitivities and specificities. PRMs are a potential cost effective means of guiding referrals for primary care physicians. The ultimate goal is to define the PRM as another tool to add further evidence to a presumed diagnosis. If a provider encounters a quickDASH score below 25 and BCTQ scores below 2, s/he could begin to develop a differential before the initial patient visit. As with the referral patient mentioned in the introduction, knowledge of the PRM guided the referring physician toward a different diagnosis than the initial referral diagnosis. The patient's PRM scores available prior to the consultation visit were: BCTQ function 2.88, BCTQ symptom severity 3.27, and quickDASH 54.5. On the basis of the outcomes of this study, the senior author would have

known, without an examination, that Dupuytren's disease, the referring diagnosis, was low on the differential as the sole cause of the patient's symptoms. In fact, the patient was experiencing carpal tunnel syndrome (in addition to mild Dupuytren's) and his symptoms were relieved following carpal tunnel release.

A limitation of the study is that the COT conditions were grouped and compared to Dupuytren's disease with ordinal data sets, thus increasing the likelihood of a type 1 error. However, we are confident that because the p value is so low, the possibility of differences or associations is most likely remote. In addition, while the data were prospectively collected, the study design and analysis were done retrospectively. The data was recorded by the patients at the time of their visit thus eliminating any potential recall bias and increasing the reliability of the study. Another limitation is that the PRMs were used to make the original diagnosis, biasing the results toward each different diagnosis. Additionally, the senior author has significant experience with using PRMs to aid in a diagnosis. This limits the generalizability to other practices.

Although the use of PRMs as a diagnostic tool is an exciting concept, future work is necessary before the measures can alone predict a specific diagnosis. While a specific PRM score was not correlated with each studied diagnosis, we were able to show trends in scores for certain ailments. These trends can be used by the provider to narrow the differential and to generate a discussion with the patient about where they fall in relation to other patients who take the same PRM. The diagnostic potential of PRMs may become more evident if each component of the PRM (i.e. pain) is studied among the groups. In the future, it would be helpful for multiple surgeons with various degrees of experience to prospectively look at differences in each component of the PRMs and assign a diagnosis on the basis of each component. It will also be imperative to avoid utilizing the PRM under investigation to aid in the final diagnosis. Doing so may show the ability of PRMs to help with treatments and diagnoses.

REFERENCES

1   Descatha A, Dale AM, Franzblau A, Coomes J, Evanoff B. Diagnostic strategies using physical examination are minimally useful in defining carpal tunnel syndrome in population-based research studies. Occup Environ Med 2009; 67: 133-135.

2   Mondelli M, Passero S, Giannini F. Provocative tests in different stages of carpal tunnel syndrome. Clin Neurol Neurosurg 2001; 103: 178-183.

3   Katz JN, Larson MG, Sabra A, Krarup C, Stirrat CR, Sethi R et al. The carpal tunnel syndrome: diagnostic utility of the history and physical examination findings. Ann Intern Med 1990; 112: 321-327.

4   Ansari NN, Adelmanesh F, Naghdi S, Mousavi S. The relationship between symptoms, clinical tests and nerve conduction study findings in carpal tunnel syndrome. Electromyogr Clin Neurophysiol 2009; 49: 53-57.

5   Kuhlman KA, Hennessey WJ. Sensitivity and specificity of carpal tunnel syndrome signs. Am J Phys Med Rehabil 1997; 76: 451-457.

6   Levine DW, Simmons BP, Koris MJ, Daltroy LH, Hohl GG, Fossel AH et al. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. J Bone Joint Surg 1993; 75A: 1585-1592.

7   Ollivere BJ, Logan K, Ellahee N, Miller-Jones JC, Wood M, Nairn DS. Severity scoring in carpal tunnel syndrome helps predict the value of conservative therapy. J Hand Surg 2009; 34B: 511-515.

8   Sambandam SN, Priyanka P, Gul A, Ilango B. Critical analysis of outcome measures used in the assessment of carpal tunnel syndrome. Int Orthop 2008; 32: 497-504.

9   Gay RE, Amadio PC, Johnson JC. Comparative responsiveness of the disabilities of the arm, shoulder, and hand, the carpal tunnel questionnaire, and the SF-36 to clinical change after carpal tunnel release. J Hand Surg 2003; 28A: 250-254.

10   Dias JJ, Rajan RA, Thompson JR. Which questionnaire is best? The reliability, validity and ease of use of the Patient Evaluation Measure, the Disabilities of the Arm, Shoulder and Hand and the Michigan Hand Outcome Measure. J Hand Surg 2008; 33B: 9-17.

11   Greenslade JR, Mehta RL, Belward P, Warwick DJ. Dash and Boston questionnaire assessment of carpal tunnel syndrome outcome: what is the responsiveness of an outcome questionnaire? J Hand Surg 2004; 29B: 159-164.

12   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; 14: 128-146.

13   SooHoo NF, McDonald AP, Seiler JG, 3rd, McGillivary GR. Evaluation of the construct validity of the DASH questionnaire by correlation to the SF-36. J Hand Surg 2002; 27A: 537-541.

14   Bot SD, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. Ann Rheum Dis 2004; 63: 335-341.

15   Roy JS, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. Arthritis Rheum 2009; 61: 623-632.

16   Gummesson C, Ward MM, Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. BMC Musculoskelet Disord 2006; 7: 44.

17   Beaton DE, Wright JG, Katz JN. Development of the QuickDASH: comparison of three item-reduction approaches. J Bone Joint Surg 2005; 87A: 1038-1046.

18   Angst F, Goldhahn J, Drerup S, Flury M, Schwyzer HK, Simmen BR. How sharp is the short QuickDASH? A refined content and validity analysis of the short form of the disabilities of the shoulder, arm and hand questionnaire in the strata of symptoms and function and specific joint conditions. Qual Life Res 2009; 18: 1043-1051.

19   Gabel CP, Yelland M, Melloh M, Burkett B. A modified QuickDASH-9 provides a valid outcome instrument for upper limb function. BMC Musculoskelet Disord 2009; 10: 161.

20   Mehta S, Macdermid JC, Carlesso LC, McPhee C. Concurrent validation of the DASH and the QuickDASH in comparison to neck-specific scales in patients with neck pain. Spine (Phila Pa 1976) 2010; 35: 2150-2156.

21   Niekel MC, Lindenhovius AL, Watson JB, Vranceanu AM, Ring D. Correlation of DASH and QuickDASH with measures of psychological distress. J Hand Surg 2009; 34A: 1499-1505.

22   Fan ZJ, Smith CK, Silverstein BA. Responsiveness of the QuickDASH and SF-12 in Workers with Neck or Upper Extremity Musculoskeletal Disorders: One-Year Follow-Up. J Occup Rehabil 2011; 21: 234-243.

23   Polson K, Reid D, McNair PJ, Larmer P. Responsiveness, minimal importance difference and minimal detectable change scores of the shortened disability arm shoulder hand (QuickDASH) questionnaire. Man Ther 2011; 15: 404-407.

24   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; 18: 920-926.

25   Wu A, Edgar DW, Wood FM. The QuickDASH is an appropriate tool for measuring the quality of recovery after upper limb burn injury. Burns 2007; 33: 843-849.

26   Budd HR, Larson D, Chojnowski A, Shepstone L. The QuickDASH score: a patient-reported outcome measure for Dupuytren's surgery. J Hand Ther 2011; 24: 15-20; quiz 21.

27   SF Health Surveys. http://wwwqualitymetriccom/WhatWeDo/GenericHealthSurveys/tabid/184/Defaultaspx?gclid=CK305-Oit6wCFcjb4Aod9D_DHg 2011.

28   Jester A, Harth A, Wind G, Germann G, Sauerbier M. Disabilities of the arm, shoulder and hand (DASH) questionnaire: Determining functional activity profiles in patients with upper extremity disorders. J Hand Surg 2005; 30B: 23-28.

Figure Legend

Figure 1. The number of hands assessed by the indicated PRM.

Figure 2.   Histogram showing distribution of quickDASH scores (binned in increments of 5 on the x axis) for the two diagnostic groups. Percentage of patients displayed on the y axis.

Figure 3. Histogram showing distribution of BCTQ symptom severity scores (binned in increments of 0.5 on the x axis) for the two diagnosis groups. Percentage of patients displayed on the y axis.

Figure 4. Histogram showing distribution of BCTQ function scores (binned in increments of 0.5 on the x axis) for the two diagnosis groups. Percentage of patients displayed on the y axis.

Table 1. Average PRM score for the indicated diagnostic group

Diagnostic Group	BCTQ Function Average (SD)	BCTQ Symptom Average (SD)	qDASH (SD)	SF-8 Mental (SD)	SF-8 Physical (SD)
Carpal Tunnel	2.3 (0.7)	2.9 (0.7)	40.2 (18.7)	53.1 (10.0)	42.8 (10.4)
Tenosynovitis/ Trigger Finger	2.6 (1.0)	2.9 (0.5)	40.9(24.5)	59.8 (5.6)	36.0 (12.9)
Osteoarthritis	2.7 (1.0)	2.6 (1.1)	58.6 (25.6)	53.9 (9.2)	43.0 (10.3)
Dupuytren's	1.6 (0.5)	2.0 (0.6)	11.3 (10.0)	56.9 (5.0)	50.0 (7.9)
Multiple Dx	2.3 (1.0)	2.7 (0.8)	37.4 (18.5)	53.3 (9.3)	43.0 (10.4)
ANOVA p value	< 0.0001	< 0.0001	< 0.0001	0.0007	< 0.0001

. SD = Standard deviation. ANOVA = analysis of variance.

Table 2.   Post-hoc analysis comparing PRM scores for each diagnosis to another.

	Tenosynovitis	Osteoarthritis	Dupuytren's Disease
Osteoarthritis	BCTQ fxn = 1.00 BCTQ Sx = 1.00 qDASH = 0.789 SF-8 Mental = 1.00 SF-8 Physical = 1.00		
Dupuytren's Disease	BCTQ fxn = 0.014 BCTQ Sx = 0.019 qDASH = <0.001 SF-8 Mental = 1.00 SF-8 Physical = <0.001	BCTQ fxn = 0.014 BCTQ Sx = 0.581 qDASH = <0.001 SF-8 Mental = 1.00 SF-8 Physical = 0.624	
Carpal Tunnel	BCTQ fxn = 1.00 BCTQ Sx = 1.00 qDASH = 1.00 SF-8 Mental = 0.219 SF-8 Physical = 0.434	BCTQ fxn = 1.00 BCTQ Sx = 1.00 qDASH = 1.00 SF-8 Mental = 0.219 SF-8 Physical = 0.434	BCTQ fxn = <0.001 BCTQ Sx = <0.001 qDASH = <0.001 SF-8 Mental = 0.002 SF-8 Physical = <0.001

Patients with multiple diagnoses were eliminated from this analysis.   Numbers indicate p-values for comparison of the indicated PRMs. Highlighted numbers indicate statistical significance.

Table 3. Logistic Regression Analysis comparing PRM scores between Duputytren's Disease and

Inflammatory Conditions

	OR	95% CI	P value	# Hands Inflammatory	# Hands with Dupuytrens
BCTQ Function	5.04	2.42 to 10.51	< 0.001	185	42
BCTQ Symptom	7.48	2.98 to 18.78	< 0.001	166	23
quickDASH	1.18	1.10 to 1.25	< 0.001	177	39
SF-8 Mental	0.94	0.90 to 0.98	0.008	175	117
SF-8 Physical	0.90	0.87 to 0.93	< 0.001	175	117

Odds ratios (OR), 95% confidence intervals (CI), and p values for indicated PRM score comparisons

between the Dupuytren's and inflammatory cohorts. Also listed is the number of hands in each cohort for
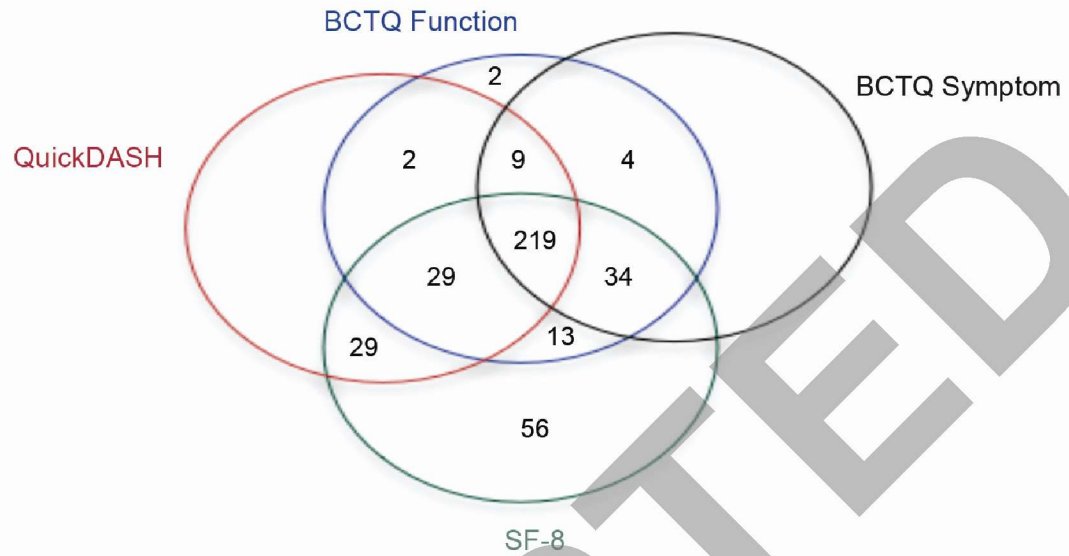
each PRM comparison.

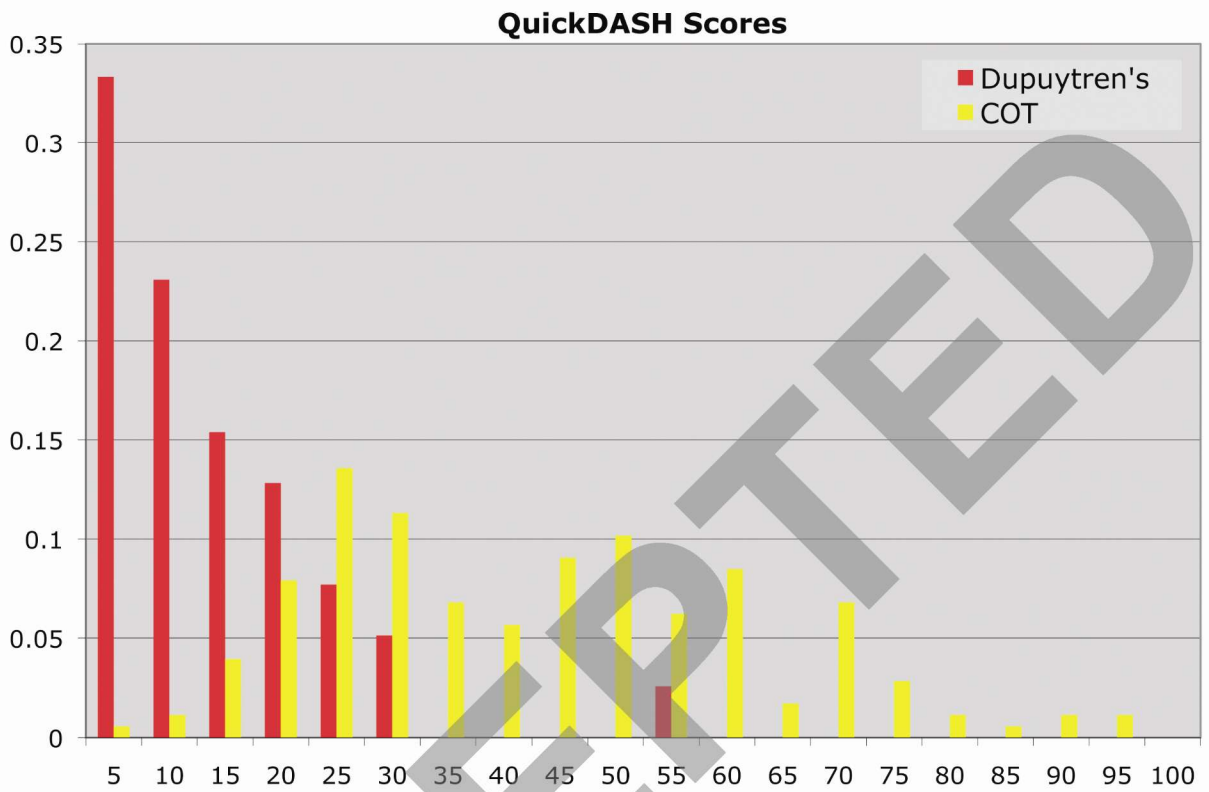Table 4. Sensitivity, Specificity and the AUC Ranges for Given PRM Score Range

	Sensitivity	Specificity	Area Under Curve
BCTQ Function PRM			
Score ≥ 1.75	77.8%	61.9%	0.699
Score ≥ 2.00	67.6%	78.6%	0.731
Score ≥ 2.25	42.2%	90.5%	0.663
BCTQ Symptom PRM			
Score ≥ 2.00	89.2%	52.2%	0.707
Score ≥ 2.25	81.3%	73.9%	0.776
Score ≥ 2.50	69.9%	91.3%	0.806
DASH PRM			
Score ≥ 20	86.4%	84.6%	0.855
Score ≥ 25	82.5%	92.3%	0.874

Highlighted values indicate PRM score with greatest AUC. Percents represent sensitivity or specificity

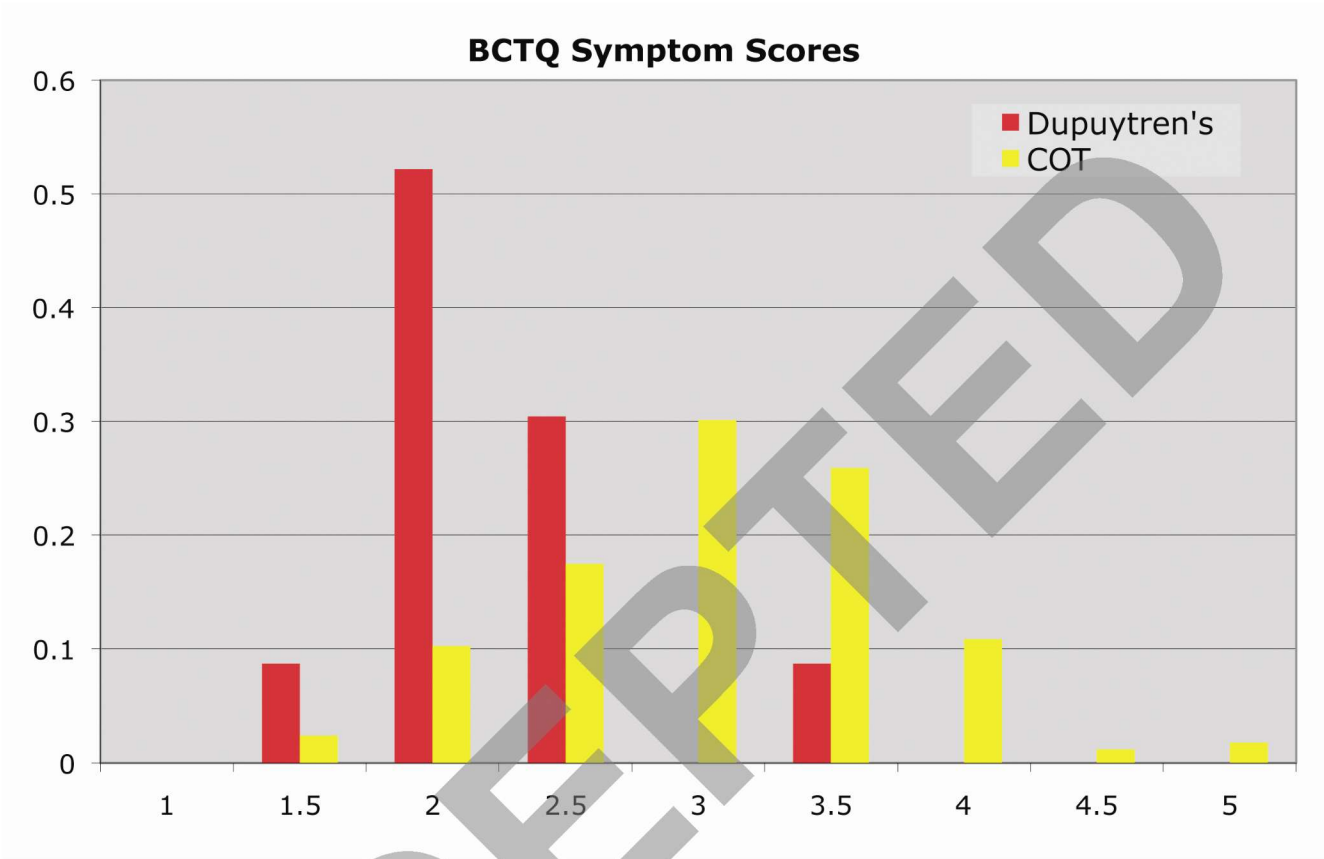for predicting and inflammatory condition.
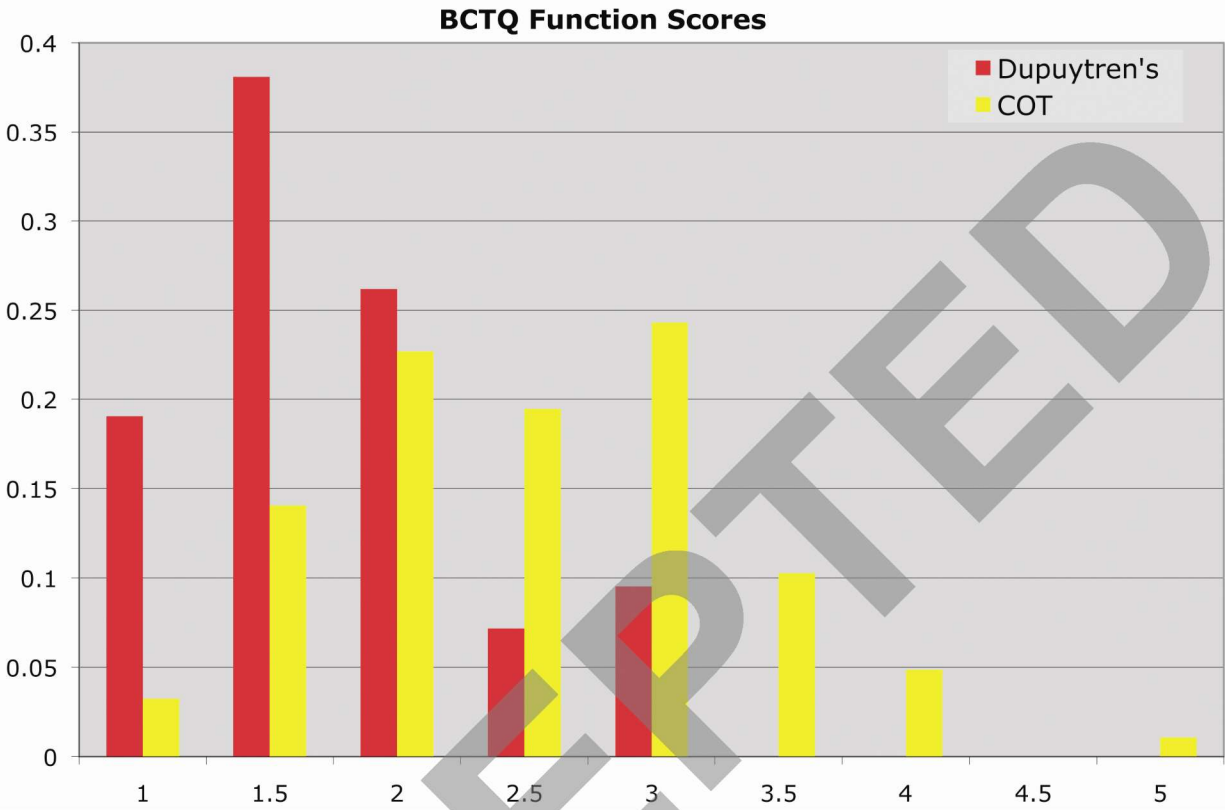
Total Hands Surveyed: 397

QuickDASH Scores

BCTQ Symptom Scores

BCTQ Function Scores