Published in final edited form as:

J Appl Biobehav Res. 2011 December 1; 16(3-4): 148–166. doi:10.1111/j.1751-9861.2011.00068.x.

USE OF THE MINIMAL CLINICALLY IMPORTANT DIFFERENCE (MCID) FOR EVALUATING TREATMENT OUTCOMES WITH TMJMD PATIENTS: A PRELIMINARY STUDY¹

Megan Ingram, MS, Yun Hee Choi [Ph.D. Candidate], Chung-Yi Chiu, Ph.D., CRC, Rob Haggard, MS, Angela Liegey Dougall, Ph.D., Peter Buschang, DDS, and Robert J. Gatchel, Ph.D., ABPP²

Abstract

Temporomandibular joint and muscle disorder (TMJMD) is one of the most prevalent types of musculoskeletal disorders. The major goal of the study was to more objectively quantify clinically meaningful relief for TMJMD treatment outcomes by using the new metric of minimal clinically important difference (MCID). Pre- to post-treatment changes on a number of self-report measures were evaluated in a cohort of 101 acute TMJMD patients. An anchor-based MCID approach was employed, with an objective chewing performance measure serving as the clinical outcome of interest. Using a Receiver Operating Curve analysis, it was found that the Physical Component Scale (PCS) of the SF-36 was the most robust self-report measure to use as the MCID in a TMJMD patient population.

The American Academy of Orofacial Pain (American Academy of Orofacial Pain, 2004) has estimated that 75% of the U.S. population experiences symptoms of temporomandibular joint and muscle disorder (TMJMD), 5–10% of whom require professional pain management procedures and treatment. Moreover, TMJMD ranks as one of the highest commonly occurring musculoskeletal conditions resulting in pain and disability (second only to chronic low back pain), and it is the most common cause of facial pain (National Institute of Dental and Craniofacial Research, 2008). The NIDCR (2008) also estimated TMJMDs cost an average of \$4 billion annually. Even though TMJMD is both a common and a costly disorder, it has traditionally been difficult to objectively and reliably measure the perceived effectiveness of its treatment. Although there has been an increase in the research of pain and chronic pain conditions such as TMJMD, clinical research in the field has been limited to self-reported descriptions by patients. Research in the field of neuroscience has assessed the basic neural and biochemical mechanisms involved and, with the development of the biopsychosocial model of pain, new approaches to the management of pain are being made. As noted by Gatchel, Peng, Peters, Fuchs, and Turk (2007):

“each individual experiences pain uniquely, and a range of psychological and socioeconomic factors can interact with physical pathology to modulate a patient’s report of symptoms and subsequent disability.” (p.607)

This subjectivity in measures of pain has stimulated clinical researchers to develop “interpretive guidance from a concrete value that might indicate a clinically significant/important outcome” (Gatchel & Mayer, 2010, p. 322). It is important to not only develop better pain management methods, but also to develop reliable measures that objectively

¹This research was supported by Grant 5U01DE010713014 from the National Institute of Dental and Craniofacial Research.

²Address correspondence to Dr. Robert J. Gatchel, Department of Psychology, College of Science, The University of Texas at Arlington, Arlington, TX. 76019-0528. gatchel@uta.edu.

document treatment efficacy. Objective measures that both qualitatively describe and quantitatively measure pain symptoms are vital for understanding how patients experience pain, and how much relief may be expected by applying different treatments.

Pain is often measured clinically by evaluating treatment responsiveness over time, such as a response to a therapy protocol or a surgical procedure. Beaton, Bombardier, and colleagues (Beaton et al., 2001) have argued that evaluations of responsiveness are designed to assess how a target of interest (such as joint pain and disability) has changed over time. Similar to pain in other conditions (e.g., spinal pain disorders), TMJMD pain can be assessed across three major measurement domains: self-report, overt behavior (e.g., activities of daily living, maintaining responsibilities at work, and social relationships), and physical indices (measures of physical and functional performance; Gatchel, Mayer, & Chou, in press). Self-report measures have traditionally been used in health-care research for a number of reasons. First, patients are often the best and sometimes only source of information when evaluating phenomena such as pain. Therefore, the information may only be obtained by directly asking them. Second, objective measurements may not always match the patient's evaluation. As a consequence, the concept of a minimal clinically important difference (MCID) has been recently introduced to better document treatment outcome. MCID is defined as the smallest change or difference that patients perceive as clinically beneficial (Jaeschke, Singer, & Guyatt, 1989). The MCID is intended to provide outcome measures that are more clinically meaningful than measurements based simply on mean improvement on some outcome. Indeed, a basic concept behind the MCID is that statistically significant differences in measures do not necessarily reflect clinically meaningful benefits. This relatively new metric has been the recent focus of many studies, especially in the orthopedic literature, using a standardized MCID to define outcomes for patients who have had "successful" or "unsuccessful" surgical treatment (Gatchel et al., in press).

On a global level, an early recommendation of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials consensus review was that a 30% change in self-reported pain or disability be used as a general barometer of positive clinical change (Dworkin et al., 2005, 2008; Haythornthwaite, 2010). However, this was merely a consensus decision based on a general review of numerous studies that used different treatment modalities and statistical methods. Subsequently, a number of other recommendations have been made, using both anchor-based and distribution-based MCID approaches (Gatchel & Mayer, 2010). A distribution-based approach focuses on the statistical characteristics of a patient sample, and then compares an observed change to an index of variability in order to determine whether the change is substantial and clinically meaningful. Such an approach has been utilized for a variety of measures of statistical variability, including the standard error of measurement, the standard error of mean change, the standard deviation of change, and the standard deviation of the sample group. The effect size is a standardized statistical measure that compares a score change, after treatment, relative to a pretreatment standard deviation. Guidelines for the interpretation of effect size are somewhat arbitrary, although criterion proposed by Cohen (1977) have been widely accepted and used in clinical research.

In contrast to the above approach, the anchor-based approach requires the use of an independent, objective criterion to use in order to determine a threshold value for the MCID. This can be determined in one of two ways. The first is by use of a mean change score (the mean change of patients who improved and, therefore, one can set its cutoffs on the basis of the change score of patients who were determined to have had a small, moderate, or large change). The second is by use of a Receiver Operator Curve (ROC) method, where the ROC determines a cutoff score for dichotomizing a particular cohort into patients who improved and patients who did not improve on the same external criterion or anchor. The reader is referred to Copay and colleagues (Copay, Subach, Glassman, Polly, & Schuler, 2007) for a

more comprehensive overview of these two approaches. Gatchel, Lurie, and Mayer (2010) have also discussed these two approaches.

The purpose of the present study was to evaluate the clinical utility of the MCID for use with TMJMD patient outcomes in a randomized clinical trial of a biobehavioral treatment program, using an external anchor measure of chewing performance. This is the first study that has applied the new MCID metric to a TMJMD population, and also the first to apply MCID to the objective measure of chewing performance.

METHODS

Participants

The current study consisted of 101 patients with acute TMJMD who were referred by private dental practices and community dental clinics located throughout the Dallas-Fort Worth Metroplex. These various dental clinics collaborated with the Acute TMJMD Treatment Program at the University of Texas Arlington, in Arlington, Texas and the Texas A&M Health Science Center/Baylor College of Dentistry in Dallas, Texas. Participants were also recruited via word-of-mouth, the Internet, public flyers, and a single radio advertising campaign. Individuals qualified to participate in the program if their symptoms of TMJMD, based on the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD; Dworkin & LeResche, 1992), had been present for approximately 6 months or less. Some participants were considered “high risk” on the basis of earlier clinical research revealing that individuals experiencing a certain level of symptom severity were at a significantly higher risk for developing a chronic TMJMD condition if they did not participate in an early intervention program for their condition (Epker, Gatchel, & Ellis, 1999; Wright et al., 2004). Participants who had the same symptom timeframe of approximately less than 6 months, but whose symptoms were not as severe (“low risk” patients) were monitored while they participated in a standard care program. Although the majority of participants were Caucasian women, both men and women were considered eligible if they were over 18 years of age ($M = 45$ years, $SD = 16$ years), and met the criteria just described. Table 1 presents the demographic variables for this entire sample. These initial criteria involved: (1) participant report of the presence of at least one of the three cardinal signs of TMJMD: jaw pain, limited range of motion in the jaw, or noise in the temporomandibular joint; (2) feelings of stiffness, tightness, or fatigue in the jaw; and (3) had these symptoms for less than 6 months at the time of entry into the study. Participants were excluded if they had significant comorbid physical conditions; cancer; low-back pain; or fibromyalgia (that may have had exacerbated the participant’s pain symptoms); or have had a history of jaw pain before the most recent episode. Individuals who did not meet the RDC/TMD diagnostic criteria were also excluded. Dentists and research associates/clinicians evaluated each participant to determine eligibility. High risk versus low risk status of participants was determined by an algorithm initially developed by Epker et al. (1999).

Associate clinical researchers were Masters level clinicians, formally educated in the fields of Clinical Social Work or Counseling. The clinicians performing the examinations were trained by both a clinical psychologist and a licensed counselor experienced with the RDC/TMD. They observed each rater reliably conduct a physical exam on a nonsubject volunteer prior to the beginning of the study in order to ensure that correct evaluation methods were followed. The clinicians were given a detailed training video in addition to their supervised training. “Recalibration” sessions were regularly held to make certain that there was continued interrater reliability, as well as quality control through the random selection and reevaluation of selected cases. In an effort to ensure quality and protocol adherence, the clinicians were provided weekly supervision by a licensed psychologist through the review of session recordings.

Procedure

Participants filled out a general information form and the history form before scheduling a series of pre-intervention biopsychosocial (BPS) evaluations. The pre-intervention BPS evaluations were completed preferably within 1 week, and they included physical, psychosocial, and quantitative functional measures. Trained clinicians administered the RDC/TMD, which consisted of the “at risk” algorithm in order to identify participants who were either at “high” or “low” risk for developing chronic problems (Epker et al., 1999; Wright et al., 2004). The algorithm used to assess risk is composed of Question 3 from the RDC History Questionnaire, the Characteristic Pain Intensity (CPI), and the assessment of oral facial pain from muscle palpation on Items 1, 8, and 10 of the Oral Facial exam. The Functional Evaluation of Chewing Performance was administered as a physical measure, while the psychosocial measures of this study included the Graded Chronic Pain Scale (GCPS); the CPI; and the SF-36 Health Survey; Symptom Checklist; Headache Questionnaire, Orthodontic History Questionnaire; and Treatment Cost data. Of specific interest to this current study were the CPI, GCPS, SF-36, and Chewing Performance.

These acute TMJMD participants were randomly assigned to one of two treatment groups based on the results of their initial pre-intervention evaluations and screening: biobehavioral treatment or self-care treatment. Overall, this created three separate groups: high risk, biobehavioral treatment; high-risk, self-care treatment; or low risk, nonintervention. All three groups were continuously matched through randomization for age, gender, race, and time-since-original-onset of TMD. Dentists collaborating on this study were kept blind to the group assignments. The basic demographic characteristics of each of these groups are presented in Table 2.

Biobehavioral Treatment

Those “high risk” participants assigned to the biobehavioral group ($n = 28$) were provided 6 intervention sessions, that consisted of individual meetings with a trained clinician who adhered to a standardized treatment protocol. Session 1 focused on the clinician providing the participant with an overview and rationale for biobehavioral treatment. At the conclusion of the first session, the participant was educated on diaphragmatic breathing techniques for both the purpose of relaxation and pain management. Session 2 involved more relaxation training, specifically progressive muscle relaxation. In Session 3, the participant was introduced to the idea of using relaxation skills in everyday situations. Biofeedback was also introduced to reinforce the acquisition of skills. During Session 4, biofeedback was integrated, and the clinician assisted the participant in learning distraction methods and activity scheduling. The participant received an explanation of the rationale for cognitive interventions in Session 5. Participants were then taught how to identify destructive, automatic thoughts and correct them during the fifth session. In the final session (Session 6), the clinician reviewed with the participant what skills he or she had learned in the past five meetings. Discussions on how to maintain therapeutic gains, and a plan for coping with reoccurrences or pain flares in the future, also took place. Between each session, participants were required to complete assigned homework and handouts. Participants received a comprehensive workbook that included an overview of what was taught in each session, and daily logs that provided space for the participants to record their pain, stress level, frequency of stress, pain triggers, amount of sleep, and coping mechanisms. The sessions were typically conducted within 3-to-12-week time period, depending upon participant scheduling ability.

Self-Care Treatment

Those “high risk” participants who were assigned to the self-care group ($n = 25$) were also administered 6 intervention sessions, the difference being that these participants did not

receive any training on coping techniques like progressive muscle relaxation or cognitive interventions. Rather, participants in this group received readings over the course of their treatment that were geared toward educating the patient about TMJMD, self-care activities, medications, nutrition, treatment options, and patient–physician communication. Clinicians then reviewed the major points of the readings in each session, and requested feedback on the participant’s reactions to them. Participants were required to fill out a daily log recording their pain/discomfort, stress, and tension.

Non-Intervention

The “low risk” patients were assigned to a nonintervention group ($n = 48$) that received the standard-of-care that would normally have been offered to them and that they accepted. Any provider with whom the participant consulted would offer this care. The participant was required to document/log all care pertaining to their jaw pain.

For all participants, immediately following treatment, they were readministered a postintervention BPS evaluation that was identical to the one preintervention BPS evaluation.

Measures

Patient Information Form—The Patient Information Form gathered data such as demographics, education, contact information, employment status, workers compensation, personal injury litigation, history of jaw pain (including onset, date of treatment, and type of treatment), and chronic health conditions.

RDC/TMD History Questionnaire and Evaluation—The major outcome measures, such as pain, disability, and limitations in mandibular function, were derived from scores on a History Questionnaire and TMJMD Examination. The examination form was based on Laskin’s criteria of pain and tenderness in the mastication muscles and the TMJ with limited mandibular movement (Laskin, 1969). The Examination Form documented numerous items including the opening pattern, presence of joint sounds (clicking or popping) at opening or closing, additional measurements of clicks, vertical range of motion, and distance of protrusions (Gardea, Gatchel, & Mishra, 2001).

Both the TMJMD Examination Form and the History Questionnaire used the standard RDC/TMD (Dworkin & LeResche, 1992; Schiffmann et al., 2010). The RDC/TMD is a multiaxial system. Axis I assesses physiological diagnoses across three groups: (1) Group I assesses the presence or absence of muscle disorders, such as myofascial pain, myofascial pain with limited opening, or no diagnosis; (2) Group II assesses displacements of discs within the joint which can include disc displacement with reduction, without reduction limited opening, without reduction or limited opening, and no diagnosis; and (3) Group III assesses miscellaneous joint conditions like arthralgia, osteoarthritis of the TMJ, osteoarthritis of the TMJ, or no diagnosis. The RDC/TMD also provided a “rule-out” against systematic arthritic disease and acute traumatic injury. Axis II assessed psychosocial status and pain-related disability issues. The combination of assessing both physical and psychosocial components of TMD is consistent with the biopsychosocial health model (Schiffmann et al., 2010). The History Questionnaire assessed psychosocial aspects of TMJMD, which included pain intensity, disability, depression, and nonspecific physical symptoms and limitations to mandibular functioning.

The Characteristic Pain Inventory (CPI)—The CPI (Dworkin & LeResche, 1992) is a self-report measure that quantifies aspects of pain over the past 3 months. The RDC/TMD Examination’s Questions 7, 8, 9 composed the Characteristic Pain Inventory. Pain severity

was measured with a range from 0 to 100 (100 being the most intense pain). The score was then calculated by taking the mean of current pain, worst pain, and average pain and multiplying them by 10.

Graded Chronic Pain Scale (GCPS)—The Graded Chronic Pain Scale is derived from Axis II of the RDC/TMD. Pain intensity, interferences or changes in activities (work related, family, or leisure), and disability days due to pain were measured on the Graded Chronic Pain Scale (GCPS) which was derived from Questions 10, 11, 12, and 13 on the RDC Examination. Scores ranged from 0 (no change) to 100 (extreme change). Scoring rules categorize pain severity in the following categories: Grade I, low intensity with little pain-related impediment; Grade II, high intensity pain associated with little pain-related impediment; Grade III, high pain intensity and pain-related disability; and Grade IV, with severely limiting pain and high disability.

The SF-36—The Short Form-36 Health Survey (SF-36) was designed by the Medical Outcomes Study, and is a multipurpose health survey that is used to assess quality of life in relation to health status as viewed by the health-care recipient (McHorney, Ware, & Raczek, 1993; Stewart & Ware, 1992). The SF-36 was constructed as a representation of both multidimensional health concepts and the full range of health states (McHorney et al., 1993). The SF-36 results in an 8-scale profile of functional health and well-being, in addition to psychometrically based physical and mental health summary measures—Mental Component Scale (MCS) and the Physical Component Scale (PCS). The MCS and PCS were the scales used in the present study. Normative data were based on various medical populations, making this measure useful for comparative purposes. Lower scores indicate a greater degree of disability. Scores range from 0 to 100, with a mean of 50 and a standard deviation of 10. The SF-36 is considered to have high test-retest reliability coefficients with a Chronbach's alpha exceeding .70, usually above .80 (Ware, Snow, Kosinski, & Gandek, 1993).

Chewing Performance—Chewing performance is the most commonly used measure to assess masticatory performance (Bates, Stafford, & Harrison, 1976; Buschang, 2006). Chewing performance was used as the external physical measure for subsequent use in determining an MCID value. This measure of physical function was composed of an evaluation of median particle size, broadness of distribution, and the participant's self-rating of pain. Participants were asked to chew a standardized, tasteless tablet (5 mm thick and 20 mm in diameter) of new, softer CutterSil® for a total of 20 chews. A total of 20 chews was selected because the majority of studies have estimated masticatory performance using a total of 20 chewing cycles. (Akeel, Nilner, & Nilner, 1992; Buschang, 2006; Durante Gayiã, Raymundo, & Sobrinho, 2001; English, Buschang, & Throckmorton, 2002; Henrikson, Ekberg, & Nilner, 1998; Owens, Buschang, Throckmorton, Palmer, & English, 2002; Shiere & Manly, 1952). The new CutterSil is a condensation silicon impression material that is formed using a Plexiglass template. After hardening for an hour, the tablets are cut into quarters. Five portions containing three quarter-tablets each were then packaged for each participant (Buschang, Throckmorton, Travers, & Johnson, 1997). The time for each of the five trials were recorded. Once the chewed samples were obtained from the participants, they were air dried in filter papers over a stainless steel colander.

For purposes of quantification, the samples were separated using a series of sieves with mesh sizes of 5.6 mm, 4.0 mm, 2.8 mm, 2.0 mm, .85 mm, .425 mm, and .25 mm, which were stacked on a mechanical stacker and vibrated for 2 minutes. Once separated, the samples were weighed to the nearest .01 gm and the cumulative weight percentages (defined by the amount of the sample that can pass through each successive sieve) were calculated for each sample. This allows for the determination of median particle size and broadness of

particle distribution using Rosin–Rammler equation (Oltoff, Van der Bilt, Bosman, & Kleizen, 1984). During the chewing test, pain was also rated on a scale from 1 (no pain) to 10 (pain as bad as it could be). In addition, patients were asked to state which side of their mouth they chewed on, and which felt most comfortable while they were chewing. Then, the assessment was performed a second time, the only difference being that the participant chewed on the opposite side of the mouth than they did in the first round. The change scores were calculated as the difference between the initial and the follow-up scores to determine the patient's functional impairment, followed by classifying that data as “unchanged” (with posttreatment scores falling outside of .5 SD) or “improved” (with posttreatment scores falling inside .5 SD). The cutoff criteria of .5 SD corresponded to the MCID across several studies according to Norman, Sloan, and Wyrwich (2003), and was equivalent to 1 SEM for a reliability of .75 (Copay et al., 2007). It should be noted that a review conducted by Miller (1956) found the limit of human discrimination to be equivalent to an effect size between .36 and .63.

Statistical Analyses

The change scores were calculated for each self-report measure (the CPI, GCPS, and the PCS and MCS of the SF-36). All clients were classified as improved or not improved based on the same criteria as one-half SD of chewing performance of posttreatment scores of broadness of distribution (using broadness of distribution as the functional measure), and were classified dichotomously to conduct ROC analyses as “unchanged/slightly improved” or “unchanged/improved.” The ROC curve provides MCID scores on self-report measures. Any differences in a score greater than the minimal change detected were considered as substantially clinically beneficial. ROC curves are limited to a two-class classification: “unchanged” or “improved.” The asymptotic values address the significance level and should be less than .05. However, due to a sample size limitation of the current study, asymptotic values corresponding to the 90% confidence interval at .1 were also considered. The area under the ROC provides the probability of correctly discriminating between improved and nonimproved patients, ranging from .5 to 1, with .5 representing the ability to discriminate by chance, and 1 being the ability to perfectly classify all the patients (Gatchel & Mayer, 2010). The area under the curve of .7 to .8 was considered acceptable (Hosmer & Lemeshow, 2000). It should be pointed out that the values of the “area under the curve” varied, and they need to be interpreted in conjunction with the asymptotic values. The emphasis in clinical practice on accurately identifying individuals with a condition was evaluated through sensitivity (true positives) and specificity (true negatives). Intervention outcomes are evaluated with an emphasis on accurately identifying patients who truly improve as the result of an intervention, as well as correctly identifying those who do not. So, the current study evaluated MCID values of the PCS, MCS, CPI, and GCPS that would have, at least, the highest sensitivities. Also, the higher possibility of accurately detecting a positive outcome is needed. Therefore, the bigger area under the ROC indicates that measure's accuracy.

RESULTS

Of the 101 participants with chewing performance data, only 98 had complete data for the self-report measures at both pre- and posttreatment (the CPI, GCPS, and the PCS and MCS of SF-36). The anchor-based approach was applied to each of these four self-report measures for all the participants (using broadness of distribution as the functional measure), and .5 SD as the cutoff. To further understand the meaningful change scores calculated on the PCS, MCS, CPI, and GCPS, the participants were separated out by intervention group: either High Risk/Biobehavioral Group or High Risk/Self-Care Group. Table 3 presents the descriptive analysis for all four self-report measures for all participants. In an attempt to

further evaluate the MCID value calculated for the overall sample, the sample was separated into the two assigned treatment groups.

All Participants

For the results calculated using a .5 SD from the mean (in accordance with the review by Copay et al., 2007) for the particle distribution index of the chewing performance test, the areas under the curve (the degree of accuracy) were as follows: PCS = .641; MCS = .562; CPI = .432; and GCPS = .453. The associated change scores were: 2.745 (Sensitivity = .805) for the PCS; 1.460 (Sensitivity = .805) for the MCS; 30.000 (Sensitivity = .805) for the CPI; and 23.333 (Sensitivity = .818–.857) for the GCPS. Thus, overall, the PCS was found to be the better measure of MCID, with the best asymptotic value of .058 and corresponding ROC area of .641 when using the broadness of distribution index of chewing performance. In addition, the PCS correctly identified participants with a sensitivity of .805. Although its specificity was only .526, that was much higher than those of the other measures.

Intervention: High Risk/Biobehavioral Group

Again, using the .5 SD cutoff for particle distribution, the ROC analysis reported areas under that curve as follows: PCS = .740; MCS = .479; CPI = .422; and GCPS = .438. The associated change scores were as follows: 2.115 (Sensitivity = .833) for the PCS; 5.665 (Sensitivity = .833) for the MCS; 40.000 (Sensitivity = .833) for the CPI; 43.333 (Sensitivity = .833) for the GCPS. Overall, the PCS again continued to provide more clinically meaningful change scores, with an asymptotic value of .131 and an ROC area of .740 when using broadness of distribution. The sensitivity of the PCS also remained high (.833).

Intervention: High Risk/Self-Care Group

For broadness of distribution index of chewing performance, within the High Risk/Self-Care Group, the areas under the curve were PCS = .656; MCS = .767; CPI = .589; and GCPS = .494. The associated change scores were 1.825 (Sensitivity = .833) for the PCS; 1.245 (Sensitivity = .833) for the MCS; 41.667 (Sensitivity = .833) for the CPI; and 23.333 (Sensitivity = .833) for the GCPs. Once again, for this group, the PCS continued to provide a more robust MCID metric, with an asymptotic value of .297 and corresponding ROC area of .656. The PCS, as demonstrated with all participants and the High Risk/Biobehavioral Group, correctly identified affected participants with a sensitivity of .833. It should also be noted that the MCS provided a good MCID metric for this group, with an asymptotic value of .074 and ROC area of .767.

DISCUSSION

Determining clinically important changes continues to be difficult to evaluate due to the variation in approaches, and one standard metric has not yet emerged without some controversy. Past MCID approaches have been heavily criticized for violating psychometric properties when comparing two self-report measures (Gatchel & Mayer, 2010). Distribution-based approaches frequently fail to demonstrate clinical importance, and assume that the degree of change will hold consistent across ranges of scores. Anchor-based approaches are limited by how strong the objective criteria are on which they are based. Because a patient's perception of their health can greatly influence treatment protocols and self-evaluation, there has been increased emphasis on obtaining an objective value that will correspond to a patient's perception of improvement during treatment. The present study represents the first attempt to evaluate the clinical utility of the MCID metric with TMJMD using an objective measure of chewing performance. The results were quite promising. After reviewing the data for anchor-based MCID values of TMJMD, one objective measure of chewing performance (broadness of particle distribution) provided strong evidence of meaningful

change. Better asymptotic values and corresponding ROC areas under the curve when using this index provided change scores that could be interpreted as more clinically meaningful, with both the High Risk/Biobehavioral Group and the High Risk/Self-Care groups. This was consistent with previous research revealing that chewing performance for a standard number of cycles is one of the most powerful masticatory performance measures that can be used (Bates et al., 1976; Buschang, 2006).

The PCS of the SF-36 provided the strongest evidence of minimal clinically important differences in chewing performance. For all participants, this MCID value was 2.745. Thus, if patients' PCS scores changed from pretreatment to posttreatment by 2.745, then clinicians can expect that their chewing performance would be improved. The PCS resulted in more frequent asymptotic values closer to .1 (at the 90% confidence interval), and corresponding areas under the curve that were bigger in the ROC model testing, relative to the other three self-report measures. In addition, the PCS maintained higher sensitivity values (that corresponded to correctly identifying TMJMD patients who would improve in a clinical practice). It should be pointed out, though, that the specificity values were not as high as desired (the rate of accurately predicting patients whose TMJMD symptoms were unchanged). Therefore, further development of the chewing performance task anchors (e.g., not improved to normal, almost normal, and above normal) may be needed in future studies with larger sample sizes. Nevertheless, such results are quite promising in revealing an MCID value for future research designed to evaluate treatment effectiveness outcomes for TMJMD patients. It should also be noted that, although the PCS appears to be the best self-report measure to use, the less useful values reported for the CPI and GCPS could indicate that participants have a more difficult time providing a cognitive appraisal of their improvement, as reflected by these two measures. Again, both the PCS and MCS of the SF-36 are calculated based on 36 items. The CPI and GCPS have much fewer items that may significantly limit their utility in determining an objective, clinically meaningful change value.

Of course, in any preliminary clinical study of this type, there are often certain limitations that can be noted. One issue may have been the limited demographic composition of the sample, which was predominately Caucasian and female. However, the sample was representative of the clinic population from which it was recruited and was comparable to other large-scale studies supporting its use as a representative sample of clinical TMJMD patients (Hoffmann et al., 2011). Because of the relatively small sample size evaluated, the results should be viewed as preliminary in nature. Moreover, as stated earlier, the asymptotic values used in such studies should be less than .05. However, due to the sample size limitations, asymptotic values corresponding to the 90% confidence interval at .1 were considered. Nevertheless, future clinical research to determine objective and clinically meaningful changes should continue to emphasize the need for a carefully selected and an objective anchor such as was used in the present study (i.e., chewing performance). Moreover, in terms of subjective self-report measures, the psychometrics of the measure need to be considered or improved before applying. Consequently, researchers and clinicians should be aware of some potential limitations of an MCID approach. Due to the complexity of pain, both at the nociceptive level and the cognitive level, it may be difficult to objectively define meaningful change. For this reason, developing a consensus through the use of measures of several different domains of pain (psychological, physiological, and functional) is advised in order to reach more meaningful conclusions.

References

- Akeel R, Nilner M, Nilner K. Masticatory efficiency in individuals with natural dentitions. *Swedish Dental Journal*. 1992; 16:191–198. [PubMed: 1481126]

- American Academy of Orofacial Pain. TMD Symptoms. 2004. Retrieved December 16, 2004, 2004, from <http://www.aaop.org/index.asp?Type=SEC&SEC={1387ACCB-B805-48F6-BB7F-0031808FDEAB}&DE={79C1954A-AA74-44C9-A517-4FF98FB17183}>
- Bates JF, Stafford GD, Harrison A. Masticatory function -A review of the literature. *Journal of Oral Rehabilitation*. 1976; 3:57–67. [PubMed: 772184]
- Beaton DE, Bombardier C, Katz JN, Wright J, Wells GA, Boers M, Shea B. Looking for important change/differences in studies of responsiveness. *Journal of Rheumatology*. 2001; 28:400–405. [PubMed: 11246687]
- Buschang PH. Masticatory ability and performance: the effects of mutilated and maloccluded dentitions. *Seminars in Orthodontics*. 2006; 12(2):91–101.
- Buschang PH, Throckmorton GS, Travers KH, Johnson G. The effects of bolus size and chewing rate on masticatory performance with artificial test foods. *Journal of Oral Rehabilitation*. 1997; 24(7): 522–526. [PubMed: 9250840]
- Cohen, J. *Statistical Power for the Behavioral Sciences*. New York: Academic Press; 1977.
- Copay AG, Subach BR, Glassman SD, Polly DWJ, Schuler TC. Understanding clinically important difference: A review of concepts and methods. *The Spine Journal*. 2007; 7:541–546. [PubMed: 17448732]
- Durante Gayão MB, Raymundo VG, Sobrinho LC. Masticatory efficiency in children with primary dentition. *Pediatric Dental Journal*. 2001; 23:499–505.
- Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005; 113(1–2):9–19. [PubMed: 15621359]
- Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, Zavisic S. Interpreting the Clinical Importance of Treatment Outcomes in Chronic Pain Clinical Trials: IMMPACT Recommendations. *The Journal of Pain*. 2008; 9(2):105–121.10.1016/j.jpain.2007.09.005 [PubMed: 18055266]
- Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders. *Journal of Craniomandibular Disorders: Facial & Oral Pain*. 1992a; 6:301–355.
- English JD, Buschang PH, Throckmorton GS. Does malocclusion affect masticatory performance? *The Angle Orthodontist*. 2002; 72:21–27. [PubMed: 11843269]
- Epker JT, Gatchel RJ, Ellis E. An accurate model for predicting TMD chronicity: Practical applications in clinical settings. *Journal of the American Dental Association*. 1999; 130:1470–1475. [PubMed: 10570591]
- Gardea MA, Gatchel RJ, Mishra KD. Long-term efficacy of biobehavioral treatment of temporomandibular disorders. *Journal of Behavioral Medicine*. 2001; 24:341–359. [PubMed: 11523332]
- Gatchel RJ, Lurie JD, Mayer TG. Minimal Clinically Important Difference. *SPINE*. 2010; 35(19): 1739–1743. [PubMed: 20700084]
- Gatchel RJ, Mayer TG. Testing minimal clinically important difference: Consensus or conundrum? *The Spine Journal*. 2010; 10:321–327. [PubMed: 20362248]
- Gatchel RJ, Mayer TG, Chou R. What Does/Should the Minimum Clinically Important Difference measure? A Re-consideration of its Clinical Value in Evaluations Efficacy of Lumbar Fusion Surgery. *Clinical Journal of Pain*. (in press).
- Gatchel RJ, Peng Y, Peters ML, Fuchs PN, Turk DC. The Biopsychosocial Approach to Chronic Pain: Scientific Advances and Future Directions. *Psychological Bulletin*. 2007; 133:581–624. [PubMed: 17592957]
- Haythornthwaite JA. IMMPACT recommendations for clinical trials: opportunities for the RDC/TMD. *Journal of Oral Rehabilitation*. 2010; 37(10):799–806.10.1111/j.1365-2842.2010.02152.x [PubMed: 20887278]
- Henrikson T, Ekberg E, Nilner M. Masticatory efficiency and ability in relation to occlusion and mandibular dysfunction in girls. *International Journal of Prosthodontics*. 1998; 11:125–132. [PubMed: 9709601]

- Hoffmann RG, Kotchen JM, Kotchen TA, Cowley T, Dasgupta M, Cowley AW. Temporomandibular disorders and associated clinical comorbidities. *The Clinical Journal of Pain*. 2011; 27:268–274. [PubMed: 21178593]
- Hosmer, DW.; Lemeshow, S. *Applied logistic regression*. New York: Wiley, Inc; 2000.
- Jaeschke R, Singer JA, Guyatt GH. Measurement of health status: Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*. 1989; 10:407–415. [PubMed: 2691207]
- Laskin DM. Etiology of the pain-dysfunction syndrome. *Journal of the American Dental Association*. 1969; 79:147–153. [PubMed: 5254545]
- McHorney C, Ware JJ, Raczek AB. The MOS 36 Item Short-Form Health Survey (SF-36): II. Psychometric and Clinical Tests of Validity in Measuring Physical and Mental Health Constructs. *Medical Care*. 1993; 31(3):247–263. [PubMed: 8450681]
- Miller GA. The magical number seven, plus or minus two:some limits on our capacity for processing information. *Psychological Review*. 1956; 63(2):81–97.10.1037/h0043158 [PubMed: 13310704]
- National Institute of Dental and Craniofacial Research. Facial Pain. 2008 Dec 20.2008 Retrieved September 13, 2009, from <http://www.nidcr.nih.gov/DataStatistics/FindDataByTopic/FacialPain/>.
- Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life. The remarkable universality of half a standard deviation. *Medical Care*. 2003; 41:582–592. [PubMed: 12719681]
- Oloff LW, Van der Bilt A, Bosman F, Kleizen HH. Distribution of particle sizes in food comminuted by human mastication. *Archives of Oral Biology*. 1984; 29:899–903. [PubMed: 6596036]
- Owens S, Buschang PH, Throckmorton GS, Palmer L, English JD. Masticatory performance and areas of occlusal contact and near contact in subjects with normal occlusion and malocclusion. *American Journal of Orthodontic and Dentofacial Orthopedics*. 2002; 121:602–609.
- Schiffmann E, Truelove E, Ohrbach R, Anderson G, John M, List T, Look J. The research diagnostic criteria for temporomandibular disorders I: overview and methodology for assessment and validity. *Journal of Orofacial Pain*. 2010; 24:7–24. [PubMed: 20213028]
- Shiere FR, Manly RS. Theeffect of the changing dentition on masticatory function. *Journal of Dental Research*. 1952; 31:526–534. [PubMed: 14946296]
- Stewart, AL.; Ware, JE. *Measuring functioning and well-being: The Medical Outcomes Study approach*. Durham, NC: Duke University Press; 1992.
- Ware, JE.; Snow, KK.; Kosinski, M.; Gandek, B. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston: The Health Institute, New England Medical Center; 1993.
- Wright AR, Gatchel RJ, Wildenstein L, Riggs R, Buschang P, Ellis E. Biopsychosocial differences in high-risk versus low-risk acute TMD pain-related patients. *Journal of the American Dental Association*. 2004; 135:474–483. [PubMed: 15127871]

Table 1

Demographic Information for All Participants (N=101)

	N	%
Gender		
Male	23	22.8
Female	78	77.2
Race/Ethnicity		
White	72	71.3
Black	13	12.9
Hispanics	0	0
Asian or Pacific Islander	4	4.0
Aleut, Eskimo, American Indian	3	3.0
Other	5	5.0
Highest Grade Completed		
High School (12 th)	10	9.9
Associates (13–14yrs)	22	21.8
Bachelors (15–16yrs)	33	32.6
Graduate (17yrs+)	35	34.7
Marital Status		
Married spouse in household	53	52.5
Widowed	1	1.0
Divorced/Separated	13	12.9
Never Married	33	32.7
Total Household Income		
\$0–\$24,999	22	21.8
\$25,000–\$49,999	17	16.8
\$50,000 or more	60	59.4

Table 2

Demographic Data by Treatment Group

	High Risk/Biobehavioral Group(N=48)		High Risk/Self Care Group (N=25)		Low Risk/Non Intervention Group(N=48)	
	N	%	N	%	N	%
Age (M, SD)		43 (15.78)		45(17.16)		45(17.16)
Gender						
Male	5	17.9	4	16.0	14	29.2
Female	23	82.1	21	84.0	34	70.8
Race/Ethnicity						
White	21	75.0	17	68.0	34	70.8
Black	2	7.1	3	12.0	8	16.7
Hispanics	0	0	0	0	0	0
Asian or Pacific Islander	1	3.6	2	8.0	1	2.1
Aleut, Eskimo, American Indian	1	3.6	0	0	2	4.2
Other	2	7.1	1	4.0	2	4.2
Highest Grade Completed						
High School (12)	1	3.6	3	12.0	6	12.5
Associates (13–14)	6	21.4	4	16.0	12	25.0
Bachelors (15–16)	7	25.0	9	36.0	17	35.4
Graduate (17+)	14	50.0	8	32.0	13	27.1
Marital Status						
Married spouse in household	17	60.7	13	52.0	23	47.9
Widowed	0	.0	0	.0	1	2.1
Divorced/Separated	3	10.7	5	20.0	5	10.4
Never Married	8	28.6	6	24.0	19	39.6
Total Household Income						
\$0–\$24,999	8	28.6	5	20.0	9	18.7
\$25,000–\$49,999	2	7.2	5	20.0	10	20.9
\$50,000 or more	18	64.3	14	56.0	28	58.3

Table 3
MCIDs, as well as means and standard deviations of all four measures (at pre- and post-treatment) for the two treatment groups. Both groups combined (all) also included.

	MCIDs											
	SF-36 PCS			SF-36 MCS			CPI			GCPS		
	All	Bio-behavioral Intervention	Self-care	All	Bio-behavioral Intervention	Self-care	All	Bio-behavioral Intervention	Self-care	All	Bio-behavioral Intervention	Self-care
MCID (<i>p</i> ^a)	2.745 (.058)	2.115 (.131)	1.825 (.297)	1.460 (.405)	5.665 (.896)	1.245 (.074)	30.000 (.363)	40.000 (.622)	41.667 (.551)	23.333 (.529)	43.333 (.694)	23.333 (.970)
Area ^b	.641	.740	.656	.562	.479	.767	.432	.422	.589	.453	.438	.494
MEANS, SDs and Δ SCORES												
	SF-36 PCS			SF-36 MCS			CPI			GCPS		
	All	Bio-behavioral Intervention	Self-care	All	Bio-behavioral Intervention	Self-care	All	Bio-behavioral Intervention	Self-care	All	Bio-behavioral Intervention	Self-care
Pre	49.122 (7.933)	48.363 (8.740)	45.253 (8.426)	46.309 (10.179)	45.052 (11.697)	46.648 (9.768)	49.524 (20.929)	64.881 (10.980)	62.133 (16.857)	25.544 (22.607)	37.143 (25.055)	28.533 (19.078)
Post	49.931 (8.173)	50.940 (8.925)	46.488 (9.223)	49.321 (10.776)	47.483 (12.291)	50.767 (9.785)	35.374 (23.632)	45.238 (21.876)	45.733 (17.573)	12.619 (16.877)	15.357 (19.842)	18.800 (18.656)
Δ pre-post	.809 (6.380)	2.577 (5.837)	1.236 (7.975)	3.012 (9.033)	2.431 (9.618)	4.119 (6.000)	-14.150 (20.375)	-19.643 (19.295)	-16.400 (23.032)	-12.925 (20.201)	-21.786 (22.924)	-9.733 (18.780)
<i>t</i> test (<i>p</i>)	1.255 (.212)	2.336 (.027)	.743 (.465)	3.301 (.001)	1.337 (.192)	3.293 (.003)	-6.875 (.000)	-5.387 (.000)	-3.560 (.002)	-6.334 (.000)	-5.029 (.000)	-2.591 (.016)
<i>d</i>	0.255	0.899	0.317	0.670	0.515	1.404	-1.396	-2.073	-1.453	-1.286	-1.936	-1.058

Note:

^a Asymptotic significance;

^b Area under ROC curve

Note: p values are calculated based on paired samples t test. Effect sizes are reported as Cohen d .