Use of the Research Diagnostic Criteria for Temporomandibular Disorders for Multinational Research: Translation Efforts and Reliability Assessments in The Netherlands

Frank Lobbezoo, DDS, PhD
Professor

Maurits K. A. van Selms, MSc
Graduate Student

Department of Oral Function
Academic Centre for Dentistry Amsterdam (ACTA)
Amsterdam, The Netherlands

Mike T. John, DDS, PhD, MPH, PhD
Assistant Professor
Department of Prosthodontics and Material Science
University of Leipzig, Germany

Kimberly Huggins, RDH, BS
Clinical Research Manager
Department of Oral Medicine
University of Washington
Seattle, Washington

Richard Ohrbach, DDS, PhD
Associate Professor
Department of Oral Diagnostic Sciences
University at Buffalo, New York

Corine M. Visscher, PT, PhD
Assistant Professor

Jacques van der Zaag, DDS
Assistant Professor

Marylee J. van der Meulen, MSc
Psychologist and Assistant Professor

Machiel Naeije, PhD
Biophysicist, Professor, and Chair

Department of Oral Function, ACTA
Amsterdam, The Netherlands

Samuel F. Dworkin, DDS, PhD
Professor Emeritus
Departments of Oral Medicine and Psychiatry and Behavioral Sciences
University of Washington
Seattle, Washington

Correspondence to:
Prof Dr Frank Lobbezoo
Department of Oral Function, ACTA
Louwesweg 1
1066 EA Amsterdam, The Netherlands
Fax: +31 20 5188414
E-mail: f.lobbezoo@acta.nl

Aims: *To outline the steps taken to conduct and to culturally adapt Dutch translations of the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) history questionnaire, clinical examination form, and verbal instructions to the patients, and to assess the reliability of the clinical examination.* ***Methods:*** *For the linguistic translation from English into Dutch, the forward and back-translation approach was followed. For cultural adaptation, an expert panel reviewed the translation, and a pretest was performed on a small clinical sample. Examiner training and calibration were carried out, and the clinical reliability of a "gold standard examiner" and 3 clinicians was assessed on 18 symptomatic TMD patients and 6 asymptomatic controls. The order of the examinations was based on a quasi-random Latin square design. Intraclass correlation coefficients (ICCs) were calculated to assess the overall interexaminer reliability of the clinical examination.* ***Results:*** *A linguistically valid and culturally equivalent translation of the RDC/TMD into Dutch resulted from the above-outlined procedure. As for the clinical reliability, the ICC values obtained could mostly be considered "excellent" or, less frequently, as "fair to good." Poor reliability was found only for some of the palpation tests. For uncommon diagnoses (disc displacement without reduction and without limited mouth opening; osteoarthritis), no reliable ICC value could be calculated.* ***Conclusion:*** *The mode described by the authors for preparing clinical sites for RDC/TMD-based research is a feasible one.*
J OROFAC PAIN 2005;19:301–308

Key words: clinical examination, questionnaire, reliability, temporomandibular disorders, translation

In 1992, an expert panel published the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) in order to redress the lack of standardized diagnostic criteria for defining clinical subtypes of TMD.[1] Since then, these criteria have been widely used in epidemiological and clinical research. Several research groups are currently participating in the International Consortium for RDC/TMD-based Research ("International RDC/TMD Consortium"), which was founded in 2000. One of this consortium's goals is to establish an association of multinational centers that have the capability to conduct clinical and basic research, on an international and collaborative level, into the etiology and management of TMD. Prerequisites for a working multinational consortium are the availability of *(1)* properly conducted and culturally adapted translations of relevant parts of RDC/TMD publications, questionnaires, forms, and specifications for every participating culture and language and *(2)* reliable examiners using the RDC/TMD in all participating clinical centers.

The aim of this article is to outline the steps taken by the Clinic for Temporomandibular Disorders of the Department of Oral Function, Section of Oral Kinesiology, of the Academic Centre for Dentistry Amsterdam (ACTA) in The Netherlands to meet the 2 prerequisites formulated by the International RDC/TMD Consortium, particularly the translation of the English language version of the RDC/TMD history questionnaire, clinical examination form, and verbal instructions to the patients into the target language (viz, Dutch), and to assess the reliability of the clinical examination.

Materials and Methods

RDC/TMD

The RDC/TMD[1] consist of 4 parts: *(1)* a history questionnaire; *(2)* a clinical examination form; *(3)* specifications for the clinical examination, including a set of verbal instructions to the patients during the physical examination; and *(4)* an algorithmic protocol for scoring the RDC/TMD Axes I and II. A detailed description of all 4 parts of the RDC/TMD can be found at www.rdc-tmdinternational.org. For use in multinational studies, the history questionnaire, the clinical examination form, and the verbal instructions to patients especially need to be translated; the remainder of the specifications for the clinical examination and the scoring protocol are for the use of the clinical examiner only. In the Netherlands, most dental professionals have a thorough command of the English language, so there is no need for an official translation of these parts of the RDC/TMD. However, the self-administered history questionnaire, which is completed by the patient, needs to be translated officially to ensure equivalence with the source language. Where cultural equivalence (defined as the outcome of cross-cultural adaptation, ie, the process that considers both language and cultural issues in an attempt to prepare a questionnaire for use in another setting)[2] is a general goal for translations, this holds especially true for the demographic questions of the history questionnaire. The clinical examination form should be understandable not only for the examiner but also for the possible chairside scorer, who may not be fluent in English. Since verbal commands are used to direct the patient during the clinical examination, equivalence with the source language, and thus an official translation, is required for that part of the RDC/TMD as well.

RDC/TMD History Questionnaire. In the source language (ie, English), the history questionnaire consists of 31 questions. Briefly, questions 1 through 6 deal with the patients' perception of their general oral health (rated on 5-point ordinal scales ranging from 1 = excellent to 5 = poor), the presence or absence of facial pain in the past month (yes/no), time elapsed since onset (in years or months), its stability over time (3-point ordinal scale: 1 = persistent, 2 = recurrent, and 3 = 1 time), and any previous treatment of the facial pain. Questions 7 through 13 aim to quantify the patients' graded chronic facial pain on 11-point numerical scales, where 0 represents "No pain" or "No interference" [with daily activities] and 10 represents "Pain as bad as could be" or "Unable to carry on any activities."[3] Questions 14 through 18 are yes/no questions that give an impression of the presence or absence of joint problems such as clicking, locking, and rheumatoid arthritis. Further, they deal with injuries to the jaw or face and with headaches. Question 19 addresses the amount of jaw disability, also by means of yes/no questions. Question 20 is used to check for depression and nonspecific physical symptoms (both with "pain items included" and with "pain items excluded"), using the relevant scales of the Symptom Checklist 90 (SCL-90).[4] The answers to all SCL-90 questions are scored on a 5-point ordinal scale, ranging from 0 (not at all) to 4 (extremely). With questions 21 and 22, an impression can be obtained of the patients' opinion about their ability to take care of their own oral health (5-point ordinal scales, ranging from 1 = excellent to 5 = poor). Gender and age are documented in questions 23 and 24. Finally, questions 25 through 31 provide insight into the patients' demographic characteristics, including ancestry, level of education, marital status, income, and residential area. The examiner can use the Axis II scoring protocol for calculating the graded chronic pain score as well as the mood and psychosocial functioning scale items (SCL-90).

RDC/TMD Examination Form. In brief, the clinical examination form includes 2 questions about the patients' pain complaint (side and area). The examination of the active movements includes the opening pattern (eg, straight or uncorrected or corrected lateral deviation), the vertical range of motion (maximum unassisted opening without pain; maximum unassisted opening), excursions in the transverse plane, and the presence or absence of joint sounds. The patient's pain report is recorded for both vertical and transverse mandibular movements. In addition, 12 sites are palpated on both

sides; 8 extraoral muscle sites, 2 intraoral muscle sites, and 2 temporomandibular joint (TMJ) sites. Diagnostic scoring algorithms are available for use by the examiner to establish the Axis I clinical diagnoses by combining several clinical signs and patient-reported symptoms. Multiple diagnoses per patient and per side are possible, with a maximum of 5 per patient. Diagnoses are classified as belonging to 1 of the following groups: Group I—myofascial pain (with or without limited opening); Group II—disc displacement (with or without reduction; and, in the latter case, with or without limited opening) on the left and/or right side; or Group III—arthralgia, osteoarthritis, or osteoarthrosis on the left and/or right side.

Verbal Instruction. As part of the specifications for the clinical examination, a set of verbal instructions to the patients during the physical examination has been formulated as to ensure a standardized way of communication. For example, to determine the vertical range of motion regardless of the presence or absence of pain, the examiner instructs the patient as follows: "I would like you to open your mouth as wide as you can, even if it's painful."

Translation into Dutch. The history questionnaire, clinical examination form, and verbal instructions were translated from English into Dutch through the use of the forward and back-translation approach.[2] The Orofacial Pain Research Group of the University at Buffalo provided the Amsterdam group with a set of translation guidelines. Following these guidelines, 2 independently working persons translated the source documents into the target language. Both bilingual translators had Dutch as their native language. One of them (MS) was an expert in the orofacial pain area; for the other translator, a person with no specific experience in that field was intentionally selected to provide better generalization for the typical user of the patient questionnaires. Both translators and a coordinator (FL) then conducted a synthesis of the 2 translations, thus producing 1 common translation for further use. The common forward translation was translated back into the original language by an independent, professional translator from The International Institute of Buffalo, whose native language was English, who was blind to the source documents, and who was not an expert in the orofacial pain area. The resulting back-translation was then reviewed against the source document by a second coordinator (the reviewer, RO), who constructed a translation summary document (independent review) describing all areas of discrepancy between the back-translation and the source document. Among others, discrepancies concerned the translation of the following questions of the history questionnaire:

- Questions 1 and 2: "Fair" and "poor" were back-translated as "adequate" and "just so-so"
- Question 5: "Recurrent" was back-translated as "intermittent"
- Question 14a: "lock or catch" was back-translated as "stuck"
- Question 15a: "click or pop" was back-translated as "crack or snap"
- Question 15b: "grating or grinding" was back-translated as "grating or scraping"
- Question 15e: "ache" was back-translated as "hurt"
- Question 18: "have you had" was back-translated as "have you suffered"
- Question 20.l: "feeling blue" was back-translated as "feeling of depression"

No important discrepancies were present regarding the clinical examination form and the verbal instructions. The independent review was sent to the translators, who, in case of agreement with the recommendations, made the necessary corrections to the translation. All changes and disagreements were documented and explained in the summary document, which was sent back to the reviewer and the back-translator for evaluation. No repetition of the above procedure was needed, indicating a successful back-translation of the changes. In addition, an expert panel comprising a linguist, 2 epidemiologists/methodologists, a psychologist, and 3 TMD specialists reviewed the resulting translation with respect to semantic, idiomatic, experiential, and conceptual equivalences. A few recommendations for minor (mainly idiomatic) revisions of the Dutch translation were made. Further, because of the presence of norm scores for use within The Netherlands, the newly translated SCL-90 scales were substituted by the formerly translated and validated SCL-90 scales by Arrindell and Ettema.[5] Finally, small-scale pretesting in the Clinic for Temporomandibular Disorders at ACTA, for which about 30 TMD patients were interviewed about their opinion of the new questionnaire, did not result in any recommendations for revision. The final Dutch version was then posted on the website of the International RDC/TMD Consortium.[6]

As outlined above, questions 23 through 31 of the RDC/TMD history questionnaire are concerned with the patients' demographic characteristics. To take into account societal differences between the United States and The Netherlands with respect to

demographic factors such as level of education, marital status, income, and residential area, Statistics Netherlands (Centraal Bureau voor de Statistiek [CBS], a Dutch organization responsible for collecting, processing, and publishing statistics to be used in practice, by policymakers, and for scientific research) was approached for assistance. The resulting items can be found at the website of the International RDC/TMD Consortium.[6]

Training, Calibration, and Reliability Assessment

Since the reliability assessment needs to be preceded by a training and calibration session,[6] the University of Washington Orofacial Pain and Disorders Clinic and Research Group provided the Department of Oral Function of the Academic Centre for Dentistry Amsterdam (ACTA) with a gold standard examiner (GSE) (KH) to train and calibrate 3 examiners (FL, CV, and JZ) to become reliable in the conduct of the RDC/TMD Axis I examination and diagnosis methods. The scientific and ethical aspects of the protocol were reviewed and approved by the board of the Netherlands Institute for Dental Sciences.

Pretraining. Prior to the GSE's visit to ACTA, the examiners-in-training (EXTs) memorized the specifications for the clinical examination, including the exact verbal instructions to the patients. Further, the EXTs watched the RDC/TMD examination training videotape provided by the University of Washington Group until they felt comfortable with the protocol. Finally, the EXTs mastered the delivery of 1 lb and 2 lbs of pressure on a postage scale, which is needed for the standardized palpation of structures of the masticatory system as part of the clinical examination.

Training. The 3-hour training session started with the GSE's overview of the examination procedures. Subsequently, the GSE conducted the clinical examination on the 3 EXTs who, in turn, practiced on the GSE. The GSE provided the EXTs with the necessary feedback and clarifications. The EXTs then practiced the examination on each other. This procedure was repeated until the GSE determined the EXTs to be competent in the RDC/TMD clinical examination protocol.

Calibration. During the 3-hour calibration session, the GSE as well as the 3 EXTs individually conducted the entire clinical examination on 2 preselected symptomatic TMD patients. A recorder assigned to each chair kept record of observed variations among the EXTs in examination technique and verbal instructions, besides recording the examination itself. The examinations were then reviewed and, when necessary (ie, in case of discrepancies), re-examination of the patients was performed.

Reliability Assessment. A total of 24 subjects (5 men and 19 women from 19 to 62 years old; mean age ± SD 39.5 ± 14.5 years; 18 symptomatic TMD patients and 6 asymptomatic controls) participated in the 6-hour reliability assessment. The TMD patients were preselected so as to present a broad range of symptoms. Using a quasi-random Latin square design to control for order effects, the GSE and the 3 EXTs individually examined all 24 participants, blinded from each other's findings. The examiners moved from chair to chair, while the participant remained seated in the same place. The recorder translated the verbal instructions of the GSE into Dutch.

Data Analyses

Descriptive statistics were computed for all outcome measures, based on the number of observations—either 96 (24 participants × 4 examiners) or 192 (24 participants × 4 examiners 5 × sides). To facilitate interpretation, bilateral measures (eg, joint sounds) were combined into single variables. Further, for palpation, the 4-point ordinal scale was dichotomized into "no pain" versus "any pain" (ie, pain, regardless of intensity) categories. Similarly, joint sounds were dichotomized into "click" and "other/no joint noises" categories. For continuous measures (ie, the mandibular range of motion variables), mean, standard deviation, and range were calculated, while for dichotomous measures, the proportion of positive ratings was calculated.

Overall interexaminer reliability was assessed by calculating intraclass correlation coefficients (ICCs).[7] ICCs were used for both continuous variables and dichotomous variables, following the recommendations of John et al.[8] In addition, the agreement between all possible pairs of examiners was calculated for the dichotomous measures. ICC values were interpreted according to Fleiss[9]: ICC < 0.4 = poor reliability; 0.4 ≤ ICC ≤ 0.75 = fair to good reliability; and ICC > 0.75 = excellent reliability.

All statistical analyses were performed using STATA, release 7.0 (Stata Statistical Software).

Results

The outcomes of the reliability assessment are shown in Tables 1 through 4. The mandibular range of motion measures are presented in Table 1. For most of these measures, ICCs were larger than

Table 1 Mandibular Range of Motion Variables: Descriptive Statistics and ICCs

		Range of motion (mm)			
Variable	n	Mean	SD	Min–Max	ICC
Unassisted opening without pain	96	43.3	11.2	19–62	0.86
Maximum unassisted opening	95	50.2	9.4	32–66	0.94
Maximum assisted opening	94	52.6	8.8	35–66	0.93
Laterotrusion	190	9.7	2.3	4–14	0.71
Protrusion	96	5.3	2.9	0–16	0.88

Table 2 Joint Sounds: Percent Positive Ratings, ICCs, and Examiner Agreement

Variable	n	Positive ratings (%)	ICC	Agreement between examiner pairs (%)
Clicking during opening	183	26	0.82	93
Clicking during closing	183	25	0.75	91
Clicking during laterotrusion (contralateral side)	191	10	0.62	93
Clicking during laterotrusion (ipsilateral side)	189	1	*	99
Clicking during protrusion	189	13	0.74	94
Eliminated click	188	13	0.49	88

* Prevalence too low to calculate a reliable ICC.

Table 3 TMJ and Masticatory Muscle Palpation: Percent Positive Ratings, ICCs, and Examiner Agreement

Variable	n	Positive ratings (%)	ICC	Agreement between examiner pairs (%)
Temporalis				
Posterior	192	5	0.10	91
Middle	192	17	0.44	84
Anterior	192	17	0.56	87
Masseter				
Superior	192	27	0.75	90
Body	192	36	0.63	83
Inferior	192	31	0.58	82
Posterior mandibular region	192	18	0.35	81
Submandibular region	192	10	0.29	87
Lateral pterygoid area	187	56	0.61	80
Tendon of temporalis	189	40	0.44	73
TMJ				
Lateral	192	22	0.38	78

Table 4 RDC/TMD Diagnoses: Percent Positive Ratings, ICCs, and Examiner Agreement

Variable	n	Positive ratings (%)	ICC	Agreement between examiner pairs (%)
Myofascial pain	95	26	0.74	89
Myofascial pain with limited mouth opening	95	24	0.57	83
Disc displacement				
With reduction	183	20	0.62	88
Without reduction	192	4	0.56	97
Without reduction and without limited opening	191	1	*	99
Arthralgia	189	16	0.45	86
Osteoarthritis	190	0	*	100
Osteoarthrosis	187	3	0.79	99

* Prevalence too low to calculate a reliable ICC.

0.75, which indicates excellent reliability. Only for laterotrusion was an ICC value found that could be considered only fair to good.

Table 2 shows the reliability assessment for joint sounds. Clicking during opening, closing, and protrusion could be assessed reliably, as indicated by their high ICC values. Interestingly, these measures also had the highest proportions of positive ratings. However, even though the elimination of a click was found possible in a similar number of cases as clicking during protrusion was found (n = about 190 and percent positive ratings = 13% for both variables), the ICC value for the elimination of a click was relatively low. For clicking during laterotrusion on the contralateral side, a relatively low proportion of positive ratings went with an only "fair-to-good" ICC value. Clicking on the ipsilateral side during laterotrusion was so rare that no reliable ICC could be calculated for this measure.

The reliability assessment of the scores that were recorded during the palpation of the masticatory muscles and TMJs is presented in Table 3. Only the ICC value of the superior masseter muscle was found to be excellent; the other ICC values were either poor or fair to good.

With the examination outcome measures that are shown in Tables 1 through 3, clinical diagnoses were established using the RDC/TMD Axis I scoring protocol. The reliability assessment thereof is presented in Table 4. The proportion of positive ratings indicates that in 50% of the participants, a diagnosis of myofascial pain (with or without limited mouth opening) was established. In 20% of the participants, a disc displacement with reduction

was found, while 16% suffered from arthralgia. The other clinical diagnoses were less prevalent. Reliability was fair to good for both myofascial pain diagnoses, for disc displacement with and without reduction, and for arthralgia. Despite its low prevalence in this sample, osteoarthrosis had an ICC value that could be qualified as excellent. For the rarest conditions, disc displacement without reduction and without limited mouth opening (found in only 1% of the examinations) and osteoarthritis (not found at all), a proper reliability assessment could not be performed.

Discussion

This paper shows that preparing a clinical site for multinational research using the RDC/TMD[1] is an intensive process. It involves a thorough and time-consuming translation of the relevant parts of the RDC/TMD publications, questionnaires, forms, and specifications, in which many persons at different sites play important roles. Further, careful training, calibration, and reliability assessment of examiners at the clinical site is part of the preparation process. The fact that no iteration of (parts of) the translation process was necessary at ACTA and that, in general, the reliability of the examiners was good suggests that this mode of preparing clinical sites is a feasible one, and that ACTA is a reliable center for RDC/TMD-based research.

Translation

For the translation process, the guidelines compiled by the Orofacial Pain Research Group of the University at Buffalo were used. These guidelines are adapted from various published recommendations (eg, Guillemin et al[2]). The Buffalo guidelines advise the use of 2 translators; Guillemin et al[2] consider 2 translators (1 an expert in the field of interest and the other having no specific experience there) as an essential prerequisite for a properly performed translation process. Since 2 translators whose results were synthesized into a single translation were used, the Dutch translation of the RDC/TMD fulfills the "2-translator" criterion. Although Guillemin et al[2] suggested the use of 2 back-translators as to increase the likelihood of highlighting any imperfections, only 1 "naive" professional back-translator was used for the Dutch translation. This may be considered a weakness of the translation process. However, according to the guidelines compiled by the Orofacial Pain Research Group of the University at Buffalo, the use of a single back-translator is sufficient. The procedure following the completion of the back-translation (viz, review, corrections, and evaluation), on the other hand, is again in line with the criteria described by Guillemin et al.[2] Thus, despite the possible weakness of a single back-translator, the resulting translation fully fulfills the criteria of the University at Buffalo and thereby, of the International RDC/TMD Consortium.

The linguistically valid translation of the RDC/TMD into Dutch, as described here, can be considered as the first step in the necessary cultural adaptation process: It cannot be assumed a priori that the original (English) version and the translated (Dutch) version are culturally invariant. For cultural adaptation of the Dutch version, an expert panel therefore reviewed the translation, and a small-scale pretest was performed. Apart from the suggested final field-testing, which will be performed in a future, large-scale, multinational study of cultural influences on TMD pain, cultural equivalence of the Dutch version was thus verified. Pending the outcome of this field test, the Dutch translation of the RDC/TMD was considered ready for use within The Netherlands. Other Dutch-speaking countries or regions (eg, Flanders in Belgium) should be aware of possible cultural differences and that the demographic questions, 25 through 30, need to be adapted to the local situation before the Dutch translation can be used in full.

Reliability Assessment

In this study, ICCs were used not only for continuous variables but also for categorical (dichotomous) data. In most reliability assessments, kappa statistics are used for categorical data.[9] According to John et al,[8] however, the distinction between the 2 approaches is only arbitrary: They are mathematically equivalent, except for a term in the denominator that becomes negligible when the number of subjects increases to that used in the present study. Indeed, when both ICC and kappa values are calculated for the same variable, the differences were negligible. For example, palpation of the superior, middle, and inferior parts of the masseter muscle yielded respective kappa values of 0.74, 0.63, and 0.57, while respective ICC values of 0.75, 0.63, and 0.58 were obtained (Table 3). Besides being equivalent, ICC analysis offers some additional information beyond that obtained with kappa statistics. It enables a more in-depth analysis of, for example, order effects, and uncovers areas for potential improvement.[8] Therefore, ICC analysis was chosen over kappa analysis in the present study.

Range of Motion. Wide ranges of values were found for the mandibular range of motion variables. It is noteworthy that the lowest ICC values were found for laterotrusion. This is in line with the results of Goulet et al[10] and of John and Zwijnenburg,[11] who also found the lowest, albeit still acceptable, ICC values for the laterotrusions in their respective study samples. Apparently, subjects have more difficulties in performing maximal mandibular excursions to the sides than in making open and protrusion excursions in a reproducible manner. It is common clinical experience that patients frequently have difficulties in performing (maximum) laterotrusions. Alternatively, laterotrusions may have been relatively difficult to measure reliably for the examiners. Since the magnitude of lateral excursions is part of the RDC/TMD diagnoses of disc displacement without reduction, with or without limited opening, the TMD subtyping[1] may have been affected by the lower ICC value for laterotrusion. More explicitly, since reliability sets the upper limit of validity, the reliability problem associated with laterotrusive measurements may have a significant impact on the validity of these specific diagnoses.

Joint Sounds. TMJ sounds are frequently occurring phenomena, both in TMD populations (eg, Lobbezoo-Scholte et al[12]) and in non-TMD populations (eg, Huddleston Slater et al[13]). Several types of joint sounds can be distinguished. The most common types are sounds related to internal derangements, such as anterior disc displacement with reduction and hypermobility, while crepitus is found less frequently.[13,14] In the present study, clicking sounds were found in one fifth of the observations (disc displacement with reduction in Table 4). Contrary to this relatively common phenomenon, "coarse" crepitus was found in only 3% of the observations (osteoarthrosis in Table 4). Since this latter prevalence was too low to be considered on its own in the present reliability assessment, the joint sound measures were dichotomized into "clicking sounds" and "other/no sounds."

A high reliability was found for the observation of clicks during mandibular opening, closing, protrusion, and, to a slightly lesser extent, during laterotrusion to the contralateral side. In part, these findings are in line with previous studies (eg, Wabeke et al[15]; De Wijer et al[16]). The fact that no reliability assessment could be performed for clicking on the ipsilateral side during lateral excursive movements is most likely related to the rarity of that phenomenon. In turn, this rarity is undoubtedly related to the fact that the condyle only makes a very small (mostly lateral) movement on the ipsilateral side, which is known as the "immediate side shift" or "Bennet movement."[17] The smaller the condylar movement, the more unlikely it is for a joint sound to occur. Moreover, clicking sounds during such small movements may or may not be caused by displaced discs; this awaits further study.

Relatively low ICC values were found for the elimination test ("protrusive opening"),[1,13] a technique that can be used to differentiate between clicking sounds on the basis of an anterior disc displacement and those with another underlying mechanism (eg, hypermobility). Since clicking sounds upon opening, closing, and protrusion could be established reliably, it is difficult to explain these low ICC values. Perhaps clicks are not stable in terms of always occurring at the same amount of mouth opening, even though the occurrence of clicking was found to be stable over a period of 10 days.[18] If closing clicks sometimes occur farther away from maximal occlusion than they typically do,[19] elimination upon protrusive opening may be prevented in such occurrences and thus may be an unreliable confirmation test.

Palpation. The scores for palpation were dichotomized into "no pain" and "any pain" groups, because the measures are used as such to establish the RDC/TMD diagnoses. As for the joint sounds, a large-enough prevalence of a certain score yielded acceptable reliability values, although they were generally lower than those obtained for range of motion measurements and for the detection of joint sounds. This corroborates the results of several previous studies (eg, Goulet et al[10] and De Wijer et al[16]). It is noteworthy that palpation of the intraoral sites (of which the lateral pterygoid area is especially difficult in terms of technical performance of the palpation and the possibility of provocation of surrounding tissues instead of the muscle itself) yielded relatively high proportions of positive ratings. This means that because of the difficulties attached to the palpation of these sites, they may have contributed disproportionally to the TMD subtyping process, which raises the question of whether such palpation sites should be included in a diagnostic approach such as the RDC/TMD.[20, 21]

Diagnoses. For the common diagnoses, ie, myofascial pain, disc displacement with reduction, and arthralgia, acceptable reliability values were found. Again, the more uncommon a diagnosis was, the lower the reliability. However, the uncommon condition "osteoarthrosis" had very high ICC values. This illustrates another possible consequence of a low prevalence: by chance, ICC values can also be very high. The results, however,

cannot be interpreted. This stresses the need for a meta-analysis of the data of multiple clinics that participate in the International RDC/TMD Consortium. In the present study sample, about half of the 18 TMD patients had myofascial pain, which is about 3 times more than the number of patients with arthralgia. This relative distribution (3:1) is a corroboration of the Swedish and US data published by List and Dworkin,[22] as well as of the Asian data published by Yap et al.[23] In addition, a previous clinical study in the authors' department revealed a similar distribution of myogenous and arthrogenous TMD pain.[24] This suggests that the present study sample is representative of the TMD pain patient population of the Clinic for Temporomandibular Disorders of ACTA.

Acknowledgments

The authors thank Anna Lou van Os-Ravesloot, LLM, for taking care of the "naive" forward translation. The International Institute of Buffalo for the back-translation; and Jules R. Hesse, James J.R. Huddleston Slater, Nico Hofman, and Diana E.J. Meijer for their help during the reliability assessment. This study was supported by a US National Institute for Dental and Craniofacial Research planning and development grant (1 R21 DE13722-01), the Orofacial Pain Research Group of the University at Buffalo, and The Netherlands Institute for Dental Sciences.

References

1. Dworkin SF, LeResche L. Research Diagnostic Criteria for Temporomandibular Disorders: Review, criteria, examinations and specifications, critique. J Craniomandib Disord 1992;6:301–355.
2. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–1432.
3. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992;50:133–149.
4. Derogatis LR, Cleary PA. Confirmation of the dimensional structure of the SCL-90: A study in construct validation. J Clin Psychol 1977;33:981–989.
5. Arrindell WA, Ettema JHM. SCL-90: Handleiding bij een Multi-dimensionele Psychopathologie-indicator. Lisse, The Netherlands: Swets Test Service, 1986.
6. International Consortium for RDC/TMD-Based Research. RDC for TMD Booklet. Available at: www.rdc-tmdinternational.org/booklet/frmBooklet.htm
7. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979;86:420–428.
8. John MT, Mancl L, Dworkin SF. Temporomandibular disorders (TMD) reliability assessment using intraclass correlation coefficient. J Dent Res 2001;80:84.
9. Fleiss JL. The Design and Analysis of Clinical Experiments. New York: Wiley, 1986.
10. Goulet JP, Clark GT, Flack VF. Reproducibility of examiner performance for muscle and joint palpation in the temporomandibular system following training and calibration. Community Dent Oral Epidemiol 1993;21:72–77.
11. John MT, Zwijnenburg AJ. Interobserver variability in assessment of signs of TMD. Int J Prosthodont 2001;14:265–270.
12. Lobbezoo-Scholte AM, De Leeuw JRJ, Steenks MH, Bosman F, Buchner R, Olthoff LW. Diagnostic subgroups of craniomandibular disorders. Part I: Self-report data and clinical findings. J Orofac Pain 1995;9:24–36.
13. Huddleston Slater JJR, Lobbezoo F, van Selms MKA, Naeije M. Recognition of internal derangements. J Oral Rehabil 2004;31:851–854.
14. Okeson JP (ed). Orofacial Pain. Guidelines for Assessment, Diagnosis, and Management. The American Academy of Orofacial Pain. Chicago: Quintessence, 1996.
15. Wabeke KB, Spruijt RJ, van der Zaag J. The reliability of clinical methods for recording temporomandibular joint sounds. J Dent Res 1994;73:1157–1162.
16. De Wijer A, Lobbezoo-Scholte AM, Steenks MH, Bosman F. Reliability of clinical findings in temporomandibular disorders. J Orofac Pain 1995;9:181–191.
17. Davies SJ, Gray RJM (eds). A Clinical Guide to Occlusion. London: British Dental Association, 2002.
18. Gallo LM, Svoboda A, Palla S. Reproducibility of temporomandibular joint clicking. J Orofac Pain 2000;14:293–302.
19. Huddleston Slater JJR, Lobbezoo F, Naeije M. Mandibular movement characteristics of an anterior disc displacement with reduction. J Orofac Pain 2002;16:135–142.
20. Johnstone DR, Templeton MC. The feasibility of palpating the lateral pterygoid muscle. J Prosthet Dent 1980;44:318–323.
21. Türp JC, Minagi S. Palpation of the lateral pterygoid region in TMD—Where is the evidence? J Dent 2001;29:475–483.
22. List T, Dworkin SF. Comparing TMD diagnoses and clinical findings at Swedish and US TMD centers using Research Diagnostic Criteria for Temporomandibular Disorders. J Orofac Pain 1996;20:240–253.
23. Yap AUJ, Dworkin SF, Chua EK, List T, Tan KBC, Tan HH. Prevalence of temporomandibular disorder subtypes, psychologic distress, and psychosocial dysfunction in Asian patients. J Orofac Pain 2003;17:21–28.
24. Visscher CM, Lobbezoo F, de Boer W, van der Zaag J, Naeije M. The prevalence of cervical spinal pain in craniomandibular pain patients. Eur J Oral Sci 2001;109:76–80.