Critical Commentary 2

Reliability and Validity of the DC/TMD Axis I

Eric Schiffman, DDS, MS
Professor
Department of Diagnostic and Biological Sciences
University of Minnesota
Minneapolis, Minnesota, USA

Richard Ohrbach, DDS, PhD, Odont Dr (hon)
Professor
Department of Oral Diagnostic Sciences
School of Dental Medicine
University at Buffalo
Buffalo, New York, USA

We appreciate the opportunity to comment on the Focus Article by Drs Steenks, Türp, and de Wijer.[1] We have focused our comments on their most significant points regarding the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD).[2] First, we will summarize the principles used when developing the DC/TMD Axis I algorithms[3]:

- Include parts of the Research Diagnostic Criteria for TMD (RDC/TMD) when indicated
- Keep the DC/TMD Axis I as simple as possible so clinicians and researchers will use it
- Use well-operationalized, reliable tests
- Use case definitions, derived by consensus from experts in TMD and allied fields, to determine which tests to include or exclude from the Revised RDC/TMD Axis I diagnostic algorithms
- Test these case definitions for diagnostic accuracy using the reference standards from the Validation Project

The authors of the Focus Article state that there was no independent reference standard for the diagnoses in the Validation Project since "...the test examiners and the criterion examiners, by knowing and using the algorithms, are not independent, thus leading to a certain degree of circularity." However, we point out that the reference standard diagnoses used by the criterion examiners were independent of findings by the test examiners, and examination protocols were not based on the RDC/TMD Axis I algorithms.[3] Circularity was addressed by: *(1)* inclusion of participants who would not meet criteria for RDC/TMD diagnoses; *(2)* inclusion in the criterion examiner assessment protocol of all RDC/TMD tests (history, mobility, palpation, and noise detection) and independent tests comprised of additional history, examination procedures, and imaging (ie, panoramic radiograph, bilateral temporomandibular joint [TMJ] magnetic resonance imaging [MRI], and computed tomography [CT]); *(3)* an expanded reference standard taxonomy independent of the RDC/TMD (consistent in scope with the Expanded DC/TMD Taxonomy[4]); *(4)* independent examination of participants by two calibrated criterion examiners at each of three sites; *(5)* consensus diagnoses by the two criterion examiners at each site as the reference standard for pain-related diagnoses; and *(6)* reading of images by calibrated, blinded radiologists as the reference standard for TMJ intra-articular disorders.[3]

We also respectfully disagree with the statement by Steenks et al that "The help of imaging techniques for the criterion examiners is not useful, since the association between symptoms and imaging results is low." It is true that MRI-detected TMJ disc displacements (DD) and CT-detected degenerative joint disease (DJD) often have no clinical signs or symptoms associated with them, including TMJ noise, and the resultant algorithms had low diagnostic utility compared to imaging. However, the best available independent reference standard for detection of DD and DJD is TMJ MRI and CT, respectively, read by calibrated, blinded radiologists. The suggestion for not using imaging as a reference standard for detection of these disorders because clinical tests are not associated with these imaging findings is counter to Standards for Reporting of Diagnostic Accuracy (STARD) methods.[5] Our findings point out the inadequate diagnostic accuracy and low sensitivity of current non-imaging tests to detect DD and DJD rather than the disqualification of imaging as a reference standard.

With respect to the test population, Steenks et al[1] state:

> *Another selection bias may be the low threshold for having at least one of the three cardinal TMD symptoms; ie, jaw pain, limited mandibular movement (in most cases, restricted jaw opening), and TMJ noise . . . Due to the low threshold for being a case,*

doi: 10.11607/ofph.2018.1.cc2

the study group in the Validation Project contained many community cases without a treatment demand. In future study groups, the reliability of classified conditions needs to be tested in TMD patients with a concomitant demand for therapy in order to better simulate clinical patients.

However, we note that the above three cardinal signs of TMD were used for screening purposes for study entry in order to ensure a full spectrum of symptoms and signs among study participants.[3] The recommendation by Steenks et al would likely result in spectrum bias. Since seeking treatment is not a criterion for having any of the common pain-related or intra-articular disorders, restricting the spectrum of participants in diagnostic validation research actually reduces the generalizability of the findings. Furthermore, we are not aware of any evidence (nor do the authors cite supporting evidence) that seeking care alters the reliability or validity of TMD diagnoses. Restriction of participants to only those seeking care would lead to easier cases to classify, which, in our opinion, would worsen generalizability for the Axis I algorithms. Consequently, our sample of 705 participants (614 TMD cases and 91 controls), further apportioned to 114 cases without pain (107 community and 7 clinic cases) and 500 cases with TMD pain diagnoses (359 community and 141 clinic cases), resulted in a full spectrum and severity of TMD signs and symptoms in community and clinic cases, which was used to establish credible reference standards.[3] For the above reasons, we believe that the test population used for assessing the criterion validity of the DC/TMD Axis I algorithms was credible and appropriate.

In the case of rendering pain diagnoses, the Focus Article states: "The presence of both conditions (pain on palpation and during mandibular function) seems more adequate as a criterion to apply in the confirmation of TMD." We respond by pointing out that the current criteria for pain-related TMD require only one provocation test (palpation or range of motion) to be positive, but also require that, by history, pain be altered by jaw movement, function, or parafunction. From one perspective, the current criteria achieve what Steenks et al recommend, but not in the way they want the criteria to be structured. From another perspective, they disagree with the expert consensus that led to the current criteria and also with empirical results supporting the present structure of the criteria. They provide no data in support of their recommendation, so we encourage them to test their hypothesis empirically.

The Focus Article also comments on additional classification categories. It first notes that the utility of the three myalgia subgroups has been questioned in a Letter to the Editor[6] by some authors of the DC/TMD. The three myalgia subgroups are local myalgia, myofascial pain, and myofascial pain with referral. We believe that Steenks el al may be confusing utility and validity. The Letter to the Editor concerned the addition of local myalgia and myofascial pain, with unknown criterion validity, to the DC/TMD, which was made at the final stage prior to publication to be consistent with the section regarding the most common TMD in the expanded DC/TMD taxonomy.[4] In addition, in one of the DC/TMD workshops, we presented these three myalgia subgroups, and some workshop participants wanted these retained so that different pathophysiologic mechanisms that may be occurring in individuals with these different diagnoses could be investigated. Finally, the current structure of three myalgia subdiagnoses permits an inclusive set of conditions that provides an inclusive framework for empirically testing the different ways in which muscle pain may manifest. For example, we are investigating these diagnoses, and our preliminary findings indicate that the clinical diagnosis of myofascial pain with referral has clinical utility for identifying muscle-pain subjects with complex biopsychosocial characteristics when compared to normal and myalgia subjects.[7] Thus, the three subdiagnoses have research utility and probable clinical utility.

The Focus Article also notes that clinically widely used and sound criteria for disc displacement with reduction (the elimination of a click while opening and closing with the mandible in protrusion and loudness of a closing click with counterforce on the mandible) have been abandoned. We note, however, that these recommended tests were fully evaluated by the criterion examiner protocol, but their inclusion into the criteria did not improve the diagnostic accuracy (ie, sensitivity, specificity) for DD with reduction. In addition, the reliability for click elimination is low, and the loudness of a click is extremely difficult to operationalize, which reduces its reliability. Ultimately, the data determined the final criteria. We encourage the authors of the Focus Article to perform the necessary research to support their recommendations. Finally, adding tests to an algorithm as Boolean "AND" (for example, click elimination) necessarily reduces sensitivity and increases specificity—but the authors' point is oriented toward better detection (ie, increased sensitivity), which cannot be realized through the means they are suggesting. No matter what we did with the clinical tests for intra-articular disorders, these tests largely had inadequate diagnostic accuracy with poor sensitivity and serve, at best, as screening tests. In clinical and research settings, the reference standards for DD and DJD are MRI and CT, respectively.

The Focus Article also states:

> *But what is the need for the separate subgroup for headache attributed to TMD when the pain has already been classified as TMD-related pain? . . . It seems more appropriate to use "myalgia (in the temporalis region)" in such cases. In the IHS classification this subgroup makes sense, but not in the DC/TMD.*

Steenks et al appear to be suggesting that the IHS and the DC/TMD should not have any overlap, yet it is such silo views that contribute to much of the controversy and confusion in pain diagnoses. The ideal situation at this time is that all health professionals—clinicians and researchers across disciplines—use the same nomenclature so that they can succinctly communicate with each other. Thus, it was imperative that this diagnosis be included, as it now is, in both the International Classification of Headache Disorders 3rd Edition (ICHD-3) Beta version[7] and the DC/TMD classification.[2]

Another point in the Focus Article is that Steenks et al argue for elimination of the need to palpate intra-oral muscles and to reduce the number of palpation sites. We agree that the DC/TMD does not require palpation of the intraoral muscles; only palpation of the temporalis and masseter muscles in order to achieve the stated sensitivity and specificity. However, as noted in the DC/TMD publication, the intraoral muscles can be included as part of the examination when clinically indicated or for specific research questions.

The Focus Article also raises the question of whether the DC/TMD is ready for clinical use and states: "Thus, although the DC/TMD represents an improvement over the RDC/TMD, its immediate implementation in research and clinical care does not yet appear to be adequately substantiated." We respond by pointing out that if it is believed that the DC/TMD should not be used in these settings, then it follows that the RDC/TMD should not have been used in the research setting. This is because the RDC/TMD has only face and content validity, but the DC/TMD Axis I has face, content, and criterion validity (and Axis II measures have concurrent validity). Also, the RDC/TMD has been used in numerous research projects and has provided consistency of assessment and diagnosis for years, which allows comparisons between findings from different studies. Furthermore, the US National Institutes of Health has essentially required its use in all TMD-based clinical grant proposals. Therefore, we strongly believe that the RDC/TMD should have been used in prior research studies and the DC/TMD should now be used in research and clinical settings. It is important that clinicians and researchers use the same core assessment protocol so they can all communicate their ideas and findings to each other using a common "language." It is relevant that in the third (beta) version of the ICHD that Dr Jes Olesen, chairman of the Headache Classification Committee of the IHS, stated: "Already now, clinicians and researchers should start using the criteria of ICHD-3 (beta)."[8] The ICHD-3 diagnostic criteria for headache and facial pain have face and content validity, just like the RDC/TMD; however, the ICHD-3 largely lacks criterion validity, unlike the DC/TMD. Yet, the ICHD-3 is the best available taxonomy for use in clinical and research settings, where criteria with criterion validity do not exist.

Finally, the authors of the Focus Article offer neither any alternative comprehensive Axis I TMD diagnostic criteria for the most common TMD nor data in support of their recommendations. In addition, neither the DC/TMD nor any other diagnostic taxonomy that we are aware of is claimed to substitute for clinical decision-making skills to rule out other disorders—which we clearly state is a necessary first step in the diagnostic process before applying the DC/TMD diagnostic protocol.[2] Therefore, we believe that the DC/TMD is currently the best available diagnostic criteria for use in both clinical and research settings. Finally, we again point out that the DC/TMD is not the final version for TMD classification—it will need revision as more research becomes available.[2] We welcome the authors to rigorously test their suggestions for possible revision of the DC/TMD; however, we should all keep in mind that, ultimately, the goal is to have a DC/TMD diagnostic taxonomy inclusive of etiology, mechanism, and prognosis.

References

1. Steenks MH, Türp JC, de Wijer A. Reliability and validity of the Diagnostic Criteria for Temporomandibular Disorders Axis I in clinical and research settings: A critical appraisal. J Oral Facial Pain Headache 2018;32:7–18.
2. Schiffman E, Ohrbach R, Truelove E, et al. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research applications: Recommendations of the International RDC/TMD Consortium Network and Orofacial Pain Special Interest Group. J Oral Facial Pain Headache 2014;28:6–27.
3. Schiffman EL, Truelove EL, Ohrbach R, et al. The Research Diagnostic Criteria for Temporomandibular Disorders. I: Overview and methodology for assessment of validity. J Orofac Pain 2010;24:7–24.
4. Peck CC, Goulet JP, Lobbezoo F. Expanding the taxonomy of the Diagnostic Criteria for Temporomandibular Disorders. J Oral Rehabil 2014;41:2–23.
5. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. Standards for Reporting of Diagnostic Accuracy. Clin Chem 2003;49:1–6.
6. Svensson P, Michelotti A, Lobbezoo F, List T. The many faces of persistent orofacial muscle pain. J Oral Facial Pain Headache 2015;29:207–208.
7. Varun S et al. Cross-sectional study on the association between masticatory muscle pain diagnoses and biopsychosocial factors [poster]. Scottsdale: American Academy of Orofacial Pain, 2017.
8. International Headache Society Classification ICHD-3 Beta. https://www.ichd-3.org.