AHEAD OF PRINT

Original paper

Cross-language readability of the Diagnostic Criteria for Temporomandibular Disorders Axis II questionnaires: Comparing the original English version and validated translations

Giuseppe Minervini[1,A–D], Hande Uzunçıbuk[2,A,D], Nicola Fiorito[1,B,D], Giulia Maria Nicoletti[1,B–D], Diana Russo[1,C,D], Emilio Alfredo Cafferata[3,B,D], Gabriele Cervino[4,E,F], Maria Maddalena Marrapodi[5,6,A,E]

[1] Multidisciplinary Department of Medical-Surgical and Odontostomatological Specialties, University of Campania Luigi Vanvitelli, Naples, Italy
[2] Department of Orthodontics, Faculty of Dentistry, Trakya University, Edirne, Turkey
[3] Oral Peri-Implant Research Group, School of Dentistry, Universidad Científica del Sur, Lima, Peru
[4] Department of Biomedical and Dental Sciences, Morphological and Functional Images, University of Messina, G. Martino Polyclinic, Italy
[5] Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, India
[6] Department of Woman, Child and General and Specialist Surgery, Pediatric Unit, University of Campania Luigi Vanvitelli, Naples, Italy

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Dental and Medical Problems, ISSN 1644-387X (print), ISSN 2300-9020 (online) *Dent Med Probl.*

Address for correspondence
Giuseppe Minervini
E-mail: giuseppe.minervini@unicampania.it

Funding sources
The study was funded by the University of Campania Luigi Vanvitelli, Naples, Italy (STARTING GRANTS 2025; Project: MARS-Q; CUP: B63C25001050005).

Conflict of interest
None declared

Acknowledgements
The authors would like to thank the University of Campania Luigi Vanvitelli for providing funding for the study.

Received on October 8, 2025
Reviewed on November 13, 2025
Accepted on December 11, 2025

Published online on November 25, 2026

Cite as
Minervini G, Uzunçıbuk H, Fiorito N, et al. Cross-language readability of the Diagnostic Criteria for Temporomandibular Disorders Axis II questionnaires: Comparing the original English version and validated translations [published online as ahead of print on November 25, 2026]. *Dent Med Probl.* doi:10.17219/dmp/215511

DOI
10.17219/dmp/215511

Copyright
Copyright by Author(s)
This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (https://creativecommons.org/licenses/by/3.0/)

Abstract

Background. The Axis II of the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) uses standardized self-report questionnaires to assess psychosocial burden, jaw function and oral behaviors. Although many translations have been psychometrically validated, their readability and potential ease of comprehension may vary across languages.

Objectives. The aim of the study was to compare the readability of the original English version of the DC/TMD Axis II questionnaires with their validated translations and evaluate whether the translated versions require higher literacy levels.

Material and methods. Validated translations of 8 DC/TMD Axis II questionnaires (Graded Chronic Pain Scale version 2.0 (GCPS 2.0), Jaw Functional Limitation Scale (JFLS-8 and JFLS-20), Patient Health Questionnaire (PHQ-4, PHQ-9 and PHQ-15), Generalized Anxiety Disorder-7 (GAD-7), Oral Behaviors Checklist (OBC)) were analyzed across English, Italian, Spanish, French, German, Brazilian Portuguese, and Turkish. Readability was assessed using 7 standardized indices. Scores were normalized and statistically compared with those of the English version using paired *t*-tests with Benjamini–Hochberg false discovery rate correction.

Results. Translated versions generally required higher reading levels than the English versions, with a mean increase of approx. 4 school years. The largest mean grade-level differences were observed for the OBC, GCPS 2.0 and JFLS-8, whereas the largest mean standardized differences were observed for the GCPS 2.0, JFLS-20 and JFLS-8; the PHQ-4, PHQ-9 and PHQ-15 showed the smallest standardized differences.

Conclusions. Many validated translations of the DC/TMD Axis II instruments impose higher literacy demands than the original English versions, particularly those of the jaw-related instruments (GCPS 2.0, JFLS-8 and JFLS-20). To ensure equitable use of the DC/TMD measures across diverse populations, readability should be routinely audited during translation workflows, with efforts made to simplify language in widely used versions.

Keywords: temporomandibular disorders, DC/TMD, readability, orthodontics, diagnosis

Highlights

- The translated versions of the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) Axis II questionnaires require reading levels equivalent to approx. 4 additional years of education compared with their original English versions.
- Italian, Spanish and German versions present the greatest readability challenges.
- Language structure has a greater influence on readability than questionnaire type.
- Readability assessment should complement psychometric validation of translated questionnaires.
- Plain-language strategies may improve health equity and patient comprehension.

Introduction

Temporomandibular disorders (TMD) are a diverse group of conditions involving the masticatory muscles and temporomandibular joints, characterized by pain and functional limitation, and are increasingly recognized as a significant public health concern worldwide.[1,2]

The Diagnostic Criteria for Temporomandibular Disorders (DC/TMD), which evolved from the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD), offer a standardized, evidence-based dual-axis framework for both clinical practice and research. Axis I addresses physical diagnoses (e.g., pain-related and intra-articular disorders) with operational definitions and validated algorithms, while Axis II assesses pain-related disability, psychological distress, parafunctional behaviors, and jaw functional limitations through standardized self-report instruments.[3–5]

Axis II of the DC/TMD is intentionally modular and stepped, combining brief screeners with comprehensive measures to capture biobehavioral domains that influence pain and disability. The protocol retains validated screeners and adds instruments evaluating jaw function (e.g., the Jaw Functional Limitation Scale (JFLS)), psychological distress (Patient Health Questionnaire-4 (PHQ-4)), anxiety (Generalized Anxiety Disorder-7 (GAD-7)), depression (PHQ-9), somatic symptom burden (PHQ-15), and parafunctional behaviors (e.g., Oral Behaviors Checklist (OBC)), enabling stratified assessment from quick screening to an in-depth evaluation.[3] Given the increasing integration of temporomandibular disorder assessment into orthodontic diagnostics and treatment planning, the DC/TMD framework has become particularly relevant for orthodontists seeking standardized, multidimensional evaluation of patients presenting with pain, dysfunction or parafunctional habits. For instance, the Polish adaptation of the DC/TMD outlines the Axis I/II structure and includes officially translated instruments to support consistent implementation.[6]

Given the global use of the DC/TMD, rigorous translation and cultural adaptation procedures are essential to ensure both conceptual and linguistic equivalence. Published guidance from the RDC/TMD/INfORM (International Network for Orofacial Pain and Related Disorders Methodology) network describes forward and backward translation, expert panel review for cultural equivalence, bilingual pre-field testing, and field testing for readability and language congruency before adoption.[6] Empirically, these procedures have been applied, in full or adapted forms, to specific Axis II tools. For example, the Portuguese translation of the OBC demonstrated cross-cultural equivalence and acceptable psychometric properties following the INfORM guidelines,[7] and the Axis II instruments are increasingly used to characterize the psychosocial profiles of TMD patients.[8] The DC/TMD assessment instruments have also been formally introduced in national contexts, including Poland, to standardize clinical and research applications.[6] Beyond adult care, the INfORM recommendations provide child and adolescent adaptations, underscoring the breadth of validated use across populations.[9,10] The INfORM recommendations have also been made available in different languages, e.g., Polish.[11]

However, even when translations follow best practices and achieve psychometric validity, readability can vary meaningfully across languages and versions, with potential consequences for patient comprehension and data quality. Readability limitations are well documented across medical contexts: analyses of patient-facing web content in dentistry (including bruxism) routinely find materials written above recommended reading levels, contrary to the National Institutes of Health (NIH) guidance to target a 6th–7th grade level.[12] Similarly, a 20-year review of patient education materials in high-impact journals showed that most articles require high reading levels, limiting accessibility for average readers.[13] Outside dentistry, website and instrument readability frequently falls short of accepted thresholds, as shown for trauma-surgery content and other specialties.[14]

Crucially, regarding the question of whether validated translations maintain readability comparable to that of the original versions, evidence from questionnaire domains adjacent to orofacial pain suggests that translated, validated instruments can still impose substantial reading demands. The analysis of 19 validated pelvic-floor questionnaires translated into Spanish found that many instruments failed minimum readability thresholds on language-appropriate indices (Fernández-Huerta, Szigriszt-Pazos, INFLESZ, μ), implying elevated literacy demands despite formal validation.[15] These findings underscore the need for a systematic cross-language readability appraisal of the DC/TMD Axis II instruments, challenging the assumption that validation alone ensures equal linguistic accessibility.

The potential impact of readability disparities must be considered in the context of global language distribution and literacy. Large-scale network analyses indicate that a relatively small set of languages serves as hubs of global communication; rankings by native and non-native speakers highlight which translated versions are likely to reach the greatest numbers of patients and research participants.[16] This is relevant because DC/TMD-based screening is increasingly used in general dental practice,[17] and TMD symptoms and psychosocial profiles vary across ethnic and cultural groups.[18]

Recently compiled, open datasets on linguistic connectivity and multilingual communities further emphasize the breadth of cross-linguistic reach relevant to instrument dissemination.[19] At the same time, adult literacy skills vary substantially across countries, with marked differences between countries in reading proficiency that may influence who can realistically comprehend research and clinical questionnaires.[20]

Health services research also shows that language-concordant care improves understanding and outcomes, reinforcing that linguistic accessibility, including readable patient-reported outcome measures, is not merely a technical detail but a determinant of equity.[21] Conversely, the routine exclusion of non-English speakers from studies diminishes generalizability and perpetuates disparities, underscoring the need for validated, accessible translations.[22]

Taken together, the DC/TMD provide a robust, internationally endorsed diagnostic framework with a comprehensive Axis II battery. Translations have been validated across populations and languages (e.g., Portuguese OBC, national adoption in Poland and pediatric adaptations), yet prior work in cognate fields shows that readability can be compromised by translation, and even small textual shifts can raise the literacy threshold.[6,7,9,10,23,24]

Therefore, the present cross-sectional study aims to (1) quantify the readability of the original English DC/TMD Axis II questionnaires and (2) compare readability across a wide range of validated translations. By applying a uniform set of readability indices to all language versions, we aim to identify which validated translations remain linguistically onerous, despite psychometric adequacy, and to highlight opportunities for plain-language refinement that preserves validity while expanding access.

Material and methods

This descriptive cross-sectional study analyzed the readability of the DC/TMD Axis II questionnaires in their original English version and validated translations, using standardized readability indices applied at a single point in time. The methodological approach was consistent with that used in previous studies on readability in dentistry, particularly our prior study on online bruxism-related information, but also incorporated a standardization and within-questionnaire comparison framework to enable valid cross-language comparisons with the English version.

Selection of questionnaires

The study included selected DC/TMD Axis II instruments. These consist of standardized self-report questionnaires used to evaluate the psychological and behavioral components associated with TMD. Specifically, the instruments assessed were:

- Graded Chronic Pain Scale version 2.0 (GCPS 2.0);
- Jaw Functional Limitation Scale (JFLS; 20- and 8-item versions);
- Patient Health Questionnaire (PHQ; 4-, 9- and 15-item versions);
- Generalized Anxiety Disorder-7 (GAD-7); and
- Oral Behaviors Checklist (OBC).

All questionnaires were retrieved in their official, validated forms from the INfORM consortium website, which provides freely accessible, standardized diagnostic instruments. Each translation had undergone a rigorous process of forward and backward translation, expert panel review and field testing, as recommended by the INfORM guidelines, to ensure semantic and conceptual equivalence with the original English version. Only validated translations officially listed on the INfORM platform were included in the present analysis.

Languages included

To better represent the global distribution and use of the DC/TMD instruments, we compiled a multilingual dataset comprising the following target languages: English (reference language); Italian; Spanish; French; German; Brazilian Portuguese; and Turkish. Arabic, Chinese, Japanese, and Korean were excluded based on the methodological constraints of the readability indices used in the study. These formulas require alphabetic scripts, word spacing and syllabic segmentation. The abovementioned languages do not meet these requirements, as their orthographic and morphological systems (e.g., logographic characters, lack of fixed word boundaries and script-level variation) render the predominantly Anglo-centric indices invalid or unreliable.

Readability analysis

The primary methodological approach consisted of applying a fixed set of 7 widely used readability indices to each questionnaire in every language:

- Gunning Fog index;
- Coleman–Liau index (CLI);
- Flesch–Kincaid grade level (FKGL);
- automated readability index (ARI);
- simple measure of gobbledygook (SMOG);
- average grade level (as reported in source tables); and
- Flesch reading ease (FRE).

These formulas were selected because of their general utility in health communication research, their ability to capture different aspects of text complexity (e.g., sentence length, syllables per word and character counts), and their established application in studies of patient-oriented medical materials. Other indices, such as LIX and RIX readability formulas, which are considered more suitable for application to Germanic languages, were excluded because of the limited availability of medical research applying these indices to readability assessment.

Each metric captures a different aspect of textual complexity, including sentence length, letter and/or character density, syllables per word, and the frequency of polysyllabic words. To allow results from the various scales to be meaningfully compared, the FRE scores were inverted (–FRE).

All questionnaires were analyzed in their entirety. Texts were transcribed into plain-text format to avoid layout artifacts. No modifications to the original wording were made; the analysis was based on the validated wording of each questionnaire.

Data processing and standardization

To enable valid cross-language comparisons within each questionnaire, we implemented a two-step harmonization procedure:

1. Direction alignment: FRE score was inverted (–FRE) so that higher values indicated greater difficulty, consistent with the other indices;
2. Within-questionnaire and within-metric standardization: for each questionnaire–metric combination, values across the included languages were z-standardized (subtracting the cross-language mean and dividing by the cross-language standard deviation). This yielded z-scores, where 0 represented the average difficulty across languages for a given questionnaire and metric.

For each questionnaire–language combination, we then computed a composite standardized difficulty score as the mean of the available z-scores across the 7 indices (allowing for occasional missingness). This composite represented a standardized measure of relative readability burden, conceptualized in relation to the estimated educational level required to understand the information presented in the text. This measure could be readily calculated using the statistical software package.

Statistical analysis

All inferential comparisons were performed within each questionnaire, using English as the reference. For a given questionnaire, we compared each language's vector of metric-level z-scores against the corresponding English vector using a paired *t*-test across metrics (i.e., metric-by-metric pairing within the questionnaire, with a typical maximum of 7 pairs). This yielded a *p*-value comparing each language with English for each questionnaire. To account for multiple language comparisons, the Benjamini–Hochberg false discovery rate correction was applied to the resulting *p*-values.

Descriptive statistics (means and counts of available metrics) were computed for the composite standardized difficulty scores in each questionnaire–language cell. For visualization, we used violin plots of per-metric standardized differences (z_language – z_English) and forest plots of mean differences with 95% confidence intervals (95% *CI*s), both at the questionnaire level.

Ethical considerations

The research did not involve human participants and/or patient data and thus did not require approval from the institutional ethics committee. All analyzed data was derived from publicly available, validated instruments hosted on the INfORM consortium website.

Summary of methodological approach

In summary, we applied 7 established readability indices to the English DC/TMD Axis II questionnaires and validated translated versions of the DC/TMD Axis II questionnaires, standardized the results to enable within-questionnaire cross-language comparisons, and statistically tested differences relative to English using paired *t*-tests. This framework provided a standardized approach to the cross-language analysis of linguistic accessibility in the validated DC/TMD Axis II instruments.

Results

Overall descriptive findings

Regarding the DC/TMD Axis II questionnaires, readability indices demonstrated substantial heterogeneity across languages (Table 1). The English versions generally demonstrated the lowest educational requirements. Most validated translations required 1–6 additional years of education (range: –0.28–15.63 years), with Italian and Spanish versions frequently showing the highest reading difficulty ($p < 0.05$). Standardized differences relative to the English versions were positive for almost all translations (Fig. 1–8; Table 2). Italian and Spanish translations showed the largest standardized differences, followed by German and Brazilian Portuguese. Turkish demonstrated moderate increases in reading difficulty, whereas French showed the smallest increases. Differences tended to cluster by language rather than by instrument, underscoring the influence of linguistic structure (Table 2).

Table 1. Descriptive readability indices for the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) Axis II questionnaires across languages

Variable		English	French	German	Italian	Spanish	Brazilian Portuguese	Turkish
GCPS 2.0	characters, *n*	1,353	2,102	2,146	2,013	1,868	1,724	1,604
	words, *n*	382	448	388	443	436	418	335
	Gunning Fog index	9.20	10.40	13.00	14.00	13.90	13.80	10.80
	Coleman–Liau index	2.14	7.63	12.53	7.67	6.22	4.67	4.09
	Flesch–Kincaid grade level	2.93	6.55	6.79	9.38	11.49	7.38	4.65
	automated readability index	3.56	8.97	12.38	8.49	11.57	6.70	7.10
	simple measure of gobbledygook	6.99	8.57	10.07	11.13	12.53	11.15	8.07
	average grade level	5.62	9.74	11.16	11.53	11.87	9.91	8.07
	Flesch reading ease	104.00	78.00	74.00	59.00	59.00	74.00	84.00
JFLS-8	characters, *n*	548	878	950	961	748	729	817
	words, *n*	185	235	226	232	212	211	207
	Gunning Fog index	8.50	10.00	10.10	14.60	15.30	13.70	10.50
	Coleman–Liau index	0.00	2.85	5.21	5.35	2.69	1.41	1.71
	Flesch–Kincaid grade level	0.00	2.76	2.24	6.13	9.30	4.36	1.88
	automated readability index	1.77	4.56	5.90	6.37	10.33	4.43	4.06
	simple measure of gobbledygook	6.27	8.56	8.02	11.47	13.61	11.05	8.15
	average grade level	4.92	7.91	8.29	10.51	11.37	9.25	7.93
	Flesch reading ease	127.00	106.00	106.00	81.00	83.00	98.00	107.00
JFLS-20	characters, *n*	967	1,369	1,455	1,429	1,265	1,250	1,322
	words, *n*	382	436	429	432	416	418	403
	Gunning Fog index	7.50	8.80	9.20	12.00	11.50	11.80	9.50
	Coleman–Liau index	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Flesch–Kincaid grade level	0.00	0.00	0.00	2.01	1.73	1.35	0.00
	automated readability index	0.00	1.74	2.49	2.46	1.94	1.74	1.22
	simple measure of gobbledygook	4.76	7.06	6.39	9.49	9.33	9.26	7.32
	average grade level	4.18	5.09	5.15	7.95	7.77	7.74	5.29
	Flesch reading ease	144.00	127.00	129.00	111.00	115.00	118.00	125.00
PHQ-4	characters, *n*	702	887	726	680	680	645	616
	words, *n*	145	172	112	130	161	133	105
	Gunning Fog index	10.10	10.00	13.00	17.90	12.40	13.40	17.40
	Coleman–Liau index	9.90	10.07	16.10	12.37	5.15	8.86	12.58
	Flesch–Kincaid grade level	7.70	7.79	9.33	14.83	10.09	11.16	12.62
	automated readability index	9.43	11.46	13.77	14.04	9.96	10.91	13.70
	simple measure of gobbledygook	8.71	9.04	8.77	14.31	11.16	10.95	12.35
	average grade level	9.57	9.66	12.42	14.72	10.65	11.95	13.91
	Flesch reading ease	69.00	70.00	45.00	28.00	64.00	49.00	32.00

Variable		English	French	German	Italian	Spanish	Brazilian Portuguese	Turkish
PHQ-9	characters, *n*	1,102	1,290	873	1,304	1,158	1,080	1,316
	words, *n*	230	264	152	246	268	227	223
	Gunning Fog index	9.50	9.50	13.10	16.40	17.20	13.60	15.20
	Coleman–Liau index	9.62	8.55	13.93	11.89	5.94	8.43	11.71
	Flesch–Kincaid grade level	7.58	7.20	9.52	12.60	13.99	11.03	11.21
	automated readability index	8.80	9.83	12.53	10.77	13.81	9.71	12.23
	simple measure of gobbledygook	8.15	8.04	9.92	11.82	15.11	10.55	10.40
	average grade level	9.31	9.03	11.97	13.05	13.59	11.69	12.78
	Flesch reading ease	69.00	73.00	52.00	31.00	48.00	48.00	36.00
GAD-7	characters, *n*	766	1,004	842	1,018	852	775	680
	words, *n*	162	193	133	189	186	152	115
	Gunning Fog index	10.60	11.40	12.50	17.30	14.80	13.80	15.30
	Coleman–Liau index	8.98	10.20	15.04	12.53	7.01	10.03	12.30
	Flesch–Kincaid grade level	6.94	7.69	8.10	13.06	9.68	11.30	11.37
	automated readability index	7.59	10.49	12.82	11.21	9.44	10.18	12.17
	simple measure of gobbledygook	8.45	9.15	8.55	12.41	11.85	11.14	10.38
	average grade level	8.99	10.01	11.61	13.56	11.75	12.19	12.97
	Flesch reading ease	70.00	67.00	53.00	28.00	59.00	42.00	35.00
PHQ-15	characters, *n*	759	930	1,002	977	1,103	713	889
	words, *n*	150	171	143	180	245	146	162
	Gunning Fog index	8.00	9.70	13.10	15.40	13.40	12.00	10.80
	Coleman–Liau index	8.64	9.83	17.60	10.80	5.81	6.85	6.89
	Flesch–Kincaid grade level	4.17	4.86	8.56	10.30	7.90	6.83	6.17
	automated readability index	6.35	8.46	14.82	8.42	6.98	5.87	7.94
	simple measure of gobbledygook	6.55	7.24	7.83	9.59	10.62	8.05	7.09
	average grade level	7.43	8.54	12.71	11.82	10.50	9.14	8.88
	Flesch reading ease	80.00	76.00	46.00	37.00	65.00	62.00	64.00
OBC	characters, *n*	758	1,844	1,900	1,817	1,802	1,758	1,738
	words, *n*	173	379	334	356	451	361	296
	Gunning Fog index	6.00	9.60	15.70	23.80	13.70	20.50	19.70
	Coleman–Liau index	5.93	9.03	14.72	12.44	3.68	10.08	12.80
	Flesch–Kincaid grade level	3.31	7.19	15.15	29.53	7.66	20.16	17.31
	automated readability index	3.76	10.10	20.55	32.28	7.64	21.56	19.68
	simple measure of gobbledygook	5.07	8.11	12.56	23.28	11.70	17.96	16.17
	average grade level	5.15	8.93	14.96	20.78	10.01	16.34	16.44
	Flesch reading ease	88.00	75.00	41.00	0.00	77.00	22.00	19.00

GCPS 2.0 – Graded Chronic Pain Scale version 2.0; JFLS – Jaw Functional Limitation Scale; PHQ – Patient Health Questionnaire; GAD-7 – Generalized Anxiety Disorder-7; OBC – Oral Behaviors Checklist.
Data presented as raw index scores calculated on the full text of each questionnaire version. Higher scores on the grade-level indices (Gunning Fog index, Coleman–Liau index, Flesch–Kincaid grade level, automated readability index, simple measure of gobbledygook, and the average grade level) indicate greater reading difficulty, whereas higher Flesch reading ease scores indicate greater readability. No transformations or exclusions were applied at this stage; these values served as the unstandardized input data for subsequent analyses.

Table 2. Standardized differences in readability between the translated Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) Axis II questionnaires and their original English versions

Questionnaire	Language	Metrics compared, *n*	Mean z-score for the translated version	Mean z-score for the English version	*MD*	*t*-statistic	*p*-value	BH-adjusted *p*-value
GAD-7	Italian	7	1.141	−1.274	2.415	11.765	<0.001	<0.001
	Turkish	7	0.683	−1.274	1.957	9.405	<0.001	<0.001
	Brazilian Portuguese	7	0.238	−1.274	1.512	7.409	<0.001	0.001
	German	7	0.063	−1.274	1.337	3.425	0.014	0.021
	French	7	−0.695	−1.274	0.579	3.09	0.021	0.021
	Spanish	7	−0.154	−1.274	1.120	3.094	0.021	0.021
GCPS 2.0	Italian	7	0.681	−1.555	2.236	12.305	<0.001	<0.001
	Spanish	7	0.987	−1.555	2.542	10.88	<0.001	<0.001
	German	7	0.632	−1.555	2.187	8.608	<0.001	<0.001
	French	7	−0.176	−1.555	1.379	7.277	<0.001	<0.001
	Brazilian Portuguese	7	0.111	−1.555	1.666	7.653	<0.001	<0.001
	Turkish	7	−0.681	−1.555	0.874	7.381	<0.001	<0.001
JFLS-20	Italian	6	1.133	−1.479	2.612	20.504	<0.001	<0.001
	Spanish	6	0.839	−1.479	2.318	20.791	<0.001	<0.001
	Brazilian Portuguese	6	0.705	−1.479	2.184	13.494	<0.001	<0.001
	Turkish	6	−0.405	−1.479	1.074	4.172	0.009	0.013
	French	6	−0.446	−1.479	1.033	3.441	0.018	0.022
	German	6	−0.348	−1.479	1.131	2.808	0.038	0.038
JFLS-8	Italian	7	0.920	−1.418	2.338	14.07	<0.001	<0.001
	Spanish	7	1.285	−1.418	2.703	11.393	<0.001	<0.001
	French	7	−0.33	−1.418	1.088	8.738	<0.001	<0.001
	Brazilian Portuguese	7	0.136	−1.418	1.554	7.854	<0.001	<0.001
	Turkish	7	−0.484	−1.418	0.934	8.16	<0.001	<0.001
	German	7	−0.109	−1.418	1.309	4.767	0.003	0.003
OBC	Turkish	7	0.553	−1.302	1.856	20.341	<0.001	<0.001
	Italian	7	1.368	−1.302	2.671	15.13	<0.001	<0.001
	German	7	0.296	−1.302	1.598	12.028	<0.001	<0.001
	Brazilian Portuguese	7	0.572	−1.302	1.874	12.464	<0.001	<0.001
	French	7	−0.73	−1.302	0.572	10.237	<0.001	<0.001
	Spanish	7	−0.756	−1.302	0.546	2.394	0.054	0.054
PHQ-15	German	7	1.107	−1.044	2.150	8.788	<0.001	0.001
	French	7	−0.568	−1.044	0.475	6.749	0.001	0.002
	Italian	7	0.996	−1.044	2.040	5.27	0.002	0.004
	Turkish	7	−0.425	−1.044	0.619	3.034	0.023	0.034
	Spanish	7	0.214	−1.044	1.258	2.781	0.032	0.038
	Brazilian Portuguese	7	−0.280	−1.044	0.764	2.638	0.039	0.039
PHQ-4	Italian	7	1.304	−0.972	2.276	8.602	<0.001	<0.001
	Turkish	7	0.919	−0.972	1.891	9.473	<0.001	<0.001
	German	7	0.233	−0.972	1.206	4.282	0.005	0.008
	Brazilian Portuguese	7	−0.079	−0.972	0.893	4.257	0.005	0.008
	French	7	−0.794	−0.972	0.178	1.206	0.273	0.295
	Spanish	7	−0.610	−0.972	0.362	1.147	0.295	0.295
PHQ-9	Italian	7	0.701	−1.044	1.745	7.642	<0.001	0.001
	Turkish	7	0.521	−1.044	1.565	7.869	<0.001	0.001
	German	7	0.230	−1.044	1.274	7.073	<0.001	0.001
	Brazilian Portuguese	7	−0.125	−1.044	0.919	3.584	0.012	0.017
	Spanish	7	0.823	−1.044	1.866	3.262	0.017	0.021
	French	7	−1.106	−1.044	−0.063	−0.538	0.610	0.610

MD – mean difference; BH – Benjamini–Hochberg. The table presents the results of within-questionnaire comparisons between translated versions and the original English versions after direction alignment. The FRE values were reversed so that higher scores indicate greater estimated reading difficulty across all readability indices. Z-standardization was performed across languages for each questionnaire–readability index combination. Comparisons were conducted using paired *t*-tests across metrics. Positive *MD*s denote greater reading difficulty in the translated version relative to the original English version.

Questionnaire-specific descriptive statistics and results

Across all instruments, translations generally required more years of education to understand than the corresponding English versions. Italian and Spanish versions showed the greatest increases in reading difficulty (the largest standardized difference in 4 and 3 of the 8 questionnaires, respectively), followed by German and Brazilian Portuguese ($p < 0.05$ in most comparisons). The largest mean grade-level differences were observed for the OBC, GCPS 2.0 and JFLS-8, whereas the largest mean standardized differences were observed for the GCPS 2.0, JFLS-20 and JFLS-8; the PHQ-4, PHQ-9 and PHQ-15 showed the smallest standardized differences. In the following subsections, the values in parentheses represent average grade levels (Table 1), and Δ denotes the difference from the English version in years of education. Figures 1–8 present the corresponding metric-specific standardized differences.

GAD-7

The comprehension of the English version of the GAD-7 required an average of 8.99 years of education. European translations showed higher educational requirements: Italian (13.56, Δ +4.57); Turkish (12.97, Δ +3.98); Brazilian Portuguese (12.19, Δ +3.20); Spanish (11.75, Δ +2.76); German (11.61, Δ +2.62); and French (10.01, Δ +1.02) (Fig. 1).

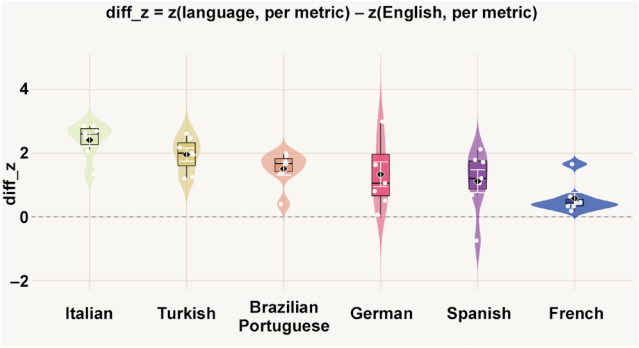

Fig. 1. Standardized readability differences relative to the original English version of the Generalized Anxiety Disorder-7 (GAD-7)

Violin plots show the distribution of metric-specific standardized readability differences between each translated version and the original English version. The readability indices were as follows: Gunning Fog index; Coleman–Liau index (CLI); Flesch–Kincaid grade level (FKGL); automated readability index (ARI); simple measure of gobbledygook (SMOG); average grade level; and Flesch reading ease (FRE). The sign of the FRE scores was reversed before z-standardization to ensure directional consistency across all indices. Boxes represent the interquartile range (*IQR*), with horizontal lines indicating the median (*Me*). Jittered points indicate individual metric-specific differences, and black diamonds with error bars indicate the mean ±standard error (*M* ±*SE*). The dashed horizontal line at 0 represents no difference from the English version; values above 0 indicate greater reading difficulty. Statistical comparisons with the original English version were performed using paired *t*-tests across readability indices for each language version, with Benjamini–Hochberg false discovery rate correction.

GCPS 2.0

With regard to the GCPS 2.0, the English version had an average grade level of 5.62, whereas all translated versions exhibited higher average grade levels: Spanish (11.87, Δ +6.25); Italian (11.53, Δ +5.91); German (11.16, Δ +5.54); Brazilian Portuguese (9.91, Δ +4.29); French (9.74, Δ +4.12); and Turkish (8.07, Δ +2.45) (Fig. 2).

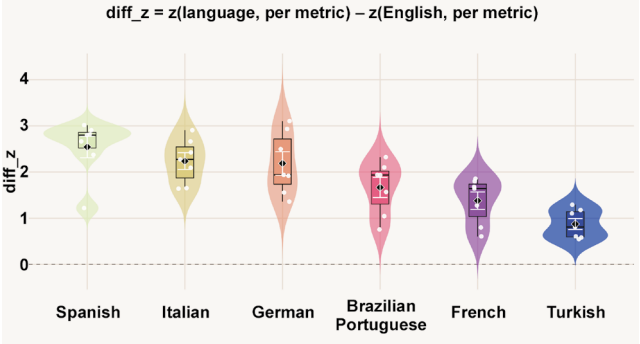

Fig. 2. Standardized readability differences relative to the original English version of the Graded Chronic Pain Scale, version 2.0 (GCPS 2.0)

Boxes represent the *IQR*, with horizontal lines indicating the *Me*. Jittered points represent individual metric-specific differences, and black diamonds with error bars indicate the *M* ±*SE*. The dashed horizontal line at 0 represents no difference from the English version; values above 0 indicate greater reading difficulty.

JFLS-20

The English version of the JFLS-20 was highly accessible, requiring 4.18 years of education. Moderate increases were observed for Italian (7.95, Δ +3.77), Spanish (7.77, Δ +3.59) and Brazilian Portuguese (7.74, Δ +3.56), whereas smaller increases were observed for Turkish (5.29, Δ +1.11), German (5.15, Δ +0.97) and French (5.09, Δ +0.91) (Fig. 3).

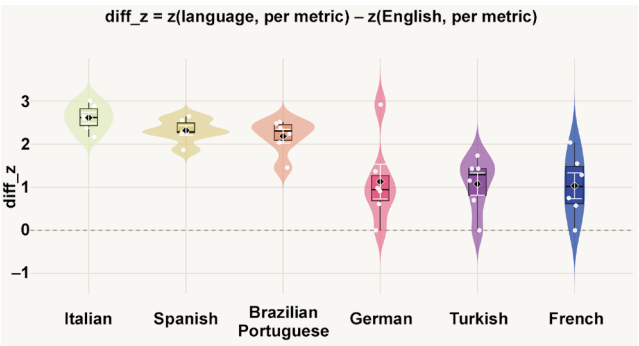

Fig. 3. Standardized readability differences relative to the original English version of the Jaw Functional Limitation Scale-20 (JFLS-20)

Boxes represent the *IQR*, with horizontal lines indicating the *Me*. Jittered points represent individual metric-specific differences, and black diamonds with error bars indicate the *M* ±*SE*. The dashed horizontal line at 0 represents no difference from the English version; values above 0 indicate greater reading difficulty.

JFLS-8

The comprehension of the English version of the JFLS-8 required 4.92 years of education. Higher educational requirements were observed for Spanish (11.37, Δ +6.45), Italian (10.51, Δ +5.59), Brazilian Portuguese (9.25, Δ +4.33), German (8.29, Δ +3.37), Turkish (7.93, Δ +3.01), and French (7.91, Δ +2.99) (Fig. 4).

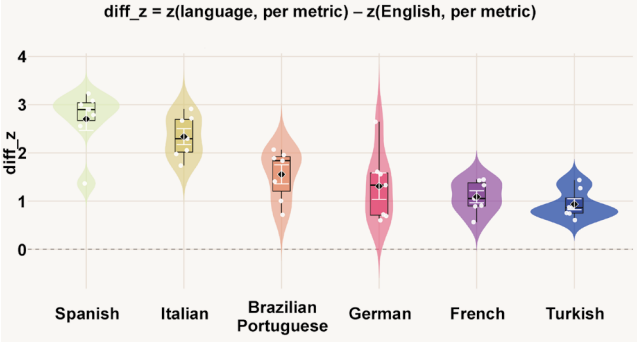

Fig. 4. Standardized readability differences relative to the original English version of the Jaw Functional Limitation Scale-8 (JFLS-8)

Boxes represent the *IQR*, with horizontal lines indicating the *Me*. Jittered points represent individual metric-specific differences, and black diamonds with error bars indicate the *M* ±*SE*. The dashed horizontal line at 0 represents no difference from the English version; values above 0 indicate greater reading difficulty.

PHQ-9

The English baseline for the PHQ-9 was 9.31 years of education. Most translations showed higher requirements: Spanish (13.59, Δ +4.28); Italian (13.05, Δ +3.74); Turkish (12.78, Δ +3.47); German (11.97, Δ +2.66); and Brazilian Portuguese (11.69, Δ +2.38). The French version was comparable to the English version (9.03, Δ –0.28) (Fig. 6).

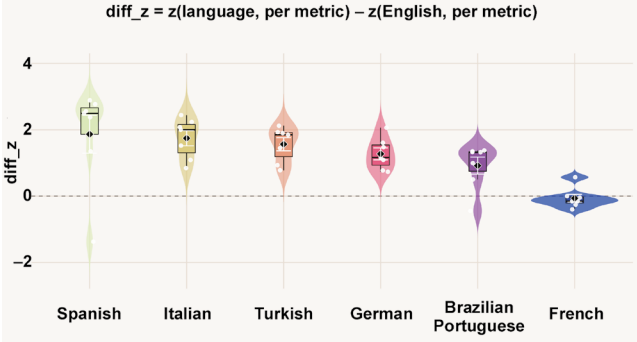

Fig. 6. Standardized readability differences relative to the original English version of the Patient Health Questionnaire-9 (PHQ-9)

Boxes represent the *IQR*, with horizontal lines indicating the *Me*. Jittered points represent individual metric-specific differences, and black diamonds with error bars indicate the *M* ±*SE*. The dashed horizontal line at 0 represents no difference from the English version; values above 0 indicate greater reading difficulty.

PHQ-15

With regard to the PHQ-15, the comprehension of the English version required an average grade level of 7.43 years. Translations generally showed higher educational requirements: German (12.71, Δ +5.28); Italian (11.82, Δ +4.39); Spanish (10.50, Δ +3.07); Brazilian Portuguese (9.14, Δ +1.71); Turkish (8.88, Δ +1.45); and French (8.54, Δ +1.11) (Fig. 5).

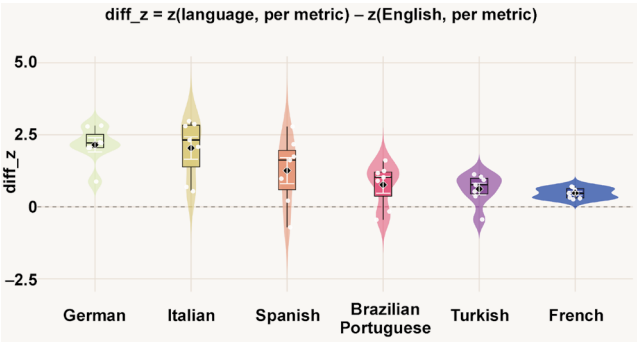

Fig. 5. Standardized readability differences relative to the original English version of the Patient Health Questionnaire-15 (PHQ-15)

Boxes represent the *IQR*, with horizontal lines indicating the *Me*. Jittered points represent individual metric-specific differences, and black diamonds with error bars indicate the *M* ±*SE*. The dashed horizontal line at 0 represents no difference from the English version; values above 0 indicate greater reading difficulty.

PHQ-4

The English version of the PHQ-4 required 9.57 years of education. Translations were more difficult to read: Italian (14.72, Δ +5.15); Turkish (13.91, Δ +4.34); German (12.42, Δ +2.85); Brazilian Portuguese (11.95, Δ +2.38); and Spanish (10.65, Δ +1.08). The French version was comparable to the English version (9.66, Δ +0.09) (Fig. 7).

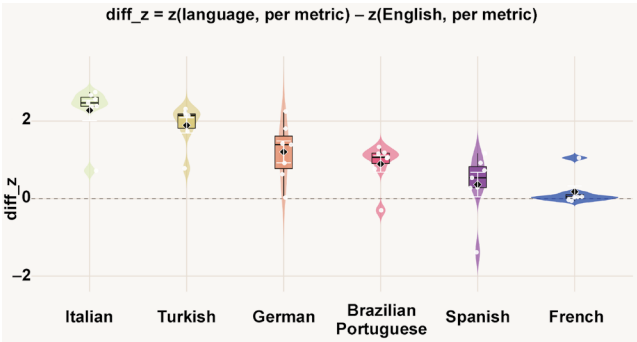

Fig. 7. Standardized readability differences relative to the original English version of the Patient Health Questionnaire-4 (PHQ-4)

Boxes represent the *IQR*, with horizontal lines indicating the *Me*. Jittered points represent individual metric-specific differences, and black diamonds with error bars indicate the *M* ±*SE*. The dashed horizontal line at 0 represents no difference from the English version; values above 0 indicate greater reading difficulty.

OBC

The English version required 5.15 years of education. Higher educational requirements were observed for Italian (20.78, Δ +15.63), Turkish (16.44, Δ +11.29), Brazilian Portuguese (16.34, Δ +11.19), German (14.96, Δ +9.81), Spanish (10.01, Δ +4.86), and French (8.93, Δ +3.78) (Fig. 8).

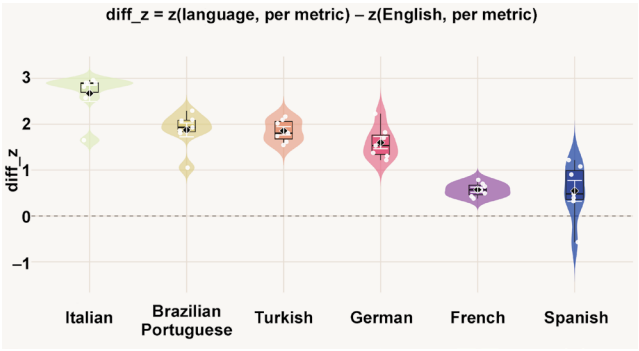

Fig. 8. Standardized readability differences relative to the original English version of the Oral Behaviors Checklist (OBC)

Boxes represent the *IQR*, with horizontal lines indicating the *Me*. Jittered points represent individual metric-specific differences, and black diamonds with error bars indicate the *M* ±*SE*. The dashed horizontal line at 0 represents no difference from the English version; values above 0 indicate greater reading difficulty.

Comparative analysis

Across all instruments and included languages, the majority of translations required more years of education than the English versions. The mean difference in readability was approx. +4.0 years of education (median: +3.6 years), with statistically significant increases observed in 91.7% of language–instrument comparisons after the Benjamini–Hochberg correction (adjusted $p < 0.05$). At the questionnaire level, the largest mean grade-level differences were observed for the OBC, GCPS 2.0 and JFLS-8, whereas the largest mean standardized differences were observed for the GCPS 2.0, JFLS-20 and JFLS-8. Detailed standardized differences between each translation and the corresponding English version, including statistically significant values, are summarized in Table 2.

Summary interpretation

The DC/TMD Axis II translations were consistently more difficult to read than their English counterparts. Figures 1–8 provide visual summaries of these cross-language differences.

Discussion

The DC/TMD provides a dual-axis framework in which Axis II comprises validated self-report instruments to capture psychosocial and functional dimensions of TMD, making their linguistic accessibility integral to valid assessment.[3] The recent INfORM recommendations emphasize structured adaptation and the clinical usability of the DC/TMD instruments across age groups, reinforcing the expectation that translations should preserve both psychometric properties and practical comprehensibility.[9,10] This concern is reflected in studies analyzing patient-facing digital resources, which have shown that platforms like YouTube often provide content with variable quality and readability, potentially affecting patient understanding and trust.[25,26] Similarly, evidence from dental patient-facing materials shows that readability frequently exceeds recommended levels for the general public, indicating a systemic risk that can also affect patient-reported instruments.[12] Minervini et al. documented that online information in orofacial pain contexts is commonly written above the 6th–7th grade comprehension level often recommended for health materials, highlighting the need for systematic readability assessment of the DC/TMD Axis II questionnaires.[12] Our comparative results, showing that many validated translations are more difficult to read than the original English versions, are consistent with the broader observation that readability can deteriorate during translation, even when formal validation is successful.[15] This concern aligns with the INfORM's emphasis on careful cross-cultural linguistic adaptation, which implicitly includes attention to language clarity and patient comprehension as essential components of successful implementation.[10]

Cross-language differences in readability are plausibly linked to the structural properties of languages, such as morphology, orthography and typical sentence patterns, which can influence the distribution of word and sentence lengths captured by standard readability indices.[27] Koplenig's cross-linguistic analysis suggests that language structure covaries with speaker communities in ways that could systematically affect the textual features measured by readability, providing a theoretical basis for the heterogeneity observed among translations.[27] Because the DC/TMD instruments are disseminated worldwide, the practical impact of readability differences is proportional to the size and connectivity of the language communities that use them.[16] Network analyses of the world's languages indicate that a relatively small set of hubs mediates large fractions of global communication. Consequently, increases in readability demands in widely used languages could affect disproportionate numbers of patients and research participants.[16] Multilingual datasets allow researchers to map how readability gaps affect instrument reach and equity.[19]

Readability is only one aspect of the access equation; the other is population literacy, which varies substantially across nations and influences who can realistically understand patient questionnaires at given grade levels.[20] Large cross-national assessments show wide variation in adult reading proficiency, suggesting that a three- to

four-grade increase in the education level required to understand a translated questionnaire can materially reduce the share of adults capable of self-administering that instrument without assistance.[20] From a health services perspective, language-concordant communication improves comprehension and learning outcomes, underscoring that translation quality must encompass readability to avoid undermining equitable care.[21] When instruments remain linguistically dense, the risk of excluding patients with lower literacy levels or those completing questionnaires in a second language is increased, potentially introducing bias into the collected data and widening disparities despite formal validation.[21]

INfORM's guidance and the core DC/TMD documentation were designed to support standardized, valid assessments across settings. Our findings suggest that an explicit plain-language checkpoint could be added to translation workflows without compromising psychometrics.[3] The recommendations for adolescents further illustrate how tailoring instruments to the reading capabilities of specific populations is part of best practice, a principle that can be extended to adult translations by targeting realistic grade levels for the intended setting.[10] Several practical strategies for improving the readability of the translated DC/TMD questionnaires while preserving construct validity can be identified from the existing literature: the use of shorter sentences; reduced use of multiple subordinate clauses; replacement of low-frequency polysyllabic terms with common synonyms; and comparison of post-editing readability scores with predefined thresholds.[12] Incorporating these steps into forward and backward translation and expert panel review would align with the INfORM's structured adaptation pathway and allow readability assessment to be documented alongside psychometric testing.[10] At the same time, readability indices have limitations: they are proxies based on surface features and do not directly measure comprehension, numeracy or cultural interpretation; therefore, triangulation with user testing remains essential.[12] Given the established comorbidity between TMD and conditions such as obstructive sleep apnea, as observed in patients using mandibular advancement devices,[28] readable assessment instruments are particularly important for accurate symptom reporting in multidisciplinary contexts.

Nevertheless, because these indices correlate with comprehension burden at the population level, they offer a practical screening tool to identify versions that likely exceed the literacy levels attainable by populations in their target countries.[20] Languages with broad transnational reach should be prioritized for plain-language refinement to maximize the impact.[16] Open linguistic-connectivity resources can further support scenario analyses, for example, estimating how many patients would potentially gain access to an instrument if a translation's required grade level were reduced by 1 or 2 years.[19] Clinically, clearer questionnaires may improve self-administration rates and reduce the need for clinician mediation, which in turn may enhance data quality for Axis II domains such as distress, disability and jaw function.[3] Educational interventions that strengthen language-concordant communication also suggest that simplifying patient instruments could yield measurable gains in understanding and engagement.[21] Standardizing scores across languages enabled the identification of consistent difficulty patterns that clinicians can easily compare.[12] Future work should combine readability screening with cognitive interviewing in each language to verify whether simplified wording preserves the constructs specified in the DC/TMD, thereby balancing accessibility and validity.[10] By situating our cross-language comparisons within the global distribution of languages and literacy, these findings support a practical agenda for plain-language optimization of the DC/TMD translations within the INfORM framework, with the goal of expanding equitable access to accurate psychosocial assessment in TMD care.[16]

Conclusions

This cross-sectional analysis demonstrates that validated translations of the DC/TMD Axis II questionnaires often require higher reading levels than the original English versions, generally corresponding to approx. 4 additional years of schooling. The greatest standardized increases occurred in the jaw-related instruments (GCPS 2.0, JFLS-20 and JFLS-8), while the PHQ screeners (PHQ-4, PHQ-9 and PHQ-15) remained closer to the readability levels of the English versions. Readability patterns clustered by language rather than instrument, indicating that linguistic structure plays a key role. These differences are clinically meaningful. In settings with lower literacy, higher educational requirements may limit self-completion and affect data quality, triage decisions and equitable access. Accordingly, readability should be evaluated alongside psychometric properties during questionnaire translation and cultural adaptation. Incorporating plain-language strategies into the INfORM framework and reporting readability for each language version may help optimize accessibility, particularly in widely spoken languages. Future research should combine readability estimates with cognitive interviewing and user testing to ensure that improved clarity preserves construct validity and supports accurate, patient-centered assessment across diverse linguistic contexts.

Ethics approval and consent to participate

Not applicable.

Data availability

All analyzed questionnaires were retrieved from the INfORM (International Network for Orofacial Pain and Related Disorders Methodology) consortium website (https://inform-iadr.com) and are publicly accessible.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

During the preparation of this manuscript, the authors used DeepL Translate (DeepL SE, Cologne, Germany) for language checking and grammar correction. Following its use, the authors thoroughly reviewed, verified and revised all content to ensure accuracy and originality.

ORCID iDs

Giuseppe Minervini https://orcid.org/0000-0002-8309-1272
Hande Uzunçıbuk https://orcid.org/0000-0001-9265-1772
Nicola Fiorito https://orcid.org/0000-0002-9452-9698
Giulia Maria Nicoletti https://orcid.org/0000-0001-8801-8456
Diana Russo https://orcid.org/0000-0002-0915-4642
Emilio Alfredo Cafferata https://orcid.org/0000-0001-9243-1382
Gabriele Cervino https://orcid.org/0000-0003-4619-4691
Maria Maddalena Marrapodi https://orcid.org/0000-0002-9494-6942

References

1. Rajaraman V, Nallaswamy D, Shenoy A. Artificial intelligence and machine learning in diagnosing and managing temporomandibular disorders: A systematic review and meta-analysis. *J Oral Biol Craniofac Res*. 2025;15(6):1591–1600. doi:10.1016/j.jobcr.2025.09.013
2. Kaynak BA, Taş S, Tüfek MT, Dağ F, Dikici R. Fatty infiltration and morphology of cervical muscles in patients with temporomandibular disorders: A case–control study. *Dent Med Probl*. 2024;61(4):525–532. doi:10.17219/dmp/166266
3. Schiffman E, Ohrbach R, Truelove E, et al.; International RDC/TMD Consortium Network, International association for Dental Research; Orofacial Pain Special Interest Group, International Association for the Study of Pain. 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(1):6–27. doi:10.11607/jop.1151
4. Yap AU, Lei J, Liu C, Huang Z, Fu KY. Evaluating temporomandibular disorder severity: A four-dimensional model. *J Oral Pathol Med*. 2025;54(10):982–991. doi:10.1111/jop.70050
5. Ohrbach R, Dworkin SF. The evolution of TMD diagnosis: Past, present, future. *J Dent Res*. 2016;95(10):1093–1101. doi:10.1177/0022034516653922
6. Osiewicz M, Ciapała B, Bolt K, Kołodziej P, Więckiewicz M, Ohrbach R. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD): Polish assessment instruments. *Dent Med Probl*. 2024;61(1):5–8. doi:10.17219/dmp/181149
7. Barbosa C, Manso MC, Reis T, Soares T, Gavinha S, Ohrbach R. Cultural equivalence, reliability and utility of the Portuguese version of the Oral Behaviours Checklist. *J Oral Rehabil*. 2018;45(12):924–931. doi:10.1111/joor.12716
8. Yu Y, Wang L, Chen C, Wang M, Nie S, Fu L. Correlations of oral behaviors, pain, and psychological distress among patients with temporomandibular disorders: Clinical investigation of axial II evaluation. *Front Public Health*. 2025;13:1604995. doi:10.3389/fpubh.2025.1604995
9. Nilsson IM, Ekberg EC, Michelotti A, et al.; International Network for Orofacial Pain and Related Disorders (INfORM). Diagnostic Criteria for Temporomandibular Disorders – INfORM recommendations: Comprehensive and short-form adaptations for children. *J Oral Rehabil*. 2023;50(2):99–112. doi:10.1111/joor.13390
10. Ekberg EC, Nilsson IM, Michelotti A, et al.; International Network for Orofacial Pain and Related Disorders Methodology (INfORM). Diagnostic Criteria for Temporomandibular Disorders–INfORM recommendations: Comprehensive and short-form adaptations for adolescents. *J Oral Rehabil*. 2023;50(11):1167–1180. doi:10.1111/joor.13488
11. Nitecka-Buchta A, Baron S, Walczyńska-Dragon K, et al. Temporomandibular disorders: INfORM/IADR key points for good clinical practice based on standard of care: The Polish language version. *Dent Med Probl*. 2025;62(4):565–567. doi:10.17219/dmp/208377
12. Minervini G, Marrapodi MM, Cicciù M. Online bruxism-related information: Can people understand what they read? A cross-sectional study. *J Oral Rehabil*. 2023;50(11):1211–1216. doi:10.1111/joor.13519
13. Rooney MK, Santiago G, Perni S, et al. Readability of patient education materials from high-impact medical journals: A 20-year analysis. *J Patient Exp*. 2021;8:2374373521998847. doi:10.1177/2374373521998847
14. Eltorai AEM, Ghanian S, Adams CA, Born CT, Daniels AH. Readability of patient education materials on the American Association for Surgery of Trauma website. *Arch Trauma Res*. 2014;3(2):e18161. doi:10.5812/atr.18161
15. Checa-Moreno V, Díaz-Mohedo E, Suárez-Serrano C. Analysis of the readability of questionnaires on symptoms of pelvic floor dysfunctions adapted to Spanish. *Int J Environ Res Public Health*. 2021;18(19):10320. doi:10.3390/ijerph181910320
16. Ronen S, Gonçalves B, Hu KZ, Vespignani A, Pinker S, Hidalgo CA. Links that speak: The global language network and its association with global fame. *Proc Natl Acad Sci U S A*. 2014;111(52):E5616–E5622. doi:10.1073/pnas.1410931111
17. Emodi-Perlman A, Eli I. Temporomandibular disorders and bruxism – up-to-date assessment and screening tools the general dentist should be aware of. *Dent Med Probl*. 2024;61(2):169–171. doi:10.17219/dmp/175582
18. Yap AU, Chan SW, Alias AA, et al. Ethnic variations in temporomandibular disorder symptoms, oral behaviours, personality traits, and psychosocial profiles among young adults: Insights from a comparative analysis. *J Oral Rehabil*. 2026;53(1):116–126. doi:10.1111/joor.70067
19. Gurevich T, Herman PR, Toubal F, Yotov YV. A dataset on linguistic connectivity across and within countries. *Sci Data*. 2025;12(1):542. doi:10.1038/s41597-025-04692-8
20. Kyröläinen AJ, Kuperman V. Predictors of literacy in adulthood: Evidence from 33 countries. *PLoS One*. 2021;16(3):e0243763. doi:10.1371/journal.pone.0243763
21. Molina RL, Kasper J. The power of language-concordant care: A call to action for medical schools. *BMC Med Educ*. 2019;19(1):378. doi:10.1186/s12909-019-1807-4
22. Bahji A, Acion L, Laslett AM, Adinoff B. Exclusion of the non-English-speaking world from the scientific literature: Recommendations for change for addiction journals and publishers. *Nordisk Alkohol Nark*. 2023;40(1):6–13. doi:10.1177/14550725221102227
23. Gałczyńska-Rusin M, Pobudek-Radzikowska M, Czajka-Jakubowska A. Polish language adaptation and validation of the Fonseca Anamnestic Index for individuals with temporomandibular disorders. *Dent Med Probl*. 2024;61(5):705–711. doi:10.17219/dmp/177287
24. da Cunha TA, Chaves TC, Pereira Júnior FJ, de Godoi Gonçalves DA, Alstergren P, Biasotto-Gonzalez DA. Brazilian Portuguese version of the Diagnostic Criteria for Temporomandibular Disorders Axis II: Translation, cross-cultural adaptation and measurement properties. *J Oral Rehabil*. 2025;52(5):712–721. doi:10.1111/joor.13921
25. Di Stasio D, Romano A, Paparella RS, et al. How social media meet patients questions: YouTube review for mouth sores in children. *J Biol Regul Homeost Agents*. 2018;32(2 Suppl. 1):117–121. PMID:29460528.
26. Tok ÖN, Yurttaş AN, Taş S. YouTube videos as an information source about exercises for temporomandibular disorders. *Dent Med Probl*. 2024;61(6):855–863. doi:10.17219/dmp/170922
27. Koplenig A. Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. *R Soc Open Sci*. 2019;6(2):181274. doi:10.1098/rsos.181274
28. Langaliya A, Alam MK, Hegde U, Panakaje MS, Cervino G, Minervini G. Occurrence of temporomandibular disorders among patients undergoing treatment for obstructive sleep apnoea syndrome (OSAS) using mandibular advancement device (MAD): A systematic review conducted according to PRISMA guidelines and the Cochrane Handbook for Systematic Reviews of Interventions. *J Oral Rehabil*. 2023;50(12):1554–1563. doi:10.1111/joor.13574