Received: 14 May 2023 | Revised: 18 July 2023 | Accepted: 18 August 2023

DOI: 10.1111/odi.14729

ORAL DISEASES Leading in Oral, Maxillofacial, Head & Neck Medicine WILEY

ORIGINAL ARTICLE

# Validity and internal consistency of four scales in patients with TMD: PHQ8, GAD7, PHQ15 and JFLS20

Dina Taimeh[1,2] | Richeal Ni Riordain[1,3] | Stefano Fedele[1,4,5] | Rachel Leeson[1]

[1]Department of Maxillofacial Medicine and Surgery, UCL Eastman Dental Institute, University College London, London, UK

[2]Department of Oral and Maxillofacial Surgery, Oral Medicine and Periodontology, Faculty of Dentistry, University of Jordan, Amman, Jordan

[3]Oral Medicine Unit, Cork University Dental School and Hospital, Cork, Ireland

[4]Oral Medicine Unit, Eastman Dental Hospital, University College London Hospitals Trust, London, UK

[5]NIHR University College London Hospitals Biomedical Research Centre, London, UK

**Correspondence**
Dina Taimeh, UCL Eastman Dental Institute, University College London, Rockefeller Building, 20 University Street, London WC1E 6DE, UK and School of Dentistry, and University of Jordan, Queen Rania Street, 11972 Amman, Jordan.
Email: dina.taimeh.17@ucl.ac.uk

**Funding information**
University of Jordan

## Abstract

**Objectives:** The aim of this cross-sectional study was to explore the structural validity and internal consistency reliability of General Anxiety Disorder-7, Patient Health Questionnaire-8, 15 and Jaw Functional Limitation Scale-20 in patients with chronic pain of temporomandibular disorders.

**Materials and Methods:** Validity and reliability were assessed in 129 patients diagnosed according to the diagnostic criteria for temporomandibular disorders. Structural validity was explored using factor analysis, and internal consistency by calculating Cronbach $\alpha$.

**Results:** Confirmatory factor analysis revealed a suitable 2-factor model for Patient Health Questionnaire-8, with Cronbach $\alpha$ of 0.89, and 0.86. One and 2-factor models were suitable for General Anxiety Disorder-7, with overall Cronbach $\alpha$ of 0.93 for the 1-factor model, and 0.91 and 0.84 for both factors in a 2-factor model. A 4-factor solution was appropriate for Patient Health Questionnaire-15, with Cronbach $\alpha$ of 0.72, 0.57, 0.71 and 0.73 for each factor separately. Exploratory factor analysis was conducted to explore the factor structure of Jaw Functional Limitation Scale 20, and a 3-factor solution was appropriate.

**Conclusions:** This study provides positive evidence of structural validity and internal consistency of these questionnaires in patients with pain of temporomandibular disorders. However, additional testing is required to explore further psychometric properties.

KEYWORDS
patient-reported outcome measures, reliability, temporomandibular disorders, validity

## 1 | INTRODUCTION

The term 'Temporomandibular disorders' (TMD) is an umbrella term which encompasses a range of conditions affecting the temporomandibular joint, surrounding musculature or both (Leeuw, 2013). They could manifest in an array of symptoms, including pain, limitation of mouth opening, deviation in mandibular movements, joint sounds and otalgia (Durham, 2013). The prevalence of this group of conditions is reported to be from 5% to 50% of the general population, and is estimated third among chronic pain conditions after common headaches and backache (Dworkin, 2011). The Diagnostic Criteria for TMD (DC/TMD) classifies TMD broadly into pain-related TMD and headache, intra-articular joint disorders and degenerative joint disorder (Schiffman et al., 2014). Being a chronic pain condition, TMD is theorised to fall under the biopsychosocial model of pain, which depicts pain as a reciprocating interplay among physiological, psychological and social factors, eventually giving rise to complex pain syndromes such as fibromyalgia, temporomandibular disorders

and back pain (Edwards et al., 2016). Research has also supported a bidirectional link between mood disorders and enduring pain. As pain persists, a higher risk of an effective disorder is expected, and similarly, psychosocial variables such as distress and anxiety are among the most robust predictors of the transition from acute to chronic pain (Edwards et al., 2016). A good deal of empirical evidence supports this argument and several reviews have highlighted the importance of such elements in shaping pain-related experiences and associated treatment outcomes (Pincus et al., 2002; Vissers et al., 2012). However, in practice, psychological components are often viewed as a reaction to the pain and are assigned secondary status (Edwards et al., 2016). Evidence also suggests that chronic orofacial pain is not only a local phenomenon, but an organismic response (Dworkin, 2011), where central sensitisation (CS) is involved. The International Association for the Study of Pain (IASP) defines central sensitisation as 'Increased responsiveness of nociceptive neurons in the central nervous system to their normal or subthreshold afferent input' (IASP, 2011). Chronic pain conditions which are considered to fall within the spectrum of CS include fibromyalgia, irritable bowel syndrome and TMD. These represent conditions with overlapping clinical and pathophysiological features where central factors may play a role in their aetiology (Harte et al., 2018).

PROMs are 'standardised, validated questionnaires that are completed by the patients to ascertain perceptions of their health status, perceived level of impairment, disability, and health-related quality of life' (Kingsley & Patel, 2017). Health measurement instruments have many applications in clinical research and practice, based on which, treatment decisions may be made. Hence, it is important that these instruments are well-designed and psychometrically sound. This is usually demonstrated by providing sufficient evidence of a rigorous development process and satisfactory psychometric properties such as validity and reliability (Mokkink et al., 2010a, 2010b). One aspect of validity is structural validity, which is defined as 'the degree to which the scores of a health-related PRO instrument are an adequate reflection of the dimensionality of the construct to be measured'. As for internal consistency reliability, it is 'the degree of the interrelatedness among the items' (Mokkink et al., 2010a, 2010b). PROMs are often tested in a specific population to study the acceptability of their behaviour and performance in that population, as it cannot be assumed that they will perform well across all cohorts. For example, an instrument which measures depression or health-related quality of life may require modifications to its factor structure or standard cut-off points (Dyer et al., 2016). The present authors could not locate any articles which tested the structural validity of General Anxiety Disorder-7 (GAD7), Patient Health Questionnaire-8 (PHQ8) and Patient Health Questionnaire-15 (PHQ15) in a TMD population.

Thus, the aim of this study was to explore the structural validity and internal consistency reliability of GAD7, PHQ8, PHQ15 and Jaw Functional Limitation Scale-20 (JLS20) in patients with chronic pain related to TMD, to investigate pain levels, functional limitation and psychological co-morbidities in this cohort of patients, and to study the associations between these variables.

## 2 | MATERIALS AND METHODS

### 2.1 | Study design

This was a questionnaire-based study with a cross-sectional design. It received ethical approval from the Southeast Scotland Research Ethics Committee 1 (REC reference: 19/SS/0130) and was conducted in accordance with the Declaration of Helsinki. The eligible participants were recruited from Oral Surgery and Facial Pain clinics at UCLH Eastman Dental Hospital from March to September 2021. The research activity was conducted at a single time point alongside standard of care after the conclusion of the participants' routine clinical visits. They were informed about the study orally and provided with a Patient Information Sheet (PIS). Those who showed interest were invited to sign an informed consent form and complete the study questionnaires.

### 2.2 | Participants

Participant recruitment was based upon convenience sampling. Patients were eligible to take part if they were over the age of 18 and had a diagnosis of TMD with associated pain according to the DC/TMD criteria (Schiffman et al., 2014). Patients having at least one of the following diagnoses were eligible: myalgia (local myalgia, myofascial pain, myofascial pain with referral), arthralgia and headache attributed to TMD. Additionally, a good command of the English language, and an ability to give informed consent were required. Eligible patients were invited to take part whether they were new or follow-up patients at the clinical centre. They were excluded if they had a recent history of surgery or trauma to the head and neck region, if they had intra-articular complaints with no associated pain, poor command of the English language or inability to give informed consent.

The sample size was determined according to the guidance of the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) initiative risk of bias checklist, where an adequate sample when testing the structural validity through factor analysis would be at least five times the number of items and ≥100; or at least six times the number of items but <100 (Mokkink et al., 2018). The longest tool in this study was the JFLS20, containing 20 items. Therefore, a minimum of 100 participants were needed.

### 2.3 | Procedure and outcomes

A comprehensive clinical exam was carried out to confirm the diagnosis of TMD and classify the participants according to the DC/TMD criteria. This popular diagnostic criteria consists of two axes; axis I offers a classification system obtained from a thorough history and an examination checklist, and axis II contains a set of PROMs to record several parameters, including pain intensity, jaw function, psychological status and psychosocial function (Schiffman et al., 2014).

The participants were then asked to complete a set of paper-based questionnaires including a demographics form collecting the age of the participants, sex, ethnicity, smoking status (never, previous, current smoker), alcohol consumption (Y/N) and the medical co-morbidities including other systemic chronic pain conditions. Additionally, the participants completed the GCPS version 2.0, PHQ8, GAD7, PHQ15 and JFLS20. These five scales compose a part of the second axis of the DC/TMD criteria.

The study was reported according to the COSMIN reporting guidelines for studies on measurement properties of PROMs (Gagnier et al., 2021).

## 2.4 | Patient-reported outcome measures

*Graded Chronic Pain Scale* (*GCPS*) *version 2.0* is a composite score which takes into account the Characteristic Pain Intensity (CPI) and Interference Score (IS). Two versions exist, 6-month and 1-month versions. The 1-month version was chosen as the axis 1 of the DC/TMD criteria is based on a 1-month reference, therefore, this version would be consistent with the classification system used.

*Patient Health Questionnaire 8* (*PHQ8*) is a measure of depressive disorders in the general population (Kroenke et al., 2009). The scores range from 0 to 24, with values of 5, 10, 15 and 20 representing cut-off points for mild, moderate, moderately severe and severe depression respectively. A 10-point cut-off score was used in the second part of this study to identify positive cases of clinical depression, hence categorising the scores into a binary variable (<10 and ≥10) (Feingold et al., 2017). This scale has been tested repeatedly in various populations, with several alternative factor structures suggested, including 1-factor, 2-factor and bifactor models. (Chilcot et al., 2013; Granillo, 2012; Krause et al., 2011).

*General Anxiety Disorder 7* (*GAD7*): This tool is a simple and short questionnaire which was developed to increase recognition of general anxiety disorder in primary care (Spitzer et al., 2006). Scores range from 0 to 21, and values of 5, 10 and 15 represent cut-off points for mild, moderate and severe anxiety. Similar binary categorisation was applied to the scores with a 10-point cut-off value to identify positive cases of general anxiety disorder (Feingold et al., 2017). Akin to PHQ-8, GAD-7 has more than one suggested factor solution found in the literature.

*Patient Health Questionnaire-15* (*PHQ-15*): A brief and self-reported PROM to screen and monitor somatisation and somatic symptom severity (Kroenke et al., 2002). Several studies have examined the psychometric properties of the scale, contributing to support its validity and reliability in various settings (Han et al., 2009; Hyphantis et al., 2014). Scores range from 0 to 30, with scores of 5, 10 and 15 represent cut-off points for low, medium and high physical symptoms respectively. Similar to the previous two scales, a 10-point cut-off score was applied to establish patients with medically unexplained physical symptoms (North et al., 2019).

*Jaw Functional Limitation Scale-20* (*JFLS-20*): This instrument assess disability related to facial pain and covers a range of activities such as eating foods of various consistencies, yawning and laughing in patients with TMD (Ohrbach, Granger, et al., 2008). Two articles were identified exploring the factor structure of this instrument in the literature, and both showed a three-domain solution (Fetai et al., 2020; Xu et al., 2020). A single global score of 'jaw functional limitation' can be computed as the mean of the available items or by computing the mean of the three subscale scores (*Mastication*: mean of items 1–6, *Mobility*: mean of items 7–10 and *Verbal and non-verbal communication*: mean of items 13–20).

## 2.5 | Data analysis

The statistical analysis was performed using STATA version 17 (StataCorp). Descriptive analysis was first conducted to summarise the demographics and outcome scores of the sample. The normality of distribution of the data was also tested using the Shapiro–Wilk test.

The structural validity of PHQ15, PHQ8, GAD7 and JFLS20 was explored in this study by conducting confirmatory factor analysis (CFA). Where the factor solutions proposed in the literature did not suit any of the PROMs, exploratory factor analysis (EFA) was conducted. The following fit indices were used to judge the suitability of the proposed models in CFA: root mean square of error approximation (RMSEA), standardised root mean squared residual (SRMR), comparative fit index (CFI) and Tucker-Lewis index (TLI). RMSEA and SRMR values <0.5 are indicative of a good fit, and values <0.8 of an acceptable fit. CFI and TLI values greater than 0.95 are considered acceptable (Hu & Bentler, 1999). Satorra–Bentler correction was applied to account for non-normality in the data, as some of these indices are affected by it, especially in small to medium sample sizes (Frazier et al., 2018; Nima et al., 2020; Satorra & Bentler, 2001).

In the cases were EFA was needed, several steps were applied starting with evaluating whether the data are suitable for this type of analysis. A pair of tests were used to that end; the Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity. The KMO index ranges from 0 to 1, with values higher than 0.50 considered suitable for factor analysis. The Bartlett's test of sphericity should be significant, indicating sufficient intercorrelations for factor analysis. (Hair et al., 1995; Tabachnick & Fidell, 2007). Factor extraction was subsequently carried out using principal factor analysis, to reduce the large number of items into factors. Two rules were utilised; the Kaiser's criteria (eigenvalue >1 rule) (Kaiser, 1960) and the scree test (Cattell, 1966). The scree plot is a heuristic graph that plots the eigenvalues against the components. By inspecting the elbow of the plot—the point where the notable decline in factors levels off—the number of retained factors could be estimated (Ledesma et al., 2015).

The initially extracted loadings are usually not particularly interpretable because the items may load on multiple factors. Therefore, factor rotation was applied next. It is a mathematical transformation with the aim of obtaining an interpretable factor loading matrix that provides a simple structure solution (Finch, 2020). Factor rotation was done using Promax rotation method with a 0.45 cut-off point

for factor loadings. Items were excluded if they had weak loadings on factors (<0.45), or if they cross-loaded on more than one factor.

Cronbach $\alpha$ was used to measure the internal consistency reliability, with a minimum cut-off point of 0.7 as an acceptable value (Terwee et al., 2007).

The associations between the various PROMs and corresponding demographic data were explored by running bivariable analysis using Mann–Whitney rank sum test, Kruskal–Wallis test and Spearman's correlation for continuous variables, and chi-square and Fisher's exact tests for categorical variables. All tests were two-tailed and $p$ <0.05 were considered statistically significant.

Missing values were dealt with by mean imputation, where the mean of the rest of the items in the scale for each participant was used to estimate the missing value. The integer mean value was used.

## 3 | RESULTS

One hundred and twenty-nine participants took part in this cross-sectional study. The mean age was 39.8 years (median = 37) and ranged from 18 to 74. 82.17% ($n$ = 106) were females and 17.83% ($n$ = 23) were males. The mean duration of onset of symptoms was 7.5 years (±7.9 SD) with a range of 0.05–38 years. See Table 1 for a full description of the participants' details.

In relation to missing data, 0.3%, 0.3%, 1% and 1.8% of the data sets were missing for PHQ8, GAD7, PHQ15 and JFLS20 respectively. Mean imputation was used to replace the missing values. No participants were excluded as a result of missing data.

**TABLE 1** Descriptive statistics of the demographic variables of the study participants.

| | |
|---|---|
| Gender | *N* (%) |
| Female | 106 (82.14%) |
| Male | 23 (17.83%) |
| Ethnicity | |
| White | 84 (65.1%) |
| Asian | 27 (20.9%) |
| Black | 9 (6.98%) |
| Mixed | 5 (3.88%) |
| Other | 4 (3.10%) |
| Smoking status | |
| Never smoked | 84 (65.1%) |
| Previous smoker | 31 (24%) |
| Current smoker | 14 (10.9%) |
| Alcohol consumption | |
| Non-drinker | 69 (49.64%) |
| Drinker | 70 (50.36%) |
| TMD symptoms | |
| Pain only symptoms | 65 (50.4%) |
| Pain+ intraarticular involvement | 64 (49.6%) |

### 3.1 | PHQ8

Several factor solutions were described in the literature (Krause et al., 2011; Lamela et al., 2020; Pagán-Torres et al., 2020). Four of those were tested for this scale to identify a suitable fit in a TMD sample; 1-factor, 2-factor and bifactor models. See the supplementary material for a full description of the models tested and the CFA fit indices for each of the models. The most suitable solution was a 2-factor model. This model suggested a 2-factor solution, with items 1, 2 and 6 comprising a cognitive component, and items 3, 4, 5, 7 and 8 comprising a somatic component. Figure 1 describes this model, and Table 2 details the associated CFA fit indices. Cronbach $\alpha$ was subsequently calculated based on this model for all the items grouped together and for the factors separately. Both values were good and fell above the acceptable cut-off point of 0.7. See Table 2.

### 3.2 | GAD7

Akin to the other scales, GAD-7 was also tested for its factor structure several times in different populations (Johnson et al., 2019; Terrill et al., 2015). 1 and 2-factor solutions (f1: emotional and cognitive, f2: physical) were tested to assess the best fit model. The fit indices for both solutions were suggestive of a good fit. Figure 1 displays the tested models, and Table 2 details the resultant CFA fit indices.

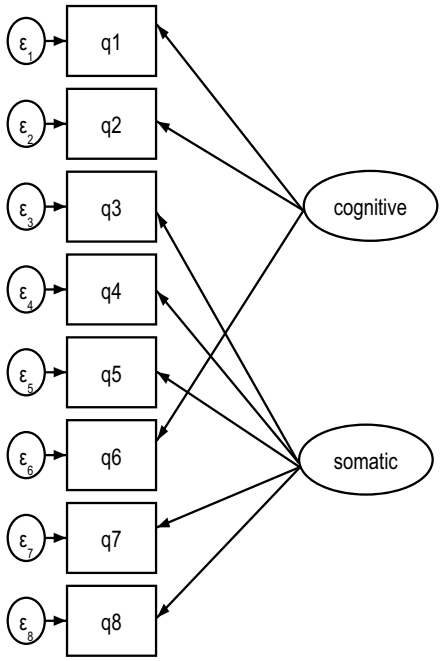
Cronbach $\alpha$ was calculated next and gave good results for both solutions as well. All alpha values were good and above 0.7. See Table 2.

### 3.3 | PHQ15

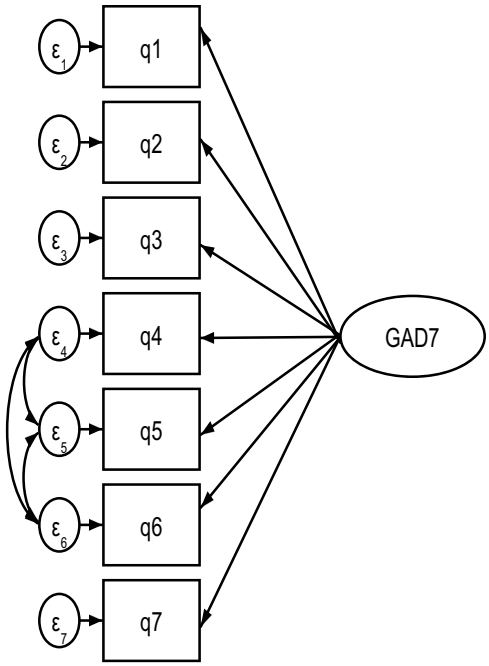
Studies exploring the factor structure of PHQ-15 were reported abundantly in the literature (Cano-García et al., 2020; Leonhart et al., 2018; Liao et al., 2016). Hence, multiple models were tested to find the best fit. 1-, 3-, 4- and bifactor models were tested. The suggested factors were gastro-intestinal, pain, fatigue and cardiopulmonary. The supplementary material describes the tested models and the fit indices for each solution.

The 1 and 3-factor solutions yielded the poorest fit. The 4-factor and bifactor solutions gave a good fit to the data. However, a couple of points were noted; item 4 was omitted from most articles as it contained gender-specific content (menstrual problems). Item 8 (fainting spells) was also omitted from the model suggested by Cano-García et al, as it had a very low base rate in their sample. This model gave very good results in the current sample, whether item 8 was included or excluded.

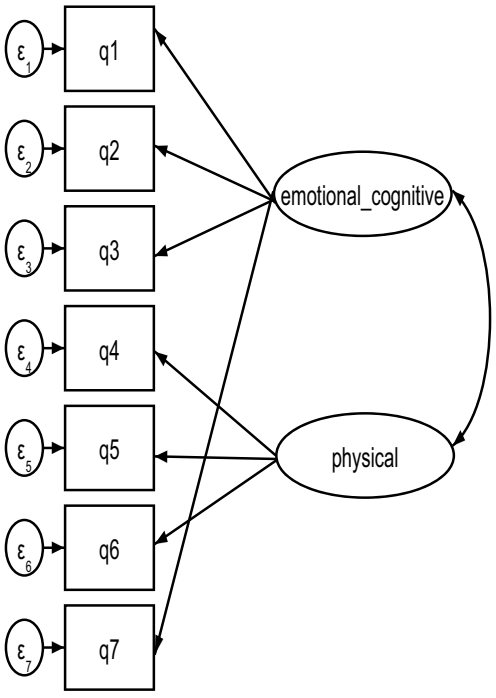
Cronbach $\alpha$ was calculated for the 4-factor and bifactor models proposed in the previous section. Based on the CFA and Cronbach $\alpha$ results, the 4-factor solution gave the most satisfactory results. See Figure 1 for a description of this model, and Table 2 for the associated CFA fit indices and Cronbach $\alpha$.



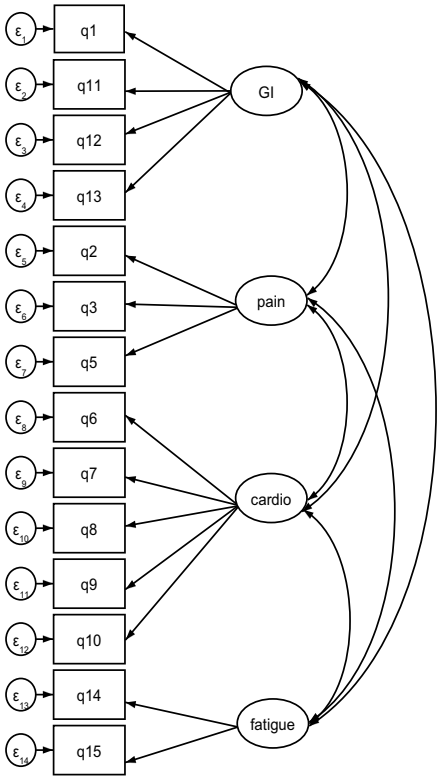

2-factor model for PHQ8




1-factor model for GAD7




2-factor model for GAD7




4-factor model for PHQ15

**FIGURE 1** Best fit models tested for the study PROMs.

### 3.4 | JFLS20

The original development articles used Rasch analysis to assess the relevance of the items. (Ohrbach, Granger, et al., 2008; Ohrbach, Larsson, et al., 2008). Two further studies explored the factor structure of JFLS-20, both describing a 3-factor solution (Fetai et al., 2020; Xu et al., 2020). The suggested models were tested in this sample of patients with TMD; however, both did not

**TABLE 2** CFA fit indices for PHQ8, GAD7 and PHQ15.

| | |
|---|---|
| PHQ8 | |
| Fit indices for the 2-factor model | |
| RMSEA-SB | 0.67 |
| CFI-SB | 0.980 |
| TLI-SB | 0.970 |
| SRMR | 0.044 |
| Cronbach $\alpha$ all items | 0.91 |
| Cronbach $\alpha$ for factor 1 (items 1, 2, 6) | 0.89 |
| Cronbach $\alpha$ for factor 2 (items 3, 4, 5, 7, 8) | 0.86 |
| GAD7 | |
| Fit indices for the 1-factor model | |
| RMSEA-SB | 0.069 |
| CFI-SB | 0.989 |
| TLI-SB | 0.979 |
| SRMR | 0.027 |
| Cronbach $\alpha$ all items | 0.93 |
| Fit indices for the 2-factor model | |
| RMSEA-SB | 0.067 |
| CFI-SB | 0.988 |
| TLI-SB | 0.980 |
| SRMR | 0.033 |
| Cronbach $\alpha$ for factor 1 (emotional-cognitive) | 0.91 |
| Cronbach $\alpha$ for factor 2 (physical) | 0.84 |
| PHQ15 | |
| Fit indices for the 4-factor model | |
| RMSEA-SB | 0.042 |
| CFI-SB | 0.960 |
| TLI-SB | 0.949 |
| SRMR | 0.059 |
| Cronbach $\alpha$ all items | 0.84 |
| Cronbach $\alpha$ factor 1 (GI): items 1, 11, 12, 13 | 0.72 |
| Cronbach $\alpha$ factor 2 (pain): items 2, 3, 5 | 0.57 |
| Cronbach $\alpha$ factor 3 (cardio): items 6, 7, 8, 9, 10 | 0.71 |
| Cronbach $\alpha$ factor 4 (fatigue): items 14, 15 | 0.73 |

Abbreviation: SB, Satorra–Bentler correction.

yield a good fit. See the Data S1. Considering the lack of abundant literature that describes the factor structure for this scale, exploratory factor analysis (EFA) was deemed appropriate in this case.

The Bartlett test of sphericity and the Kaiser–Meyer–Olkin Measure of Sampling Adequacy both gave favourable results; with the former giving a $p$ <0.05, and the latter giving a value of 0.914. Next, all the items were inserted in the analysis equation. Promax factor rotation with 0.45 cut-off value was chosen. Three factors were retained as determined by the Kaiser's test (number of eigenvalues >1) and by plotting a scree graph. See Figure 2 for the scree graph. The first domain contained the items 1–4, 7 and 12, the second domain the items 5, 6, 9–11 and the third domain the items 13–20. See

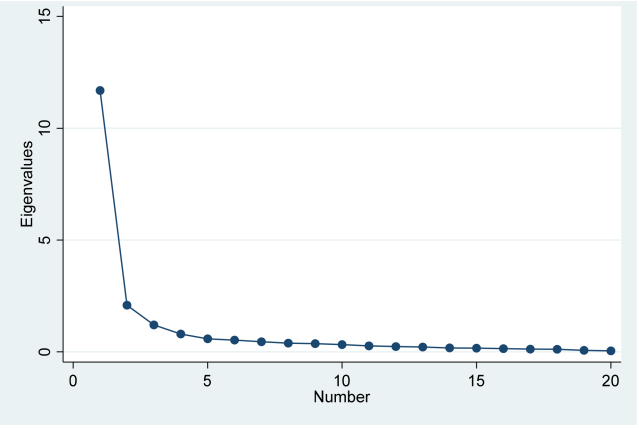



**FIGURE 2** Scree plot of retained factors during exploratory factor analysis of JFLS20.

Table 3 for the STATA output of the rotated factors containing the corresponding items. No items had weak loadings or cross loaded on more than one factor; therefore, none were deleted.

Cronbach $\alpha$ for the overall score was 0.9605. As for the individual domains, alpha for factors 1, 2 and 3 were 0.9421, 0.8940 and 0.9422 respectively.

### 3.5 | Descriptive statistics and cross-sectional analysis of the associations between pain, functional limitation, anxiety, depression and somatisation

A 10-point cut off score was chosen to identify positive cases of anxiety and depression on the GAD7 and PHQ8. However, looking at the original classification system for these two PROMs, 35.66% had minimal anxiety, 27.91% had mild anxiety, 18.60% had moderate anxiety and 17.83% had severe anxiety. As for depression, 29.46% had no depression, 18.60% had mild depression, 12.40% had moderate depression, 12.40% had moderately severe depression and 8.53% had severe depression. Table 4 describes the mean, median, skewness, kurtosis and normality of the data for the different scales used. Most of the varibales were not normally distributed, hence, non-parametric tests were used.

The relationships between the demographic variables and the different scales were studied by running the following tests: Chi square, Fisher's exact test, Kruskal–Wallis test, Mann–Whitney rank sum test and Spearman's correlation as detailed in Table 5. Most of the results gave non-significant $p$-values, apart from the relationship of CPI scores with gender and the relationship of GCPS (total, CPI and IS), JFLS and PHQ-15 with the smoking status. A post hoc Dunn test with Bonferroni adjustment was applied next to identify the responsible pair. Generally, the mean rank on the current smokers was higher than the other two groups, as detailed in Table 5.

Further analysis was also conducted to study the relationships between the different study PROMs. Table 6 details the conducted tests and the associated p-values.

## 4 | DISCUSSION

This study aimed to explore the structural layout and internal consistency of four common scales in a TMD population. GAD-7, PHQ-8 and PHQ-15 have been validated in various samples, such as pregnancy (Soto-Balbuena et al., 2021) and atypical chest pain (Lin et al., 2021). All amounting to abundant literature describing their factor structure. In such a case, confirmatory factor analysis (CFA) was appropriate to apply (Kim & Mueller, 1978). CFA is a method used to compare the measures of construct in a current sample to a hypothesised/suggested model in previous analytical research (Kline, 2011). Several models were identified for these three scales in our search, and alternative solutions were tested for each. A 2-factor model with cognitive and somatic factors was suitable for PHQ-8, and both models proposed for GAD-7 (1 and 2-factor models) delivered good results in terms of CFA indices and Cronbach $\alpha$ values. Perhaps with a larger sample size in future research, one model could edge the other in terms of robustness in a TMD population. As for PHQ-15, seven models were tested, with 4-factor and bifactor models producing good fit indices. Cronbach $\alpha$ values were the determining factor in this study, as a 4-factor solution consisting of GI, pain, cardiopulmonary and

**TABLE 3** Factor loadings of JFLS20 items during factor analysis using Promax rotation and 0.45 cutoff point.

| Variable | Factor 1 | Factor 2 | Factor 3 | Uniqueness |
|---|---|---|---|---|
| Jflsq19 | 0.9616 | 0.0307 | -0.1522 | 0.2060 |
| Jflsq15 | 0.9296 | 0.0168 | -0.0539 | 0.1785 |
| Jflsq20 | 0.7887 | 0.2463 | -0.0912 | 0.1945 |
| Jflsq16 | 0.7834 | -0.0940 | 0.2335 | 0.2005 |
| Jflsq17 | 0.7328 | -0.1172 | 0.2348 | 0.3054 |
| Jflsq13 | 0.7231 | 0.0199 | 0.1933 | 0.2421 |
| Jflsq18 | 0.6644 | 0.0122 | 0.1904 | 0.3506 |
| Jflsq14 | 0.4664 | 0.2099 | 0.1915 | 0.4338 |
| Jflsq1 | -0.0726 | 1.0019 | -0.0486 | 0.1186 |
| Jflsq2 | 0.0255 | 0.9441 | 0.0817 | 0.1535 |
| Jflsq7 | 0.0597 | 0.9096 | -0.0984 | 0.1933 |
| Jflsq8 | 0.1039 | 0.8001 | 0.0239 | 0.2286 |
| Jflsq4 | -0.0287 | 0.7427 | 0.2431 | 0.2404 |
| Jflsq3 | -0.0816 | 0.6649 | 0.3725 | 0.2649 |
| Jflsq12 | 0.4065 | 0.5299 | -0.0772 | 0.3762 |
| Jflsq6 | -0.0311 | -0.1225 | 0.9918 | 0.1571 |
| Jflsq11 | 0.0246 | -0.0219 | 0.7897 | 0.3689 |
| Jflsq5 | -0.0157 | 0.1944 | 0.7528 | 0.2662 |
| Jflsq10 | 0.3307 | 0.0064 | 0.6234 | 0.2353 |
| Jflsq9 | 0.3246 | 0.1784 | 0.4646 | 0.3052 |

**TABLE 4** Descriptive statistics, response distribution, skewness, kurtosis, and normality of distribution for study PROMs.

| | *N* (%) | Mean (SD) | Median (Range) | Skewness | Kurtosis | S-W (*p*-value) |
|---|---|---|---|---|---|---|
| GCPS-total | 129 (100%) | 2.5 (1.2) | 2 (1–4) | 0.09 | 1.5 | 0.89 |
| Grade 1,2 | 69 (53.5%) | | | | | |
| Grade 3,4 | 60 (46.5%) | | | | | |
| GCPS-CPI | 129 (100%) | 56.0 (22.4) | 60 (3.3–96.7) | −0.44 | 2.3 | **0.002*** |
| GCPS-IS | 129 (100%) | 36.9 (30.1) | 10 | 0.46 | 2.2 | **0.005*** |
| GAD-7 | 129 (100%) | 8.1 (6.4) | 7 (0–21) | 0.45 | 2.0 | **0.0001*** |
| <10 | 82 (63.6%) | | | | | |
| ≥10 | 47 (36.4%) | | | | | |
| PHQ-8 | 129 (100%) | 8.95 (6.6) | 8 (0–24) | 0.48 | 2.3 | **0.0002*** |
| <10 | 78 (60.5%) | | | | | |
| ≥10 | 51 (39.5%) | | | | | |
| PHQ-15 | 129 (100%) | 9.58 (5.7) | 9 (0–24) | 0.51 | 2.5 | **0.002*** |
| <10 | 74 (57.4%) | | | | | |
| ≥10 | 55 (42.6%) | | | | | |
| JFLS-20 | 129 (100%) | 3.1 (2.3) | 3.14 (0–8.5) | 0.38 | 2.3 | **0.000*** |

Abbreviations: CPI, characteristic pain intensity; IS, interference score.

*Bold values represent *p*-value <0.5 indicating statistical significance.

**TABLE 5** Descriptive statistics and *p*-values of the relationships between study PROMs and patient characteristics.

| | *N* (%) | GCPS-total | CPI | IS | JFLS | GAD-7 | PHQ-8 | PHQ-15 |
|---|---|---|---|---|---|---|---|---|
| Ethnicity | | 0.34[a] | 0.13[b] | 0.28[b] | 0.21[b] | 0.25[a] | 0.27[a] | 0.36[a] |
| White | 84 (65.1%) | | | | | | | |
| Asian | 27 (20.9%) | | | | | | | |
| Black | 9 (6.98%) | | | | | | | |
| Mixed | 5 (3.88%) | | | | | | | |
| Other | 4 (3.10%) | | | | | | | |
| Gender | | 0.213[c] | **0.035**[d,*] | 0.143[d] | 0.34[d] | 0.51[c] | 0.607[c] | 0.19[c] |
| Female | 106 (82.2%) | | | | | | | |
| Male | 12 (17.8%) | | | | | | | |
| Smoking status | | **0.032**[c,*] | **0.007**[b,*] | **0.0007**[b,*] | **0.029**[b,*] | 0.228[c] | 0.123[c] | **0.008**[c,*] |
| Never[†] | 84 (65.1%) | | | | | | | |
| Previous | 31 (24%) | **0.04**[‡] | 0.057[¥] | **0.0016**[‡] | 0.07[¥] | | | 0.058[¥] |
| Current | 14 (10.9%) | 0.10[¥] | **0.009**[‡] | **0.017**[‡] | **0.048**[‡] | | | **0.011**[‡] |
| Alcohol | | 0.35[c] | 0.18[d] | 0.16[d] | 0.22[d] | 0.08[c] | 0.08[c] | 0.998[c] |
| No | 61 (47.3%) | | | | | | | |
| Yes | 68 (52.7%) | | | | | | | |
| DC-TMD | | 0.32[c] | 0.34[d] | 0.07[d] | 0.064[d] | 0.19[c] | 0.047[c] | 0.88[c] |
| Myalgia | 62 (48.1%) | | | | | | | |
| Myalgia + IA | 67 (51.9%) | | | | | | | |
| Age | | 0.64[d] | 0.79[e] | 0.76[e] | 0.78[e] | 0.14[d] | 0.40[d] | 0.88[d] |

[a]Fisher's exact test.
[b]Kruskal–Wallis test.
[c]Chi square test.
[d]Mann–Whitney rank sum test.
[e]Spearman's correlation.
*Bold values represent *p*-value <0.5 indicating statistical significance.
[†]Reference group.
[¥]*p*-value with the reference group after post hoc Dunn test with Bonferroni adjustment.
[‡]Significant *p*-value with the reference group after post hoc Dunn test with Bonferroni adjustment.

**TABLE 6** *p*-values resulting for analysis of the relationships between study PROMs.

| | GAD-7 | PHQ-8 | PHQ-15 | JFLS-20 |
|---|---|---|---|---|
| GCPS-total | **0.003**[a,*] | **0.001**[a,*] | **0.008**[a,*] | **0.0001**[b,*] |
| GCPS-CPI | **0.0269**[b,*] | **0.0001**[b,*] | **0.0004**[b,*] | **0.0001**[c,*] ($r$ = 0.52, CI: 0.381–0.642) |
| GCPS-IS | **0.0003**[b,*] | **0.0007**[b,*] | **0.0004**[b,*] | **0.0001**[c,*] ($r$ = 0.59, CI: 0.467–0.699) |
| GAD-7 | – | **0.000**[a,*] | **0.001**[a,*] | **0.0026**[b,*] |
| PHQ-8 | – | – | **0.001**[a,*] | **0.0003**[b,*] |
| PHQ-15 | – | – | – | **0.0116**[b,*] |

[a]Chi square test.
[b]Mann–Whitney rank sum test.
[c]Spearman's correlation.
*Bold values represent *p*-value <0.5 indicating statistical significance.

fatigue domains gave the best internal consistency results. Hietaharju et al reported Cronbach $\alpha$ values of 0.85 for PHQ-9, 0.81 for PHQ-15 and 0.91 for GAD-7 in their study comparing the tools of RDC/TMD to the updated version; DC/TMD (Hietaharju et al., 2021). The results of this study were in keeping with these reported values, indicating that indeed, these three scales have good internal consistency reliability in patients with chronic pain related to TMD.

EFA was applied to explore the factor structure of JFLS20 in this study, indicating a suitable 3-factor model with very good internal consistency. Studies exploring some psychometric properties of this scale reported Cronbach $\alpha$ values >0.8 (Fetai et al., 2020; Ohrbach, Granger, et al., 2008; Ohrbach, Larsson, et al., 2008; Xu et al., 2020). Again, supporting the results in this study which demonstrated good internal consistency of the scale. Xu et al., also reported good test–retest reliability as measured by interclass correlation coefficient for the domains of the scale (>0.85), providing additional evidence to support its use in patients with TMD.

GAD-7 and PHQ-8/9 are often used together to measure anxiety and depression respectively (Heindl et al., 2021; Reddy et al., 2021). Previous investigations revealed high levels of psychological disorders in patients with TMD ranging from 21.4% to 60.1% for moderate–severe depression (Canales et al., 2018) and around 30% for moderate–severe anxiety (Bertoli & de Leeuw, 2016; Simoen et al., 2020). In this study, over a third of the participants were positive for anxiety and depression (36.4% and 39.5% respectively). Additionally, participants with higher CPI scores, had higher anxiety, depression and somatisation scores as well.

Somatisation is 'the association of medically unexplained somatic symptoms with psychological distress and health-seeking behaviour' (Kroenke et al., 2002). The triad of anxiety, depression and somatisation seems to constitute the most common psychological problems encountered in primary care (Kroenke et al., 2002). Having reliable measures to recognise somatic symptoms is therefore important. The DC/TMD initiative includes PHQ-15 as a measure of the severity of somatic symptoms, due to the emerging evidence of its importance in the overall symptom reporting in individuals with TMD (Fillingim et al., 2011; Schiffman et al., 2014). Somatic awareness was shown to be elevated among patients with chronic pain and indeed chronic TMD (Macfarlane et al., 2002; Manfredini et al., 2010). A recent systematic review by Canales et al. (2018) reported prevalence between 28.5% and 76.6% for moderate–severe somatisation in patients with TMD. Most of the articles in the review, measured somatisation using Symptoms Checklist 90 (SCL-90), an instrument recommended by the original RDC/TMD. Replaced currently by PHQ-15 in the updated criteria (DC/TMD), a shift may occur in future TMD studies to reflect these alterations.

Anxiety, depression and somatisation are discussed frequently in association with chronic pain. Manfredini et al reported a strong correlation between pain-related disability (GCPS categories) and both depression and somatisation in a multi-centre study of patients with TMD (Manfredini et al., 2010). A high correlation coefficient (0.73) was also reported by Yap et al between depression and somatisation, concluding that a considerable section of clinically depressed patients with TMD describe frequent non-specific physical symptoms such as chest pain or GI problems (Yap et al., 2002). While some studies suggested a less significant role for anxiety in chronic myofascial pain (Giannakopoulos et al., 2010; Reiter et al., 2015), others noted a high correlation between anxiety and depression in patients with TMD (Simoen et al., 2020) and indicated a statistically significant rise in anxiety when compared to non-TMD patients (Resende et al., 2020; Simoen et al., 2020).

The results of this study reiterate the relevance of the psychological profile of patients with TMD and the importance of such assessment before treatment. The psychological co-morbidities are viewed as elements of the biopsychosocial model of pain, which TMD is theorised to fall under (Hampf, 1990; Suvinen et al., 2005). Therefore, it is essential to have reliable and valid instruments able to give trustworthy results, based on which treatments can be planned and support can be arranged.

No gender or ethnic differences were detected in this study in relation to anxiety, depression and jaw functional limitation. Interestingly, statistically significant results were obtained with PHQ-15, GCPS and JFLS scores when looking at the smoking status of the participants. While smoking does not offer pain relief, participants in a recent qualitative study exploring chronic pain, described it as a coping strategy (Lee et al., 2021). It serves as a cognitive distraction from the pain, resorting to it to 'calm them down'. Other studies, also report that smokers describe greater pain intensity and greater pain-related functional interference (Volkman et al., 2015; Weingarten et al., 2009), which is in keeping with the results of this study.

It is worthwhile for future research to explore the rest of the psychometric properties of these scales in a larger sample size, to provide further evidence of their suitability in a population with TMD. A longitudinal design also offers the possibility of testing other measurement properties, such as test–retest reliability and responsiveness.

### 4.1 | Limitations

All the participants in this study had myofascial TMD with approximately half having intraarticular involvement as well. Taking this into consideration, the results may not be generalisable to all types of TMD, such as those presenting with functional limitations without pain.

The study was cross-sectional in nature without long-term follow-ups. Therefore, some psychometric properties could not be explored such as responsiveness and test–retest reliability.

Other limitations were related to the sample size used. Factor analysis in general is a technique for large samples (Kyriazos, 2018). The definition of a large and sufficient sample is still however debatable. The sample used in this study was adequate according to the recommendations of the COSMIN initiative.

## 5 | CONCLUSION

Pervious literature has shown that the assessment of psychosocial functioning is an essential part of the diagnostic process of TMDs, and a useful prompt to a multidisciplinary approach to treatment (Schiffman et al., 2014). Having reliable and validated instruments can facilitate communication regarding consultations, management

and prognosis. The results from this study provide positive psychometric evidence in terms of structural validity and internal consistency for the use of PHQ8, GAD7, PHQ15 and JFLS20 in patients with chronic pain related to TMD. Future research with a longitudinal design and a larger sample size is needed to explore other psychometric properties in a TMD population such as test–retest reliability, responsiveness and construct validity.

## AUTHOR CONTRIBUTIONS

**Dina Taimeh:** Conceptualization; investigation; writing – original draft; methodology; writing – review and editing; formal analysis; data curation; funding acquisition; project administration; software. **Richeal Ni Riordain:** Conceptualization; methodology; supervision; investigation; project administration; validation; writing – review and editing. **Stefano Fedele:** Conceptualization; methodology; validation; writing – review and editing; project administration; supervision. **Rachel Leeson:** Conceptualization; investigation; methodology; validation; writing – review and editing; project administration; supervision.

## ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by the University of Jordan, in the form of a PhD student sponsorship awarded to DT. The authors declare no conflict of interest in relation to this piece of research.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

*Dina Taimeh* https://orcid.org/0000-0001-9907-0647
*Stefano Fedele* https://orcid.org/0000-0001-9006-9412

## REFERENCES

Bertoli, E., & de Leeuw, R. (2016). Prevalence of suicidal ideation, depression, and anxiety in chronic temporomandibular disorder patients. *Journal of Oral & Facial Pain and Headache*, *30*(4), 296–301. https://doi.org/10.11607/ofph.1675

Canales, G. D., Camara-Souza, M. B., Lora, V., Guarda-Nardini, L., Conti, P. C. R., Garcia, R. M. R., Del Bel Cury, A. A., & Manfredini, D. (2018). Prevalence of psychosocial impairment in temporomandibular disorder patients: A systematic review. *Journal of Oral Rehabilitation*, *45*(11), 881–889. https://doi.org/10.1111/joor.12685

Cano-García, F. J., Muñoz-Navarro, R., Sesé Abad, A., Moretti, L. S., Medrano, L. A., Ruiz-Rodríguez, P., González-Blanch, C., Moriana, J. A., & Cano-Vindel, A. (2020). Latent structure and factor invariance of somatic symptoms in the patient health questionnaire (PHQ-15). *Journal of Affective Disorders*, *261*, 21–29. https://doi.org/10.1016/j.jad.2019.09.077

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chilcot, J., Rayner, L., Lee, W., Price, A., Goodwin, L., Monroe, B., Sykes, N., Hansford, P., & Hotopf, M. (2013). The factor structure of the PHQ-9 in palliative care. *Journal of Psychosomatic Research*, *75*(1), 60–64. https://doi.org/10.1016/j.jpsychores.2012.12.012

Durham, J. (2013). Oral surgery: Part 3. Temporomandibular disorders. *British Dental Journal*, *215*(7), 331–337. https://doi.org/10.1038/sj.bdj.2013.950

Dworkin, S. F. (2011). The OPPERA study: Act one. *The Journal of Pain*, *12*(11 Suppl), T1–T3. https://doi.org/10.1016/j.jpain.2011.08.004

Dyer, J. R., Williams, R., Bombardier, C. H., Vannoy, S., & Fann, J. R. (2016). Evaluating the psychometric properties of 3 depression measures in a sample of persons with traumatic brain injury and major depressive disorder. *The Journal of Head Trauma Rehabilitation*, *31*(3), 225–232. https://doi.org/10.1097/htr.0000000000000177

Edwards, R. R., Dworkin, R. H., Sullivan, M. D., Turk, D. C., & Wasan, A. D. (2016). The role of psychosocial processes in the development and maintenance of chronic pain. *Journal of Pain*, *17*(9 Suppl), T70–T92. https://doi.org/10.1016/j.jpain.2016.01.001

Feingold, D., Brill, S., Goor-Aryeh, I., Delayahu, Y., & Lev-Ran, S. (2017). Depression and anxiety among chronic pain patients receiving prescription opioids and medical marijuana. *Journal of Affective Disorders*, *218*, 1–7. https://doi.org/10.1016/j.jad.2017.04.026

Fetai, A., Dedic, B., Lajnert, V., & Spalj, S. (2020). To what extent are the characteristics of painful temporomandibular disorders predictors of self-reported limitations in jaw function? *Cranio*, *41*, 290–297. https://doi.org/10.1080/08869634.2020.1853309

Fillingim, R. B., Ohrbach, R., Greenspan, J. D., Knott, C., Dubner, R., Bair, E., Baraian, C., Slade, G. D., & Maixner, W. (2011). Potential psychosocial risk factors for chronic TMD: Descriptive data and empirically identified domains from the OPPERA case-control study. *Journal of Pain*, *12*(11), T46–T60. https://doi.org/10.1016/j.jpain.2011.08.007

Finch, W. H. (2020). *Exploratory factor analysis* (pp. 1–12). SAGE Publications, Inc. https://doi.org/10.4135/9781544339900

Frazier, T., Hogue, C. J., & Yount, K. M. (2018). The development of the healthy pregnancy stress scale, and validation in a sample of low-income African American women. *Maternal and Child Health Journal*, *22*(2), 247–254. https://doi.org/10.1007/s10995-017-2396-7

Gagnier, J. J., Lai, J., Mokkink, L. B., & Terwee, C. B. (2021). COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Quality of Life Research*, *30*(8), 2197–2218. https://doi.org/10.1007/s11136-021-02822-4

Giannakopoulos, N. N., Keller, L., Rammelsberg, P., Kronmüller, K. T., & Schmitter, M. (2010). Anxiety and depression in patients with chronic temporomandibular pain and in controls. *Journal of Dentistry*, *38*(5), 369–376. https://doi.org/10.1016/j.jdent.2010.01.003

Granillo, M. T. (2012). Structure and function of the patient health Questionnaire-9 among Latina and non-Latina white female college students. *Journal of the Society for Social Work and Research*, *3*(2), 80–93. https://doi.org/10.5243/jsswr.2012.6

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis (4th ed.): With readings*. Prentice-Hall, Inc.

Hampf, G. (1990). A biopsychosocial approach to TMJ pain-or looking for keys in the dark. *Proceedings of the Finnish Dental Society*, *86*(3–4), 171–181.

Han, C., Pae, C.-U., Patkar, A. A., Masand, P. S., Woong Kim, K., Joe, S.-H., & Jung, I.-K. (2009). Psychometric properties of the patient health questionnaire–15 (PHQ–15) for measuring the somatic symptoms of psychiatric outpatients. *Psychosomatics*, *50*(6), 580–585. https://doi.org/10.1016/S0033-3182(09)70859-X

Harte, S. E., Harris, R. E., & Clauw, D. J. (2018). The neurobiology of central sensitization. *Journal of Applied Biobehavioral Research*, *23*(2), e12137. https://doi.org/10.1111/jabr.12137

Heindl, L. M., Trester, M., Guo, Y., Zwiener, F., Sadat, N., Pine, N. S., Traweger, A., & Rokohl, A. C. (2021). Anxiety and depression in patients wearing prosthetic eyes. *Graefes Archive for Clinical and Experimental Ophthalmology*, *259*(2), 495–503. https://doi.org/10.1007/s00417-020-04908-0

Hietaharju, M., Kivimäki, I., Heikkilä, H., Näpänkangas, R., Teerijoki-Oksa, T., Tanner, J., Kemppainen, P., Tolvanen, M., Suvinen, T., & Sipilä, K.

(2021). Comparison of Axis II psychosocial assessment methods of RDC/TMD and DC/TMD as part of DC/TMD-FIN phase II validation studies in tertiary care Finnish TMD pain patients. *Journal of Oral Rehabilitation*, *48*(12), 1295–1306. https://doi.org/10.1111/joor.13260

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hyphantis, T., Kroenke, K., Papatheodorou, E., Paika, V., Theocharopoulos, N., Ninou, A., Tomenson, B., Carvalho, A. F., & Guthrie, E. (2014). Validity of the Greek version of the PHQ 15-item somatic symptom severity scale in patients with chronic medical conditions and correlations with emergency department use and illness perceptions. *Comprehensive Psychiatry*, *55*(8), 1950–1959. https://doi.org/10.1016/j.comppsych.2014.08.042

IASP. (2011). IASP terminology. https://www.iasp-pain.org/resources/terminology/

Johnson, S. U., Ulvenes, P. G., Øktedalen, T., & Hoffart, A. (2019). Psychometric properties of the general anxiety disorder 7-item (GAD-7) scale in a heterogeneous psychiatric sample. *Frontiers in Psychology*, *10*, 1713. https://doi.org/10.3389/fpsyg.2019.01713

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kim, J.-O., & Mueller, C. W. (1978). *Introduction to factor analysis what it is and how to do it* (Vol. *13*). SAGE Publications, Inc.

Kingsley, C., & Patel, S. (2017). Patient-reported outcome measures and patient-reported experience measures. *BJA Education*, *17*(4), 137–144. https://doi.org/10.1093/bjaed/mkw060JBJAEducation

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Krause, J. S., Saunders, L. L., Bombardier, C., & Kalpakjian, C. (2011). Confirmatory factor analysis of the patient health Questionnaire-9: A study of the participants from the spinal cord injury model systems. *PM&R*, *3*(6), 533–540. https://doi.org/10.1016/j.pmrj.2011.03.003

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine*, *64*(2), 258–266. https://doi.org/10.1097/00006842-200203000-00008

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, *114*(1–3), 163–173. https://doi.org/10.1016/j.jad.2008.06.026

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(8), 2207. https://doi.org/10.4236/psych.2018.98126

Lamela, D., Soreira, C., Matos, P., & Morais, A. (2020). Systematic review of the factor structure and measurement invariance of the patient health questionnaire-9 (PHQ-9) and validation of the Portuguese version in community settings. *Journal of Affective Disorders*, *276*, 220–233. https://doi.org/10.1016/j.jad.2020.06.066

Ledesma, R. D., Valero-Mora, P., & Macbeth, G. (2015). The scree test and the number of factors: A dynamic graphics approach. *The Spanish Journal of Psychology*, *18*, E11. https://doi.org/10.1017/sjp.2015.13

Lee, M., Snow, J., Quon, C., Selander, K., DeRycke, E., Lawless, M., Driscoll, M., Ditre, J. W., Mattocks, K. M., Becker, W. C., & Bastian, L. A. (2021). I smoke to cope with pain: Patients' perspectives on the link between cigarette smoking and pain. *Wiener Klinische Wochenschrift*, *133*(19–20), 1012–1019. https://doi.org/10.1007/s00508-021-01931-x

Leeuw, R. d. (2013). American Academy of orofacial pain guidelines for assessment, diagnosis, and management.

Leonhart, R., de Vroege, L., Zhang, L., Liu, Y., Dong, Z., Schaefert, R., Nolte, S., Fischer, F., Fritzsche, K., & van der Feltz-Cornelis, C. M. (2018). Comparison of the factor structure of the patient health questionnaire for somatic symptoms (PHQ-15) in Germany, The Netherlands, and China. A transcultural structural equation modeling (SEM) study. *Frontiers in Psychiatry*, *9*, 240. https://doi.org/10.3389/fpsyt.2018.00240

Liao, S. C., Huang, W. L., Ma, H. M., Lee, M. T., Chen, T. T., Chen, I. M., & Gau, S. S. (2016). The relation between the patient health questionnaire-15 and DSM somatic diagnoses. *BMC Psychiatry*, *16*(1), 351. https://doi.org/10.1186/s12888-016-1068-2

Lin, Q., Bonkano, O., Wu, K., Liu, Q., Ali Ibrahim, T., & Liu, L. (2021). The value of Chinese version GAD-7 and PHQ-9 to screen anxiety and depression in Chinese outpatients with atypical chest pain. *Therapeutics and Clinical Risk Management*, *17*, 423–431. https://doi.org/10.2147/tcrm.S305623

Macfarlane, T. V., Blinkhorn, A. S., Davies, R. M., Ryan, P., Worthington, H. V., & Macfarlane, G. J. (2002). Orofacial pain: Just another chronic pain? Results from a Population-Based Survey. *Pain*, *99*(3), 453–458. https://doi.org/10.1016/s0304-3959(02)00181-1

Manfredini, D., Winocur, E., Ahlberg, J., Guarda-Nardini, L., & Lobbezoo, F. (2010). Psychosocial impairment in temporomandibular disorders patients. RDC/TMD axis II findings from a multicentre study. *Journal of Dentistry*, *38*(10), 765–772. https://doi.org/10.1016/j.jdent.2010.06.007

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1171–1179. https://doi.org/10.1007/s11136-017-1765-4

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010a). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010b). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, *63*(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Nima, A. A., Cloninger, K. M., Lucchese, F., Sikström, S., & Garcia, D. (2020). Validation of a general subjective well-being factor using classical test theory. *PeerJ*, *8*, e9193. https://doi.org/10.7717/peerj.9193

North, C. S., Hong, B. A., Lai, H. H., & Alpers, D. H. (2019). Assessing somatization in urologic chronic pelvic pain syndrome. *BMC Urology*, *19*(1), 130. https://doi.org/10.1186/s12894-019-0556-3

Ohrbach, R., Granger, C., List, T., & Dworkin, S. (2008). Preliminary development and validation of the jaw functional limitation scale. *Community Dentistry and Oral Epidemiology*, *36*(3), 228–236. https://doi.org/10.1111/j.1600-0528.2007.00397.x

Ohrbach, R., Larsson, P., & List, T. (2008). The jaw functional limitation scale: Development, reliability, and validity of 8-item and 20-item versions. *Journal of Orofacial Pain*, *22*(3), 219–230.

Pagán-Torres, O. M., González-Rivera, J. A., & Rosario-Hernández, E. (2020). Psychometric analysis and factor structure of the Spanish version of the eight-item patient health questionnaire in a general sample of Puerto Rican adults. *Hispanic Journal of Behavioral Sciences*, *42*(3), 401–415. https://doi.org/10.1177/0739986320926524

Pincus, T., Burton, A. K., Vogel, S., & Field, A. P. (2002). A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine*, *27*(5), E109–E120. https://doi.org/10.1097/00007632-200203010-00017

Reddy, A. S., Tomita, A., & Paruk, S. (2021). Depression, anxiety and treatment satisfaction in the parents of children on antiretroviral therapy in South Africa. *Psychology, Health & Medicine*, *26*(5), 584–594. https://doi.org/10.1080/13548506.2020.1837389

Reiter, S., Emodi-Perlman, A., Goldsmith, C., Friedman-Rubin, P., & Winocur, E. (2015). Comorbidity between depression and anxiety in patients with temporomandibular disorders according to the research diagnostic criteria for temporomandibular disorders. *Journal of Oral & Facial Pain and Headache*, *29*(2), 135–143. https://doi.org/10.11607/ofph.1297

Resende, C. M. B. M. d., Rocha, L. G. D. d. S., Paiva, R. P. d., Cavalcanti, C. d. S., Almeida, E. O. d., Roncalli, A. G., & Barbosa, G. A. S. (2020). Relationship between anxiety, quality of life, and sociodemographic characteristics and temporomandibular disorder. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, *129*(2), 125–132. https://doi.org/10.1016/j.oooo.2019.10.007

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–514. https://doi.org/10.1007/BF02296192

Schiffman, E., Ohrbach, R., Truelove, E., Look, J., Anderson, G., Goulet, J. P., List, T., Svensson, P., Gonzalez, Y., Michelotti, A., Brooks, S. L., Ceusters, W., Drangsholt, M., Ettlin, D., Gaul, C., Goldberg, L. J., Haythornthwaite, J. A., Hollender, L., Jensen, R., … Dworkin, S. F. (2014). 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 Groupdagger. *Journal of Oral & Facial Pain and Headache*, *28*(1), 6–27. https://doi.org/10.11607/jop.1151

Simoen, L., Van den Berghe, L., Jacquet, W., & Marks, L. (2020). Depression and anxiety levels in patients with temporomandibular disorders: Comparison with the general population. *Clinical Oral Investigations*, *24*(11), 3939–3945. https://doi.org/10.1007/s00784-020-03260-1

Soto-Balbuena, C., Rodríguez-Muñoz, M. F., & Le, H. N. (2021). Validation of the generalized anxiety disorder screener (GAD-7) in Spanish pregnant women. *Psicothema*, *33*(1), 164–170. https://doi.org/10.7334/psicothema2020.167

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Suvinen, T. I., Reade, P. C., Kemppainen, P., Kononen, M., & Dworkin, S. F. (2005). Review of aetiological concepts of temporomandibular pain disorders: Towards a biopsychosocial model for integration of physical disorder factors with psychological and psychosocial illness impact factors. *European Journal of Pain*, *9*(6), 613–633.

Tabachnick, B. G., & Fidell, L. S. (Producer). (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon/Pearson Education.

Terrill, A. L., Hartoonian, N., Beier, M., Salem, R., & Alschuler, K. (2015). The 7-item generalized anxiety disorder scale as a tool for measuring generalized anxiety in multiple sclerosis. *International Journal of MS Care*, *17*(2), 49–56. https://doi.org/10.7224/1537-2073.2014-008

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Vissers, M. M., Bussmann, J. B., Verhaar, J. A., Busschbach, J. J., Bierma-Zeinstra, S. M., & Reijman, M. (2012). Psychological factors affecting the outcome of total hip and knee arthroplasty: A systematic review. *Seminars in Arthritis and Rheumatism*, *41*(4), 576–588. https://doi.org/10.1016/j.semarthrit.2011.07.003

Volkman, J. E., DeRycke, E. C., Driscoll, M. A., Becker, W. C., Brandt, C. A., Mattocks, K. M., Haskell, S. G., Bathulapalli, H., Goulet, J. L., & Bastian, L. A. (2015). Smoking status and pain intensity among OEF/OIF/OND veterans. *Pain Medicine*, *16*(9), 1690–1696. https://doi.org/10.1111/pme.12753

Weingarten, T. N., Iverson, B. C., Shi, Y., Schroeder, D. R., Warner, D. O., & Reid, K. I. (2009). Impact of tobacco use on the symptoms of painful temporomandibular joint disorders. *Pain*, *147*(1–3), 67–71. https://doi.org/10.1016/j.pain.2009.08.021

Xu, L., He, Y., Fan, S., Cai, B., Fang, Z., & Dai, K. (2020). Validation of a Chinese version of the jaw functional limitation scale in relation to the diagnostic subgroup of temporomandibular disorders. *Journal of Oral Rehabilitation*, *47*(1), 1–8. https://doi.org/10.1111/joor.12868

Yap, A. U. J., Tan, K. B. C., Prosthodont, C., Chua, E. K., & Tan, H. H. (2002). Depression and somatization in patients with temporomandibular disorders. *The Journal of Prosthetic Dentistry*, *88*(5), 479–484. https://doi.org/10.1067/mpr.2002.129375

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Taimeh, D., Riordain, R. N., Fedele, S., & Leeson, R. (2023). Validity and internal consistency of four scales in patients with TMD: PHQ8, GAD7, PHQ15 and JFLS20. *Oral Diseases*, *00*, 1–12. https://doi.org/10.1111/odi.14729