Neurological Sciences (2022) 43:4157–4165
https://doi.org/10.1007/s10072-022-06066-z

REVIEW ARTICLE

Psychometric properties of the Facial Disability Index in patients with facial palsy: a systematic review and meta-analysis

Fatih Özden[1] · **İsmet Tümtürk**[2] · **Zübeyir Sarı**[3]

Received: 26 November 2021 / Accepted: 6 April 2022 / Published online: 13 April 2022

Abstract

Purpose The aim of the study was to present the systematic review and meta-analysis of the psychometrical analysis of Facial Disability Index (FDI) studies.

Methods A literature search was conducted in the relevant electronic databases "PubMed, Scopus, Web of Science (WoS), and Cochrane Library." A total of 621 articles were obtained by searching the relevant keywords (PubMed: 384, Cochrane Library: 14, Web of Science: 132, Scopus: 91). A total of 8 papers were included. The four-point classification and rating-based "COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)" tools were used to evaluate the bias risk and evidence levels.

Results Cronbach's alpha pooling of FDI total score was (ES): 0.803 (95% CI: 0.73–0.86). Heterogeneity for the Facial Disability Index-Physical Function (FDI-PF) and Facial Disability Index-Social Function (FDI-SF) subscore based on intraclass correlation coefficient (ICC) were $I^2 = 84.2\%$ (ICC: 0.88, 95% CI: 0.81–0.92) and $I^2 = 73.7\%$ (ICC: 0.87, 95% CI: 0.81–0.90), respectively. Correlational results between Sunnybrook Facial Grading System (SFGS) with FDI-PF and FDI-SF were 0.38 and 0.22, respectively. The correlations of FDI-PF with Short Form-12 Physical Component Summary (SF-12-PCS) and Short Form-12 Mental Health Component Summary (SF-12-MCS) were 0.43 and 0.28, respectively. Correlation results of FDI-SF with SF-12-PCS and SF-12-MCS were 0.23 and 0.57. The relationship results of Facial Clinimetric Evaluation with FDI-PF and FDI-SF were 0.71 and 0.57, respectively.

Conclusion FDI is a psychometrically valuable questionnaire, especially for the internal consistency, reliability, and validity. In clinical practice, the use of FDI would be valuable, in addition to clinician-based grading, to see more of patients' social influences precisely.

Keywords Bell's palsy · Facial disability · FDI, reliability · Responsiveness · Validity

Introduction

Various evaluation methods are developed to evaluate facial palsy (FP). However, most of these are clinician-based tools that are not filled by the patient and only measure motor deficiencies. FP is a disease that causes motor deficiency and psychosocial symptoms. In addition, it is essential to evaluate both motor and psychosocial outcomes in patients with FP simultaneously. One of the frequently used patient-based tools, "Facial Disability Index (FDI)," addresses the non-motor disabilities of patients with FP explicitly [1, 2]. FDI is a short, self-report questionnaire that aims to evaluate the patient's daily living experience with a facial nerve disorder, disability, psychosocial factors, and the intervention outcome.

Van Swearingen and Branch developed FDI in 1996 for patients with disorders of the facial neuromuscular system [3]. The FDI consists of 10 questions, five questions in the physical function subscore and five questions in the social function and well-being subscore. The first part of the FDI items assesses difficulty and functionality in eating, drinking, talking, and brushing teeth. The second part examines

✉ Fatih Özden
fatihozden@mu.edu.tr

[1] Department of Healthcare Services, Köyceğiz Vocational School of Health Services, Muğla Sıtkı Koçman University, Muğla, Turkey

[2] Department of Physiotherapy and Rehabilitation, Institute of Health Sciences, Ege University, İzmir, Turkey

[3] Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, Marmara University, İstanbul, Turkey

social well-being such as social life, emotional state, and sleep quality. These ten questions are scored on a 6-point Likert scale. Final scoring ranges from 0 (worst) to 100 (best) [4, 5].

Validity, reliability, and cultural adaptation studies have reported Spanish, Swedish, Italian, French, Brazilian, Turkish, and Dutch versions. Some studies reported bivariate positive correlations between FDI and House-Brackmann (H-B) scale, moderate-to-strong correlation between social areas compared to equivalent areas in Short Form-36 (SF-36). The FDI physical function subscale was correlated with the Sunnybrook Face Rating System composite score and, the social/well-being function subscale was correlated with the 12-item Short Form Health Questionnaire (SF-12) mental component." According to the H-B global scale, there was also a relationship between FDI and the degree of facial dysfunction; the FDI physical function subscale was correlated with SF-36 and H-B's physical function subscale. Moreover, the FDI social/well-being function subscale was correlated with H-B and SF-36 with all subscales except for role-functioning physical (RP) [4–10].

To our knowledge, there has not been a study that compiles and methodologically analyzes FDI's psychometric studies yet. The use of the FDI and the importance of the biopsychosocial model in treating and evaluating symptoms of FP such as motor disability, quality of life, and psychological disorder in previous studies require this scale to be presented at higher levels of evidence [11–15]. Examining the evidence level of the psychometric studies of FDI and exploring the clinical results obtained would strengthen its appropriate use in practice. The aim of the study was to present the systematic review and meta-analysis of the psychometrical analysis of FDI studies.

Methods

Search strategy and selection criteria

The present systematic review and meta-analysis was presented based on the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)" and the "Cochrane guidelines" [16, 17]. A literature search was conducted in November 2021. The electronic databases of "PubMed, Scopus, Web of Science (WoS), and Cochrane Library" were searched with relevant terms and synonyms. The search was performed with a combination of "Boolean operators" (AND, OR). The search strategy is detailed in Supplementary File 1.

Eligibility criteria

Inclusion criteria of the study were (1) studies published between 1995 and 2021, (2) studies conducted on subjects with FP, (3) studies investigating at least one psychometric feature (e.g., reliability, validity, responsiveness) of FDI, and (4) studies with full texts available and written in English. Exclusion criteria of the study were (1) studies written other than English and (2) articles with no full text available.

Study selection and data extraction

All searched studies were exported to the software (Rayyan, Rayyan Systems Inc., USA). The responsiveness of Rayyan was also reported by Ouzzani et al. [18]. The duplicate publications were detected and removed by this software. The two researchers of the study independently screened the articles on their titles and abstracts, considering the eligibility criteria. The results were determined by reaching a joint decision in the articles as "exclude, include, or maybe." The following data from these articles have been extracted: first author, publication date, setting, intervals, number of cases, age, gender, FP type, validation test.

Quality assessment

"The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)" tool was used to evaluate the studies' bias risk and methodological quality [19]. The researchers independently evaluated the eight included studies using the 4-point COSMIN checklist. The COSMIN tool provides an essential contribution to selecting and using high-quality outcome measures. The methodological quality indicated by each item was assigned as excellent, good, fair, or poor. In addition, a quantitative assessment COSMIN rating tool was used for each psychometric measure of the studies [20]. The measurements in each study were classified with positive (+), indeterminate (?), and negative (−) ratings.

Meta-analysis (quantitative analysis of studies)

Homogeneous data of the studies were pooled with R software version 4 and Microsoft Excel [21]. This meta-analysis used Fisher's z transformation of correlations [22]. Fisher's z transformation provides a reliable confidence interval for the correlational results of the studies. In addition, the random-effects model was estimated. This calculation model is utilized to provide a resume estimation of the magnitude of effect in a meta-analysis. The Schmidt-Hunter method was preferred to reveal the size of correlation coefficients [23].

In this way, bias was tried to be minimized. Heterogeneity was assessed with the I^2 value [24]. Heterogeneity shows the clinical diversity of the studies, which included in the meta-analysis.

Results

Search outcome

A total of 621 articles were obtained by searching the relevant keywords (PubMed: 384, Cochrane Library: 14, Web of Science: 132, Scopus: 91). Six hundred ten were excluded because they did not meet the elimination criteria and duplicated articles. Following the first screening, a total of 10 articles were obtained and reviewed. For review, studies that did not meet the inclusion criteria (i.e., non-English studies, studies for which the full text was not available) were eliminated. Two were excluded as full text was unavailable, and the other was German. A total of 8 papers were included in this systematic review (Fig. 1).

Study characteristics

Studies were carried out between 1996 and 2020. The lowest mean age of the study was 40 years [4], and the highest was 56.5 years [8]. The sample size was also min–max, 46–118 individuals. Four studies (50%) were performed in individuals with peripheral facial palsy [6, 8–10]. Most studies (75%) considered it appropriate to compare the FDI with the SF-36 [3, 4, 6, 8–10] in hypothesis testing. Two studies compared SF-12 with FDI, taking into account subscore compatibility [5, 7]. The most preferred (*n*: 3) clinician-based tool for comparison was H-B [4, 6, 8]. Two studies checked construct validity by comparing SFGS and Facial Clinimetric Evaluation (FaCE) with FDI,

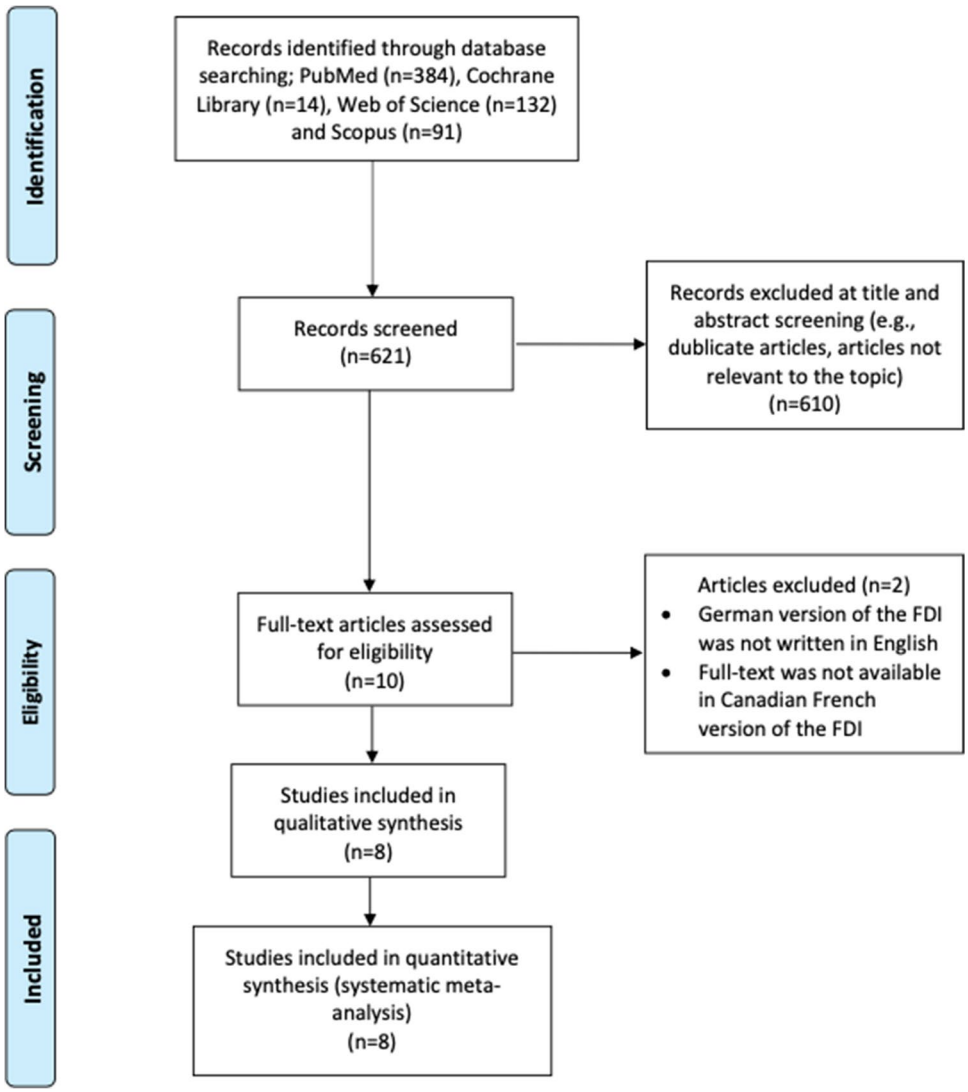

Fig. 1 PRISMA flow diagram of the study

respectively. The interval for retest was at least 1 week and at most 1 month. In all studies in which gender distribution was presented (*n*: 6) [3, 5–8, 10], the number of women was higher than the number of men (Table 1).

Quality assessment and evidence level

The 4-item COSMIN scoring yielded the quality assessment of studies' psychometric properties (Tables 2, 3, and 4). Only two studies had internal consistency scores of "poor" [3, 9].

Table 1 The characteristic overview of the studies

Author	Year	Setting	Intervals (test–retest)	*n*	Age (years)	Gender (%)	Facial palsy type	Validation tests
Barry et al	2019	Tertiary clinic	1 week	67	56.36 ± 14.2	61.2% female	Peripheral facial palsy	FDI, FaCE, and SF-36
Gonzalez-Cardero et al	2012	Clinic	N/A	79	40	N/A	Facial palsy after superficial parotidectomy	H-B, FDI, and SF-36
Graciano et al	2020	Tertiary clinic	N/A	100	48.5	N/A	Peripheral facial palsy	H-B, FDI, and SF-36
Marsk et al	2013	Clinic	2 weeks	93	56.5 ± 10.5	52.7% female	Peripheral facial palsy	FDI, FaCE, and SF-36
Özden et al	2020	Clinic	1 week	51	46.7 ± 17.1	51% female	Peripheral facial palsy	H-B, FDI, and SF-36
Pavese et al	2014	Clinic/outpatient	1 month	100	45 ± 15	72% female	Unilateral facial palsy	FDI, SFGS, and SF-12
Van Swearingen et al	1996	Clinic	N/A	46	46.8 ± 15.6	65.21% female	Facial neuromuscular dysfunction	FDI and SF-36
Van Veen et al	2020	Clinic	2 weeks	118	48.8	52.5% female	Unilateral facial palsy	FDI, SFGS, and SF-12

n, number of participants; *N/A*, not available; *SD*, standard deviation; *FDI*, Facial Disability Index; *FaCE*, Facial Clinimetric Evaluation; *SF-36*, Short Form 36; *H-B*, House-Brackmann scale; *SFGS*, Sunnybrook Facial Grading System; *SF-12*, Short Form 12

Table 2 The results of the reliability and measurement error

Author	Internal consistency		Test–retest reliability		Measurement error	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
Barry et al	FDIα = 0.84 FDI-PFα = 0.88 FDI-PFα = 0.70	Good	FDI-PF_{ICC} = 0.81 FDI-SF_{ICC} = 0.86	Good	N/A	Poor
Gonzalez-Cardero et al	FDIα = 0.83 FDI-PFα = 0.77 FDI-PFα = 0.82	Good	N/A	Poor	N/A	Poor
Graciano et al	FDIα = 0.73	Poor	FDI-PF_{ICC} = 0.79 FDI-SF_{ICC} = 0.85	Fair	N/A	Poor
Marsk et al	FDI-PFα = 0.87 FDI-PFα = 0.82	Good	FDI-PF_{ICC} = 0.93 FDI-SF_{ICC} = 0.91	Good	N/A	Poor
Özden et al	FDI-PFα = 0.82 FDI-PFα = 0.63	Fair	FDI-PF_{ICC} = 0.91 FDI-SF_{ICC} = 0.93	Fair	N/A	Poor
Pavese et al	FDI-PF Θ = 0.82 FDI-PF Θ = 0.78	Excellent	FDI-PF_{ICC} = 0.93 FDI-SF_{ICC} = 0.84	Poor	N/A	Poor
Van Swearingen et al	N/A	Poor	N/A	Poor	N/A	Poor
Van Veen et al	FDI-PFα = 0.72 FDI-PFα = 0.60	Excellent	FDI-PF_{ICC} = 0.84 FDI-SF_{ICC} = 0.78	Good	FDI-PF_{SDC} = 17.6 FDI-SF_{SDC} = 17.7	Excellent

N/A, not available; *SD*, standard deviation; *FDI*, Facial Disability Index; *FDI-PF*, Facial Disability Index-Physical Function; *FDI-SF*, Facial Disability Index-Social Function; *ICC*, intraclass correlation coefficient; *SDC*, smallest detectable change; *α*, Cronbach's alpha; *Θ*, theta coefficient; *COSMIN*, COnsensus-based Standards for the selection of health Measurement INstruments

Table 3 The results of the validity and hypotheses testing

Author	Content validity		Structural validity		Hypotheses testing	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
Barry et al	N/A	Poor	N/A	Poor	H-B vs FDI-PF: − 0.35 H-B vs FDI-SF: − 0.25 SF-36 vs FDI-PF: 0.29–0.49 SF-36 vs FDI-SF: 0.42–0.65	Fair
Gonzalez-Cardero et al	N/A	Poor	Available	Good	H-B vs FDI-PF: 0.40 H-B vs FDI-SF: 0.27 SF-36-PF vs FDI-PF: 0.43 SF-36-PF vs FDI-SF: 0.36	Good
Graciano et al	N/A	Poor	N/A	Poor	SF-36 vs FDI-PF: 0.12–0.24 SF-36 vs FDI-SF: 0.27–0.59	Good
Marsk et al	N/A	Poor	N/A	Poor	H-B vs FDI-PF: − 0.61 H-B vs FDI-SF: − 0.38 SFGS vs FDI-PF: 0.63 SFGS vs FDI-SF: 0.40 FaCE vs FDI-PF: 0.76 FaCE vs FDI-SF: 0.63 SF-36 vs FDI-PF: 0.31–0.59 SF-36 vs FDI-SF: 0.29–0.61	Good
Özden et al	N/A	Poor	Available	Good	H-B vs FDI-PF: − 0.83 H-B vs FDI-SF: − 0.35 SF-36 vs FDI-PF: 0.04–0.33 SF-36 vs FDI-PF: 0.15–0.58	Good
Pavese et al	N/A	Poor	Available	Excellent	SFGS vs FDI-PF: 0.17–0.44 SFGS vs FDI-SF: 0.07–0.33 SF-12 vs FDI-PF: 0.39–0.40 SF-12 vs FDI-SF: 0.07–0.55	Excellent
Van Swearingen et al	N/A	Poor	Available	Good	SF-36 vs FDI: 0.24–0.79	Good
Van Veen et al	Available	Poor	Available	Excellent	FaCE vs FDI-PF: 0.09 − 0.71 FaCE vs FDI-SF: 0.09 − 0.56 SF-12 vs FDI-PF: 0.16 − 0.46 SF-12 vs FDI-SF: 0.37 − 0.59 SAQ vs FDI-PF: − 0.24 SAQ vs FDI-SF: − 0.29 SFGS vs FDI-PF: 0.07 SFGS vs FDI-SF: 0.02	Excellent

N/A, not available; *SD*, standard deviation; *FDI*, Facial Disability Index; *FDI-PF*, Facial Disability Index-Physical Function; *FDI-SF*, Facial Disability Index-Social Function; *FaCE*, Facial Clinimetric Evaluation; *SF-36*, Short Form 36; *H-B*, House-Brackmann scale; *SFGS*, Sunnybrook Facial Grading System; *SF-12*, Short Form 12; *SAQ*, Synkinesis Assessment Questionnaire; *COSMIN*, COnsensus-based Standards for the selection of health Measurement INstruments

Table 4 The results of the validity and responsiveness

Author	Cross-cultural validity		Criterion validity		Responsiveness	
	Results	COSMIN score	Results	COSMIN score	Results	COSMIN score
Barry et al	Available	Fair	Available	Good	N/A	Poor
Gonzalez-Cardero et al	Available	Fair	Available	Good	N/A	Poor
Graciano et al	Available	Good	Available	Excellent	N/A	Poor
Marsk et al	Available	Good	Available	Good	N/A	Poor
Özden et al	Available	Fair	Available	Good	N/A	Poor
Pavese et al	Available	Good	Available	Excellent	Available	Good
Van Swearingen et al	Available	Poor	Available	Fair	N/A	Poor
Van Veen et al	Available	Excellent	Available	Excellent	N/A	Poor

N/A, not available; *COSMIN*, COnsensus-based Standards for the selection of health Measurement INstruments

Similarly, three studies performed the test–retest measure with poor quality [3, 4, 7]. Only one study did a statistical analysis of measurement error [5]. Three studies had poor scores in terms of construct validity [6, 8, 9].

All studies performed correlation-based hypothesis testing by comparison with the gold standard tool. Most studies (*n*: 5) were of "good" quality [3, 4, 6, 9, 10]. On the other hand, the results based on the quality analysis in terms of cultural adaptation showed that only 1 study was of poor quality [3]. According to the results based on criterion validity, half of the studies were of "good" quality, while only 1 study focused on responsiveness [7].

The evidence level of the studies is presented in Table 5. Only productive results were achieved in internal consistency, reliability, and cross-cultural adaptation. The internal consistency analysis of most studies (*n*: 7) received a positive rating [4–10], and similarly, the reliability measure of 6 studies was positive [5–10]. On the other hand, according to COSMIN, the cross-cultural adaptation measurement of all studies got negative scores.

Reliability

In the reliability analysis, internal consistency and test–retest reliability were considered. First of all, to address the internal consistency, 3 studies presented Cronbach's alpha based on the total score, although the FDI does not have a total score. Five studies gave alpha values of FDI's subscores. The meta-analysis results based on data pooling for the FDI total score (three studies) [4, 8, 9] were Cohen's effect size (ES): 0.803 (95% confidence interval (CI): 0.73–0.86). On the other hand, FDI-PF- and FDI-SF-based pooling results were $ES_{FDI\text{-}PF}$: 0.830 (95% CI: 0.77–0.87) and $ES_{FDI\text{-}SF}$: 0.728 (95% CI: 0.62–0.81), respectively [4–8, 10]. We pooled and analyzed intraclass correlation coefficient (ICC) scores for FDI subscores in test–retest reliability. According to the results of the meta-analysis of six studies in total, for the PF subscore, the heterogeneity was $I^2 = 84.2\%$ (Fisher *z*: 1.38, ICC: 0.88, 95% CI: 0.81–0.92), and for the SF subscore, $I^2 = 73.7\%$ (Fisher *z*: 1.32, ICC: 0.87, 95% CI: 0.81–0.90) was obtained [4–8, 10].

Validity

In the validity analysis, the comparison of the Facial Disability Index-Physical Function (FDI-PF) and Facial Disability Index-Social Function (FDI-SF) with Sunnybrook Facial Grading System (SFGS), SF-12, SF-36, and FaCE subscores, among the results of the correlation analysis, was performed. According to the meta-analysis results, the relationship between FDI-PF and SFGS were 0.38 (95% CI: − 0.24–0.78). In addition, it was pooled correlation coefficients 0.22 (95% CI: − 0.17–0.54) between FDI-SF and SFGS [6, 7]. Physical component of the FDI had higher relationship with clinician-based grading system. Correlation pool results of FDI-PF with Short Form-12 Physical Component Summary Score (SF-12-PCS) and Short Form-12 Mental Health Component Summary Score (SF-12-MCS) were 0.43 (95% CI: 0.31–0.53) and 0.28 (95% CI: 0.03–0.48), respectively. The results of FDI-SF with SF-12-PCS and SF-12-MCS were 0.23 (95% CI: − 0.08–0.49) and 0.57 (95% CI: 0.46–0.65), respectively [5, 7]. Namely, FDI conformed with the quality of life scores of the FD patients. On the other hand, FaCE had a correlation of 0.71 (95% CI: 0.59–79) and 0.57 (95% CI: 0.43–67) for FDI-PF and FDI-SF, respectively [5, 6]. FDI was highly related with an alternative PROM for the FD. The correlation pool results of FDI with SF-36 are presented in Table 6. The coefficient of correlation of the SF score of FDI with SF-36 was 0.27–0.65 [3, 6–10]. FDI had low and moderate correlation with SF-36, which is considered the gold standard for quality of life. Finally, there was a correlation of − 0.24 and − 0.29 between Synkinesis Assessment Questionnaire (SAQ) and FDI-PF

Table 5 Evidence level of the studies

Author	Evidence level of psychometric properties							
	Structural validity	Internal consistency	Reliability	Measurement error	Hypothesis testing	Construct validity	Cross-cultural validity	Responsiveness
Barry et al	?	+	+	?	?	?	–	?
Gonzalez-Cardero et al	?	+	?	?	?	?	–	?
Graciano et al	?	+	+	?	?	?	–	?
Marsk et al	?	+	+	?	?	?	–	?
Özden et al	?	+	+	?	?	?	–	?
Pavese et al	?	+	+	?	?	?	–	-
Van Swearingen et al	?	?	?	?	?	?	–	?
Van Veen et al	?	+	+	?	?	?	–	?

(+) positive rating, (?) indeterminate, (−) negative rating

Table 6 Correlation between SF-36 and FDI

	Physical function (*R*)	Social/well-being function (*R*)
SF-36		
Physical function	0.32	0.35
Role physical	0.35	0.42
Bodily pain	0.37	0.35
General health	0.33	0.40
Vitality	0.36	0.50
Social function	0.47	0.57
Role emotional	0.30	0.53
Mental health	0.38	0.57

R, pooled correlation coefficients

and FDI-SF, respectively [5]. The FDI represented a low-level clinical condition for the measurement of synkinesis.

Other psychometric properties

Only one study presented an analysis of measurement properties [5]. The smallest detectable change (SDC) for FDI-PF and FDI-SF were 17.6 and 17.7, respectively. Another study presented responsiveness analysis after 1 month of follow-up. In this study (Italian version of the FDI) [7], "effect size, standardized response means, and responsiveness ratio" was 1, 1.03, and 1.21 for the FDI-PF, respectively. On the other hand, it was 0.75, 0.83, and 1.15, respectively, for the FDI-SF.

Discussion

The present study aimed to demonstrate the systematic review and meta-analysis of the psychometrical properties of FDI. A patient-reported outcome measurement (PROM)–based facial palsy measurement may yield more comprehensive results, especially in the psychosocial dimension, than in clinician-based measurements. Therefore, a detailed analysis of qualitative and quantitative data with a comprehensive psychometric analysis contributes to the literature. Considering the physical and social dimensions of FDI, it would be more advantageous to provide essential guidance to clinicians for clinical practice.

So far, eight studies have revealed the psychometric properties of FDI [3–10]. In general, the sample size is expected to include individuals ten times the number of questionnaire items in such studies [20]. We observed that all eight studies formed the sample by adhering to this criterion. Another issue was the pathology type of the sample. FDI has been tested in peripheral and central palsy and specifically in individuals with unilateral involvement. A wide range of sample options strengthens the inferences about FDI. For example, Gonzalez-Cardero et al. showed that the internal consistency of the FDI was sufficient in their analysis of individuals with facial palsy after superficial parotidectomy in the post-surgical period [4]. On the other hand, a low-to-moderate correlation between the quality of life and clinician-based evaluation results and FDI revealed that it could be used healthily in the post-surgical period.

Another critical point was the gender distribution in the studies. Similar to current epidemiological comprehensive study data, it has been shown that Bell's palsy is slightly more common in women [25]. A transformative study also highlighted a better prognosis in men in its preliminary results of facial palsy [26]. In this respect, it was expected that most of the individuals who applied to the clinic for treatment were women. It is also seen that FDI is mainly applied to individuals between 40 and 60 years. The current epidemiological study has shown that this figure is mostly in individuals between the ages of 60 and 80 [25]. From these results, it can be observed that FDI has not been adequately studied in geriatric individuals. Further studies may focus on FDI in evaluating Bell's palsy in geriatrics. In this way, the psychosocial dimension of facial paralysis can be investigated in the elderly group whose possible depressive symptoms increase [27, 28].

According to the quality assessment results, it was deemed that there was sufficient examination data in terms of internal consistency, test–retest reliability, construct validity, and cultural adaptation [19, 20]. Most of the studies appeared to produce relevant analyzes of fair to excellent quality. It should be assumed that the internal consistency of the FDI is high (alpha > 0.70) [29]. The questionnaire could evaluate different items together for simply one purpose. It can be emphasized that the score in the ICC is generally > 0.80 [30], and the test–retest reliability is high. However, the evaluation time interval showed substantial variation between studies. That is, considering the results of a 1-week interval and 1-month waiting period together, we were unable to obtain precise results [20]. The quality of the studies also differed compared to COSMIN, but it has been shown that the FDI can consistently represent the same clinical situation in different measures. On the other hand, the validity-based quality analysis revealed a very broad spectrum of data. Both quality of life and clinician-based rating tools, which are accepted as a gold standard, were used for comparison. The correlation coefficient ranging from 0.02 to − 0.83 suggested that these results should be considered more comprehensively with our meta-analysis.

Evidence level review, like quality analysis, provided sufficient data for internal consistency, reliability, and validity parameters. It is regarded that psychometric analyses of internal consistency and reliability generally have positive

rankings, and studies with sufficient evidence levels are designed. In terms of cross-cultural validity, negative ranking emerged for all studies. It shows the importance of considering FDI's responsiveness and measurement error parameters in future studies.

The first of our meta-analysis evaluations was the alpha value for the total, physical, and social scores of the FDI. The fact that all values were above 0.70 quantitatively demonstrated that the internal consistency of the FDI was acceptable [20]. In the meta-analysis based on ICC, a coefficient above 0.80 and heterogeneity between studies were found to be moderate to high [24, 30]. It was thought that FDI could reflect the current results appropriately by giving similar results for the same clinical condition of the same individual, both at 1-week and 1-month intervals.

The meta-analysis of the validity results brought with it the main clinical implications. FDI-PF was more correlated with SFGS as expected. SFGS is a clinician-based questionnaire that comprehensively evaluates facial motor functions [31]. However, the FDI also focuses on different activities of daily living (ADL) symptoms, such as eye-tearing and talking. In this respect, the correlational coefficient was not high. It can be presumed that FDI-PF approaches some vital functions more comprehensively. In the comparison of SF-12 and FDI, there was a 0.57 correlation between SF-12-MCS and FDI-SF and a 0.43 correlation between SF-12-PCS and FDI-PF. Hereabouts, it can be mentioned that FDI focuses more on the social component of quality of life and is highly relational analysis compared to the common correlational classification.

On the other hand, FaCE is also a PROM that focuses on both physical and social facial palsy symptoms. The physical size of FDI was found to be more related to FaCE, and the focus of FaCE on more physical symptoms may have been effective in this situation [32]. According to these results, it can be mentioned that the FDI covers the psychosocial status and quality of life of the individual more comprehensively. In addition, high correlations of FDI-SF with vitality, social function, role emotional, and mental health (i.e., psychosocial parameters in general) of SF-36 were observed. This issue has proven again how precisely FDI can reveal the social impact and the impact on the quality of life.

Some limitations of the study should be mentioned. First, two version studies could not be handled. One of the articles was not written in English, and the other was that the full text of the version could not be reached. Second, the study was not registered in the systematic review database like the International Prospective Register of Systematic Reviews (PROSPERO). It would have been more efficient to preserve the compilation protocol's integrity and notify other researchers in this way. Pooling of some parameters could not be performed due to heterogeneous data distribution. Third, in randomized controlled studies using the FDI, the findings of monitoring the efficacy of the treatment could also reveal the responsiveness analysis of the FDI. Finally, individuals with Bell's palsy after surgery were included in the same meta-analysis. This inclusion was due to the small number of available studies. Only considering FDI separately in Bell's palsy may yield more evident results.

Conclusions

FDI was determined to be a suitable questionnaire, especially in terms of internal consistency, test–retest reliability, and construct validity. Particularly, FDI focuses more on psychosocial status regarding our validity pooling analysis. This result reveals that FDI is a unique tool compared to other facial palsy–based measurement tools. In clinical practice, we recommend the use of FDI in addition to clinician-based grading systems in order to see patients' social influences and psychological symptoms more clearly. Further study should focus on the responsiveness and measurement error of the FDI. Finally, in order to observe the psychosocial impact of FDI more comprehensively, it may be a logical methodology to consider the depression and anxiety states of individuals in correlation with FDI.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10072-022-06066-z.

Data availability All data generated or analyzed during this study are included in this published article.

Declarations

Ethical approval Not applicable.

Consent to participate Not applicable.

Conflict of interest The authors declare no competing interests.

References

1. Volk GF, Granitzka T, Kreysa H, Klingner CM, Guntinas-Lichius O (2017) Initial severity of motor and non-motor disabilities in patients with facial palsy: an assessment using patient-reported outcome measures. Eur Arch Otorhinolaryngol 274(1):45–52. https://doi.org/10.1007/s00405-016-4018-1
2. Kleiss IJ, Beurskens CHG, Stalmeier PFM, Ingels KJAO, Marres HAM (2015) Quality of life assessment in facial palsy: validation of the Dutch Facial Clinimetric Evaluation Scale. Eur Arch Otorhinolaryngol 272(8):2055–2061. https://doi.org/10.1007/s00405-015-3508-x
3. VanSwearingen JM, Brach JS (1996) The Facial Disability Index: reliability and validity of a disability assessment instrument for disorders of the facial neuromuscular system. Phys Ther 76(12):1288–1298. https://doi.org/10.1093/ptj/76.12.1288

4. Gonzalez-Cardero E, Infante-Cossio P, Cayuela A, Acosta-Feria M, Gutierrez-Perez J-L (2012) Facial disability index (FDI): adaptation to Spanish, reliability and validity. Medicina oral, patologia oral y cirugia bucal 17(6):e1006–e1012. https://doi.org/10.4317/medoral.18054
5. van Veen MM, Bruins TE, Artan M, Mooibroek-Leeuwerke T, Beurskens CHG, Werker PMN, Dijkstra PU (2020) Health-related quality of life in facial palsy: translation and validation of the Dutch version Facial Disability Index. Health Qual Life Outcomes 18(1):256. https://doi.org/10.1186/s12955-020-01502-0
6. Marsk E, Hammarstedt-Nordenvall L, Engström M, Jonsson L, Hultcrantz M (2013) Validation of a Swedish version of the Facial Disability Index (FDI) and the Facial Clinimetric Evaluation (FaCE) scale. Acta Otolaryngol 133(6):662–669. https://doi.org/10.3109/00016489.2013.766924
7. Pavese C, Cecini M, Camerino N, De Silvestri A, Tinelli C, Bejor M, Dalla Toffola E (2014) Functional and social limitations after facial palsy: expanded and independent validation of the italian version of the facial disability index. Phys Ther 94(9):1327–1336. https://doi.org/10.2522/ptj.20130254
8. Barry P, Mancini J, Alshukry A, Salburgo F, Lavieille J-P, Montava M (2019) Validation of French versions of the Facial Disability Index and the Facial Clinimetric Evaluation Scale, specific quality of life scales for peripheral facial palsy patients. Clin Otolaryngol 44(3):313–322. https://doi.org/10.1111/coa.13294
9. Graciano AJ, Bonin MM, Mory MR, Tessitore A, Paschoal JR, Chone CT (2020) Translation, cultural adaptation and validation of the facial disability index into Brazilian Portuguese☆. Braz J Otorhinolaryngol 86:602–608
10. Özden F, Karaman ÖN, Tuğay N, Savaş Ö, Sözen T, Üçüncü H (2020) The reliability and validity of the Turkish version of the Facial Disability Index. Disability and Rehabilitation:1–10. https://doi.org/10.1080/09638288.2020.1762768
11. Díaz-Aristizabal U, Valdés-Vilches M, Fernández-Ferreras TR, Calero-Muñoz E, Bienzobas-Allué E, Moracén-Naranjo T (2019) Correlations between impairment, psychological distress, disability, and quality of life in peripheral facial palsy. Neurología (English Edition) 34(7):423–428. https://doi.org/10.1016/j.nrleng.2019.03.001
12. Györi E, Przestrzelski C, Pona I, Hagmann M, Rath T, Radtke C, Tzou CHJ (2018) Quality of life and functional assessment of facial palsy patients: a questionnaire study. Int J Surg 55:92–97. https://doi.org/10.1016/j.ijsu.2018.04.061
13. Fujiwara K, Fukuda A, Morita S, Yanagi H, Hoshino K, Nakamaru Y, Furuta Y, Homma A (2021) Psychological evaluation for patients with non-cured facial nerve palsy. Auris Nasus Larynx. https://doi.org/10.1016/j.anl.2021.04.007
14. Paolucci T, Cardarola A, Colonnelli P, Ferracuti G, Gonnella R, Murgia M, Santilli V, Paoloni M, Bernetti A, Agostini F, Mangone M (2020) Give me a kiss! An integrative rehabilitative training program with motor imagery and mirror therapy for recovery of facial palsy. Eur J Phys Rehabil Med 56 (1):58–67. https://doi.org/10.23736/s1973-9087.19.05757-5
15. Steinhäuser J, Volk GF, Thielker J, Geitner M, Kuttenreich A-M, Klingner CM, Dobel C, Guntinas-Lichius O (2022) Multidisciplinary care of patients with facial palsy: treatment of 1220 patients in a German facial nerve center. J Clin Med 11(2):427
16. Henderson LK, Craig JC, Willis NS, Tovey D, Webster AC (2010) How to write a Cochrane systematic review. Nephrology 15(6):617–624. https://doi.org/10.1111/j.1440-1797.2010.01380.x
17. Moher D, Liberati A, Tetzlaff J, Altman DG, The PG (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. PLoS Med 6(7):e1000097. https://doi.org/10.1371/journal.pmed.1000097
18. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A (2016) Rayyan—a web and mobile app for systematic reviews. Syst Rev 5(1):210. https://doi.org/10.1186/s13643-016-0384-4
19. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW (2012) Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 21(4):651–657. https://doi.org/10.1007/s11136-011-9960-1
20. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60(1):34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012
21. Chambers JM (2008) Software for data analysis: programming with R, vol 2. Springer
22. Hedges LV, Pigott TD (2001) The power of statistical tests in meta-analysis. Psychol Methods 6(3):203
23. Schmidt FL (2015) History and development of the Schmidt-Hunter meta-analysis methods. Res synth methods 6(3):232–239
24. Higgins JP, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. Stat Med 21(11):1539–1558
25. Kim MH, Park SY (2021) Population-based study and a scoping review for the epidemiology and seasonality in and effect of weather on Bell's palsy. Sci Rep 11(1):1–9
26. Jones KJ (1993) Recovery from facial paralysis following crush injury of the facial nerve in hamsters: differential effects of gender and androgen exposure. Exp Neurol 121(1):133–138
27. Homeier D (2014) Aging: physiology, disease, and abuse. Clin Geriatr Med 30(4):671–686
28. Seidman SN, Walsh BT (1999) Testosterone and depression in aging men. Am J Geriatr Psychiatry 7(1):18–33
29. Connelly LM (2011) Cronbach's alpha Medsurg nursing 20(1):45–47
30. Portney LG, Watkins MP (2009) Foundations of clinical research: applications to practice, vol 892. Pearson/Prentice Hall Upper Saddle River, NJ,
31. Neely JG, Cherian NG, Dickerson CB, Nedzelski JM (2010) Sunnybrook facial grading system: reliability and criteria for grading. Laryngoscope 120(5):1038–1045
32. Ng JH, Ngo RYS (2013) The use of the facial clinimetric evaluation scale as a patient-based grading system in Bell's palsy. Laryngoscope 123(5):1256–1260