Turkish Validity and Reliability Study of WaLIDD Score

İzzet Fidancı (ORCID 0000-0001-9848-8697), Büşra Yılmaz Demiralay, Hilal Aksoy, Duygu Ayhan Başer (ORCID 0000-0002-5153-2184)

Hacettepe University, Faculty of Medicine, Department of Family Medicine, Ankara, Turkiye

Introduction

The WaLIDD score (Working ability, Location, Intensity, Days of Pain, Dysmenorrhea), developed in line with this need, is a short and applicable measurement tool that addresses the impact of menstrual pain on the individual in a broader framework. The score has been shown to be an effective tool in both diagnosing dysmenorrhea and predicting the consequences of this condition such as loss of workforce by evaluating the distribution area, duration, intensity of pain and the individual's capacity to perform daily functions together (1-7).

This study aims to fill this gap by adapting the WaLIDD score into Turkish and examining its psychometric properties. Thus, it is aimed to provide a practical, culturally appropriate and scientifically sound assessment tool for both field research and clinical applications.

Poster Presentation - PP132

Method

Research Design

This study is a cross-sectional and methodological study conducted online, aiming to evaluate the validity and reliability of the Turkish version of the WaLIDD Score used to

assess dysmenorrhea. The linguistic and cultural adaptation of the scale was carried out in accordance with the translation-adaptation principles recommended by the World Health Organization, and then its psychometric properties were tested.

Permits

Before starting the study (08.01.2025), permission was obtained from the authors who developed the score. Afterwards, permission was obtained from Hacettepe University Health Sciences Research Ethics Committee /Research Number: SBA 25/078; Session date: 18.02.2025; Session number: 2025/05; Decision Number: 2025/05-17), the study was started on 01.03.2025 and completed on 21.05.2025.

Participants and Sample Characteristics

Female individuals living in Turkey, aged 18 and over, with active internet access were included in the study on a voluntary basis. Participants were invited to study through social media channels. Inclusion criteria included: Being over 18 years of age, having regular menstruation in the last 3 months, being between the 6th and 9th day of the menstrual cycle at the time of the survey, not having a diagnosis of chronic disease and not taking regular medication. Pregnancy, postmenopausal period, being within 12 months postpartum, history of pelvic infection and incomplete or contradictory answers were defined as exclusion criteria. The sample size was determined by taking at least 10 times the number of questions in the scale (n = 4); 30% additional sample was targeted due to possible refusal, dropout and exclusion criteria. Thus, it was envisaged to reach at least 52 participants in the study.

Data Collection Process

The data collection process was conducted as an online questionnaire via Google Forms. Participants completed the questionnaire after reading and approving the informed consent text. The questionnaire consisted of four sections:

1. **Demographic Information Form**

   It includes basic variables such as age, marital status, education level, occupation, smoking and alcohol use status of the participants.

2. **WaLIDD Score - English Version**

   The WaLIDD score was first developed by Teherán et al. and is a unique scale that aims to assess dysmenorrhea with a multidimensional approach, not only based on pain intensity. The scale consists of four questions, and each question is scored between 0 and 3 points (total score range: 0-12). The dimensions assessed are:

   - **Ability to work**: The impact of pain on daily functioning.
   - **Pain localization**: The anatomical extent of the pain (e.g. lower abdomen, lower back, leg, groin).
   - **Pain intensity**: Subjective assessment based on the Wong-Baker facial pain scale.
   - **Duration of pain**: Number of days of pain during menstruation.

The WaLIDD score indicates no dysmenorrhea at low scores, mild to moderate dysfunction at moderate scores, and severe dysmenorrhea-related functional impairment at high scores. In the study in which it was developed, it was defined as a diagnostic tool with high discrimination (7).

1. **Numerical Pain Rating Scale (NRS)**

   The NRS is a simple and reliable assessment tool that asks individuals to rate their pain between 0 (no pain) and 10 (unbearable pain). It is widely preferred because it can be administered verbally and is easily understood by the patient (8). In this study, WaLIDD score was used as a reference criterion to test its relationship with pain intensity.

2. **Health-related Quality of Life Scale (SF-36)**

   SF-36 is a self-report scale that evaluates individuals' general health perceptions and quality of life in eight sub-dimensions. The Turkish version was adapted by Koçyiğit et al. and validity and reliability analyses were performed (9). Sub-dimensions: Physical Function (FF); Physical Role Restriction (FRR); Emotional Role Restriction (ERR); Body Pain (VA); Social Function (SF); Vitality (C); Mental Health (MS); General Health Perception (GS)

Each subscale is scored between 0-100, with higher scores indicating better health status. The SF-36 was used to understand the impact of dysmenorrhea on quality of life and to test the functional validity of the WaLIDD score.

Adaptation of WaLIDD Score into Turkish

The linguistic adaptation of the scale was carried out using the forward-backward translation method. In the first stage, translations were made from English to Turkish by two independent translators, and then a common text was created by the researchers. This draft was evaluated by an expert committee in terms of language, content and cultural appropriateness. After the content validity was assessed, the text translated into Turkish was translated back into English by a translator whose native language was English and who was not familiar with the scale; this version was compared with the original author to ensure the integrity of meaning. The pilot study was conducted with 10 women, and the comprehensibility of the statements was tested.

Statistical Analysis Plan

All statistical analyses were performed using SPSS (Version 23.0) and LISREL software. Descriptive analysis of the data (mean, standard deviation, media, quartiles, min-max) were calculated. Internal consistency and reliability of the WaLIDD scale were assessed using Cronbach's alpha coefficient. Construct validity, the relationship between WaLIDD

questions and NRS scores was tested with Spearman correlation analysis. Correlate validity, the relationship between WaLIDD and SF-36 subscales was examined with Spearman correlations. Confirmatory Factor Analysis (CFA), single-factor structure of WaLIDD was tested with LISREL; model fit was evaluated using CFI, TLI, RMSEA and χ^2 statistics. The significance level p< was accepted as 0.05.

Results

The mean age range of the participants was 32.07 (SD=6.58) and the age range was between 22 and 43. It is seen that 70% of the participants have at least one child and the average number of children is 0.70. Data on employment type, smoking and alcohol use were also included in the table; it was observed that the rates of smoking (mean= 1.53) and alcohol use (mean = 1.79) were particularly low. In addition, as of the day of the survey, most of the participants were between the 7th and 8th day of their menstrual cycle (Mean= 7.5 days; min:6, max:9).

The mean WaLIDD total score of the individuals participating in the study was 5.66 ± 4.04 (variance $1.40^2 = 1.96$). This value indicates that the problems experienced by individuals with dysmenorrhea are at a moderate level. In addition, the mean pain intensity measured by the NRS (Numerical Pain Rating) scale was 4.83 ± 1.01, indicating a moderate level of pain experience. Among the sub-dimensions of the SF-36 quality of life scale, the highest mean was found to be 'Mental Health' (69.98 ± 1.45), indicating that individuals felt generally well mentally. However, the meaning of the 'Vitality' sub-dimension was 62.57 ± 2.19, and it can be interpreted that their energy levels may be limited.

Based on the total scores obtained from the WaLIDD scale, participants were categorized into four levels of dysmenorrhea severity. Among the 284 respondents, 10.2% (n = 29) reported a total score of 0, indicating an absence of dysmenorrhea-related symptoms. Approximately one-third of the participants (33.8%, n = 96) fell within the mild symptom range (score: 1–4), while 19.4% (n = 55) were classified as experiencing moderate

dysmenorrhea (score: 5–7). Notably, the highest proportion of respondents (36.6%, n = 104) reported severe symptoms, corresponding to a total score between 8 and 12. This distribution suggests that although a minority of individuals are asymptomatic, most participants experience some level of functional limitation due to menstrual pain, with more than one-third exhibiting symptoms of considerable severity. These findings underscore the importance of routine functional assessment in individuals with dysmenorrhea and highlight the utility of the WaLIDD score as a multidimensional screening tool.

The results of the Spearman correlation analysis between the total score of the WaLIDD scale and the eight sub-dimensions of the SF-36 health questionnaire are presented in Table 1. Correlation coefficients (r) and p values indicate the direction and significance level of the relationship.

Table 1. Correlation between WaLIDD Total Score and SF-36 Subscales

SF-36 Subdimension	Spearman r	p value
Physical Function	0.051	0.395
Physical Role Restriction	-0.108	0.07
Emotional Role Restriction	0.071	0.235
Body Pain	0.063	0.291
Social Function	-0.066	0.268
Mental Health	-0.11	0.065
Vitality	0.049	0.413
General Health	0.038	0.522

According to the results, there was no statistically significant correlation between the total score of WaLIDD and the eight sub-dimensions of SF-36 (all $p > 0.05$). The correlation coefficients were low, and the highest absolute correlation was negative and weak with the Physical Role Restriction subscale ($r = -0.108$, $p = 0.070$). These results suggest that WaLIDD scores do not directly correlate strongly with SF-36 subscales, suggesting that

the WaLIDD reflects the specific experience of menstrual pain, especially compared to the more general health perception components of the SF-36.

Reliability Analysis

There are 3 questions in the WaLIDD scale (k=4). The standard deviation of the total score was calculated as 1.40 and the variance as 1.96. The table below shows the standard deviation and variance of each question (Table 2).

Table 2. Reliability analysis

Article	Standard Deviation	Variance
Question 1	1.0068	1.0136
Question 2	1.1038	1.2184
Question 3	1.0333	1.0677
Question 4	1.0620	1.1278

Total variance: 16.3103

Total variance of questions: 4.4275

When the formula is applied:

$$\sigma_1^2 + \sigma_2^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 = 1.0136 + 1.2184 + 1.0677 + 1.1278 = 4.4275$$

$$\alpha = (4 / 3) * (1 - 4.4275 / 16.3103) = 0.9714$$

The calculated Cronbach's Alpha value is 0.9714. This value shows that the four items of the WaLIDD scale are highly compatible with each other and that the items in the scale strongly measure the same construct (dysmenorrhea-related functioning). In this context, the internal consistency of the WaLIDD Turkish form is very high.

A very high and significant positive correlation ($r = 0.975$, $p < 0.001$) was found between the WaLIDD total score and the Numerical Pain Score (NRS). This suggests that the scale shows high agreement with instruments measuring similar concepts (pain intensity).

Validity Analysis

1. **Construct Validity (Correlation with Numerical Pain Scale)**

In the Spearman correlation analysis between the first three questions of the WaLIDD scale and the NRS; r = 0.7493 with Question 1, r = 0.6638 with Question 2 and r = 0.7077 with Question 3. These results revealed that the relevant items of the WaLIDD accurately reflected the severity of pain and showed a high level of positive correlation with the NRS. A correlation coefficient close to or above 0.70 indicates a strong relationship between the two measurement tools (Table 3).

Table 3. Spearman Correlations between WaLIDD Questions and NRS

WaLIDD Question	Correlation with NRS (r)
Question 1	0.7493
Question 2	0.6638
Question 3	0.7077

1. **Correlative Validity (Correlation with SF-36)**

When WaLIDD Question 4 (effect of pain on activities of daily living) was compared with the 8 sub-dimensions of the SF-36 quality of life scale, the highest correlation was observed with mental health (r = -0.1312) and physical role restriction (r = -0.1083). These negative correlations indicate that as pain increases, individuals' general health perception and functional roles are negatively affected. The low correlation coefficients indicate that WaLIDD measures a different dimension than SF-36 (pain-specific functioning), but the direction of the relationships is in the expected direction (Table 4).

Table 4. Correlations between WaLIDD Question 4 and SF-36 Subscales

SF-36 Subdimension	Correlation (r)
Physical Function (FF)	0.0730

Physical Role Restriction (PRR)	-0.1083
Emotional Role Restraint (ERR)	0.0706
Body Pain (VA)	0.0692
Social Function (SF)	-0.0553
Mental Health (MS)	-0.1312
Vitality (C)	0.0624
General Health (GS)	0.0695

Confirmatory Factor Analysis (CFA)

1. **Scree Plot**

In the Scree Plot analysis, the principal components plot obtained from the four items of WaLIDD showed that the first factor was clearly dominant. This result supports the scale that has a one-factor structure. The scree plot showed that the first factor was clearly dominant on the curve. This indicates that the four items in the scale represent a single construct, and a one-factor model is appropriate.

The Scree Plot above shows the results of Principal Component Analysis (PCA) for the four items of the WaLIDD scale. The proportion of variance explained by the first component is clearly larger than the others, suggesting that: WaLIDD scale has a single-factor structure, and it shows that confirmatory factor analysis (CFA) can be based on a single factor model (Table 5).

Table 5. Confirmatory Factor Analysis (CFA)

Statistics	Value
DoF	2
DoF Baseline	6
χ^2 (Chi-square)	4.2873
χ^2 p-value	0.1172
χ^2 Baseline	1459.4985
CFI (Comparative Fit Index)	0.9984

The fit statistics obtained because of Confirmatory Factor Analysis (CFA) were interpreted in various aspects to evaluate the fit of the model to the data. First, the Degrees of Freedom (DoF) of the model was determined as 2. This value indicates that the model is predictable and has sufficient degrees of freedom. The degrees of freedom represent the difference between the number of parameters included in the model and the observable information in the data structure, and a positive value is a prerequisite for the testability of the model.

The chi-square (χ^2) value calculated to test the fit of the model is 4.2873. This value measures the difference between the observed covariance matrix and the covariance matrix estimated by the model. The p-value, which is the significance level for the chi-square statistics, is 0.1172. A p-value greater than 0.05 indicates that there is no statistically significant difference between the model and the observed data, thus the model fits the data well. In other words, in this case, the hypothesis is not rejected, and the model is accepted.

In the analysis, statistics related to the baseline model were also evaluated. The degrees of freedom (DoF Baseline) of the comparative model are 6 and the chi-square value of this model is calculated as 1459.4985. This model is an independent model that is established with the assumption that there is no relationship between the variables. How well the default model performs compared to this comparative model is evaluated by the CFI (Comparative Fit Index).

The CFI value is an important index that ranges between 0 and 1 and measures model fit and was obtained as 0.9984 in this study. A CFI value above 0.95 indicates that the model provides a very good fit, while values above 0.90 indicate an acceptable fit. In this context, it can be said that the model created for the WaLIDD scale shows a very strong fit.

In general, when all these fit statistics are evaluated together, it is concluded that the one-factor model established for the WaLIDD scale has high validity both statistically and

theoretically. These findings support the structural integrity and factor structure of the scale.

1. Factor Loadings and Variance Explained

WaLIDD items were CFAed with a single factor model and factor loadings, explained variance (R^2) and error variances of each item were calculated (Table 4). The factor loadings calculated for each item of WaLIDD were quite high (between 0.91-0.96). Most of the variance of each question (R^2: 83-92%) can be explained by the factor. For example, 92.03% of the variance of Question 2 was explained by the general dysmenorrhea factor. (Table 7).

1. **Model Fit Indices**

The fit statistics of the one-factor CFA model are presented in Table 5. The fit indices show **that** the model **fits** the data **very well**. The model fit indices calculated as CFI = 0.998, TLI = 0.995 and RMSEA = 0.0636 show that the CFA model fits the data very well. CFI and TLI values above 0.95 indicate strong structural validation of the model. Since the RMSEA value is below 0.08, it shows that the model is compatible with an acceptable margin of error (Table 6).

Table 6. CFA Model Fit Statistics

Statistics	Value	Description
Chi-square	4.29	p = 0.117 (>0.05)
CFI	0.998	Perfect fit
TLI	0.995	Perfect fit
RMSEA	0.0636	Acceptable fit
AIC / BIC	Low	Advantages in model comparison

When the structural features in the model created because of CFA were evaluated, the fit indices (CFI = 0.998, TLI = 0.995, RMSEA = 0.0636) revealed that the model provided a

very good fit to the data. The χ^2 value of the model was 4.29 and the p-value was 0.117, and this result, which was not statistically significant, also supported the good fit of the model.

In conclusion, confirmatory factor analysis reveals that the Turkish version of the WaLIDD scale has a single-factor, strong and reliable structure, and its questions represent the common factor at a high level. Thanks to this structure, WaLIDD can be used as a powerful tool to assess the impact of dysmenorrhea on quality of life.

As seen in the model, the factor loadings of all question range between 0.91 and 0.96 and show quite strong relationships. The low error variances support the adequacy of the questions to represent the common factor. This structure shows the single-factor and highly coherent structure of the scale.

The data obtained show that the Turkish version of WaLIDD has strong psychometric properties in terms of both construct validity and internal consistency. The high correlation with numerical pain assessments reveals that the scale accurately reflects the dysmenorrhea-specific condition. The correlations with SF-36 sub-dimensions also support the effect of pain on quality of life. The very high fit of the model with CFA analysis reinforces the structural integrity of the scale and its success in representing the concept it is intended to measure.

The findings revealed that WaLIDD Turkish form: It has high internal consistency, is strongly correlated with the numerical pain score (NRS), shows conceptually significant and directionally expected relationships with the quality of life (SF-36) scale, and its single-factor structure was confirmed.

The WaLIDD score is capable of reliably and validly assessing dysmenorrhea-related functional limitations in the Turkish-speaking population.

Discussion

In this study, the validity and reliability of the Turkish version of the WaLIDD Score, which aims to measure dysmenorrhea-specific functional effects multidimensionally, were evaluated. The Cronbach's alpha value (.97) obtained for the internal consistency of the scale indicates a high level of homogeneity among the questions of the scale. This finding is consistent with the reliability level reported in the study in which the scale was first developed (α = .88) and similarly, strong internal consistency values (α = .91) were obtained in the Arabic validity study (7,10).

The high correlation between WaLIDD and NRS suggests that these two instruments measure similar constructions. In particular, the positive correlation between the first three questions of the WaLIDD and the NRS, which measures pain intensity, supports that the scale accurately reflects the individual experience dimension of pain. In the literature, it is emphasized that unidimensional scales such as the NRS are inadequate in the evaluation of dysmenorrhea and that measurement tools including functional effects should be preferred instead (5).

It is noteworthy that there were generally weak correlations between WaLIDD and eight sub-dimensions of SF-36. Negative and low correlations were observed especially with some sub-dimensions such as physical role limitation and mental health. These findings suggest that WaLIDD measures a more specific dimension of functionality specific to dysmenorrhea and that general health questionnaires such as SF-36 may be insufficient to directly assess this area (9).

Confirmatory factor analysis results showed that the Turkish version of the WaLIDD Score has a single-factor structure as in the original version. The high factor loadings (.91-.96) and the very good fit of the model to the data (CFI = .998, RMSEA = .0636) indicate that the scale works in accordance with the theoretical foundations. This is in line with similar fit values in the Arabic validity study (10).

The mean WaLIDD score was found to be moderate (5.66), indicating that a significant proportion of the women who participated in the study experienced functional effects related to dysmenorrhea. In previous field studies conducted in Turkey, it has been reported that dysmenorrhea is common among young women and has significant effects on quality of life (2,3). In this context, the WaLIDD Score can be considered as a short, applicable and clinically meaningful tool that can be used in dysmenorrhea evaluations.

The strengths of the study include the adequate sample size, the systematic language adaptation process, and the use of advanced statistical methods such as confirmatory factor analysis. However, the fact that the study was conducted online may cause homogeneity in the sample, and the inclusion of only Turkish-speaking individuals may limit cultural generalizability. Retrospective reporting of a subjective experience such as pain is also one of the possible biases.

Conclusion

The Turkish adapted version of the WaLIDD Score can be considered as a valid and reliable measurement tool for assessing dysmenorrhea-specific functional effects. Thanks to its strong correlation with pain severity and its single-factor, holistic structure, it can be used safely both in clinical practice and in research settings. Its addition as a complement to general health scales such as SF-36 may contribute to the multidimensional assessment of women's health.

Keywords: Dysmenorrhea, Walıdd Score, Validity, Reliability, Confirmatory Factor Analysis, Quality Of Life, SF-36

References

1. Osayande AS, Mehulic S. Diagnosis and initial management of dysmenorrhea. Am Fam Physician. 2014;89(5):341–346.
2. Ju H, Jones M, Mishra G. The prevalence and risk factors of dysmenorrhea. Epidemiol Rev. 2014;36:104–113. doi: 10.1093/epirev/mxt009
3. Unsal A, Ayranci U, Tozun M, Arslan G, Calik E. Prevalence of dysmenorrhea and its effect on quality of life among a group of female university students. Ups J Med Sci. 2010;115(2):138–145. doi: 10.3109/03009730903457218
4. Al-Jefout M, Seham A, Jameel H, et al. Dysmenorrhea: prevalence and impact on quality of life among young adult Jordanian females. J Pediatr Adolesc Gynecol. 2015;28:173–185. doi: 10.1016/j.jpag.2014.07.005
5. Chen CX, Kwekkeboom KL, Ward SE. Self-report pain and symptom measures for primary dysmenorrhea: a critical review. Eur J Pain. 2015;19(3):377–391. doi: 10.1002/ejp.556
6. De Sanctis V, Soliman A, Bernasconi S, et al. Primary dysmenorrhea in adolescents: prevalence, impact and recent knowledge. Pediatr Endocrinol Rev. 2015;13(2):512–520.
7. Teherán AA, Piñeros LG, Pulido F, Mejía Guatibonza MC. WaLIDD score, a new tool to diagnose dysmenorrhea and predict medical leave in university students. Int J Womens Health. 2018;10:35–45. doi:10.2147/IJWH.S143510.
8. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. Ann Rheum Dis. 1978;37(4):378–381. doi: 10.1136/ard.37.4.378
9. Koçyiğit H, Aydemir Ö, Fisek G, Ölmez N, Memiş A. Kısa Form-36 (KF-36)'nın Türkçe versiyonunun güvenilirliği ve geçerliliği. İlaç ve Tedavi Dergisi. 1999;12(2):102–106.
10. Bou Haidar PS, Alam DN, Obeid S, et al. Validation of the Working Ability, Location, Intensity, Days of Pain and Dysmenorrhea (WaLIDD) Scale in Arabic-Speaking Young Female Adults. Int J Womens Health. 2024;16:1565–1571. doi:10.2147/IJWH.S468803.