RESEARCH ARTICLE

Sevil Aydogan Gedik[1]
Didem Arslantas[2]
Alaettin Unsal[2]
Melih Velipasaoglu[3]

[1] Eskişehir Provincial Health Directorate, Eskişehir, Türkiye
[2] Public Health Department, Eskişehir Osmangazi University Faculty of Medicine, Eskişehir, Türkiye
[3] Gynecology and Obstetrics Department, Eskişehir Osmangazi University Faculty of Medicine, Eskişehir, Türkiye

Corresponding Author:
Sevil Aydogan Gedik
mail: aydogan.sevil@gmail.com

Received: 24.12.2024
Acceptance: 17.03.2025
DOI:10.18521/ktd.1606373

Scientific meetings where the study is presented as a paper:
1) 4th International 22nd National Public Health Congress, 13-19 December 2020.
2) 5th International and 23rd National Public Health Congress, 13-18 December 2021.

Konuralp Medical Journal
e-ISSN1309–3878
konuralptipdergi@duzce.edu.tr
konuralptipdergisi@gmail.com
www.konuralptipdergi.duzce.edu.tr

Development of the Aydoğan-Depression Screening Scale for Pregnant and Determination of Depression Risks of Pregnant Women

ABSTRACT

Objective: The aim of the study is to develop the "Aydoğan-Depression Screening Scale for Pregnant" (A-DSP) for depression screening during pregnancy and to test its validity and reliability.

Method: This methodological study was conducted with 369 pregnant. A-DSP was designed as a 4-point Likert-type self-report scale consisting of positive and negative propositions. Content, construct and criterion validity were evaluated. Internal consistency analyses, item analysis and test-retest were performed to evaluate reliability. The cut-off score was determined by ROC analysis.

Results: The results obtained from all validity and reliability analyses of A-DSP were at a sufficient level. It was found that A-DSP consisted of 4 sub-dimensions and 21 items, total explained variance was 56.3%, and Cronbach's Alpha was 0.919. An increase in scores indicates an increase in the suspicion of depression. In addition, it is accepted that there is a suspicion of depression at a score≥41. It was determined that 29.5% of the pregnant women had a suspicion of depression.

Conclusions: It was concluded that A-DSP is a valid and reliable scale that can be used to screen for depression in pregnant women. It is thought that it would be beneficial to screen for depression using A-DSP, to monitor the mental status of pregnant women by A-DSP, to perceive it as an early warning for depression when there is a change, and to refer to a psychiatrist for further examination.

Keywords: Depression, Pregnant, Pregnancy, Validity, Reliability, Screening.

Aydoğan-Gebelere Yönelik Depresyon Tarama Ölçeği'nin Geliştirilmesi ve Gebelerin Depresyon Düzeyinin Belirlenmesi

ÖZET

Amaç: Çalışmanın amacı gebelik döneminde görülen depresyon taraması için "Aydoğan - Gebelere Yönelik Depresyon Tarama Ölçeği"nin (A-GDÖ) geliştirilmesi, geçerlik ve güvenirliğinin test edilmesidir.

Yöntem: Metodolojik tipteki çalışma 369 gebe üzerinde gerçekleştirildi. A-GDÖ 4'lü Likert tipinde olumlu ve olumsuz önermelerden oluşan bir öz bildirim ölçeği olarak tasarlandı. Kapsam geçerliği, yapı geçerliği ve ölçüt geçerliği değerlendirildi. Güvenirliği değerlendirmek için iç tutarlık analizleri, madde analizi ve test - tekrar test uygulaması yapıldı. Kestirim puanı ROC analizi ile belirlendi.

Bulgular: A-GDÖ'nün tüm geçerlik ve güvenirlik analizlerinden elde edilen sonuçların yeterli düzeyde olduğu görüldü. A-GDÖ'nün 4 alt boyut ve 21 maddeden oluştuğu, toplam açıklanan varyansın % 56.3, Cronbach Alfa güvenirlik katsayısının 0.919 olduğu bulundu. Ölçekten alınan toplam puanın artışı depresyon şüphesinin arttığını göstermektedir. Ayrıca 41 puan ve üzeri depresyon şüphesinin var olduğunu göstermektedir. Çalışmada gebelerin % 29.5'inde depresyon şüphesi olduğu tespit edildi.

Sonuç: A-GDÖ'nün gebelerde depresyonun taranması amacıyla kullanılabilecek geçerli ve güvenilir bir ölçek olduğu sonucuna ulaşıldı. A-GDÖ kullanılarak depresyon taraması yapılması, gebelerin ruhsal durumlarının A-GDÖ ile izlenmesi, bir değişiklik olduğunda depresyon için erken uyarı olarak algılanması ve ileri inceleme için bir psikiyatri uzmanına yönlendirme yapılmasının faydalı olacağı düşünülmektedir.

Anahtar Kelimeler: Depresyon, Gebe, Gebelik, Geçerlik, Güvenirlik, Tarama

INTRODUCTION

Depression is a common mental disorder in pregnant women. It has been reported that approximately 10% of pregnant women worldwide have depression (1). Depression has short and long term negative effects on pregnancy, fetus and after delivery baby (2). Early diagnosis of depression during pregnancy is important in order to prevent the negative consequences (3,4).

Depression screening in pregnant women is a topic that has come to the fore recently, and it is recommended to perform routine screenings with easy-to-use tools in prenatal controls (2, 5). With the help of self-report questionnaires that can be administered by people who are not experts in mental health, it is possible to identify pregnant women at risk of depression with low costs and using less resources (6). In particular, it is important to carry out screenings by health personnel in primary care and to include psychological evaluation in the follow-up of pregnant women (2, 7).

Although the validity and reliability studies of some depression scales in pregnant have positive results or some questions for screening depression during pregnancy have been found appropriate, no scale has been found in the literature that includes pregnancy-specific questions developed only for depression seen during pregnancy. There is a lack of approved, valid and reliable screening tools to screen for depression during pregnancy (2, 8). Because it may be difficult to distinguish normal somatic and emotional symptoms of pregnancy from depression symptoms, the use of unconfirmed measurement tools may yield inaccurate results (9). Reliable tools are needed to detect pregnancy depression.

In the study, it was aimed to develop the "Aydoğan-Depression Screening Scale for Pregnant" (A-DSP) to screen for depression during pregnancy, to test its validity and reliability, and to evaluate the suspicion of depression in pregnant women.

MATERIAL AND METHODS

Study Design and Study Group: The study is a methodological type study conducted on pregnant women aged 18 and over who applied to Eskişehir Osmangazi University Health Practice and Research Hospital Gynecology and Obstetrics Polyclinic between March 2020 and November 2021.

In validity and reliability studies, reaching 5-10 times the number of items or a sample of 300 people is considered good (10, 11). In our study, as the main sample, validity and reliability analyzes were performed on 369 pregnant women. Confirmatory factor analysis was performed in another group consisting of 308 pregnant women.

The ages of the women, who constituted the main sample, ranged between 18-43, with a mean of 30.49±4.73. Gestational weeks ranged from 4 to 40, with a mean of 23.15±9.52. 44.7% (n=165) were in the second trimester.

Data Collection Tools: Data were collected by a questionnaire. In the first part of the questionnaire, there are questions about the sociodemographic characteristics, questions about pregnancy and some factors that may be related to depression. In the second section, there is the A-DSP, which will be developed for the purpose of screening for depression in pregnant women, and the Edinburgh Postnatal Depression Scale (EDS) in the third section.

EDS was developed for the recognition of postpartum depression. The Turkish validity and reliability study was performed by Aydın et al. The cut-off point was suggested as 12.5 (12). Validity and reliability studies have been conducted on pregnant women in many countries. It has been used in many studies conducted to evaluate depression in pregnant women in Turkey.

Creating the A-DSP: In order to develop A-DSP, various psychiatry books, DSM, scales used for depression screening and diagnosis, depression scales used in research on depression in pregnant women, and publications specific to depression during pregnancy were examined. An 88-item question pool was created, including the diagnosis criteria of depression and the symptoms of depression in pregnant women.

Six experts (a psychiatrist, a psychologist, a measurement/evaluation expert, a gynecologist and two public health experts who are competent in scale development) reviewed the 88-item pool of questions. It was evaluated whether the items represented the features to be measured and whether they contained information about depression during pregnancy. Highly repetitive and misleading items were eliminated. In addition, the items were also reviewed in terms of language. After the corrections were made, 53 questions remained in the question pool.

A-DSP was designed as a self-report scale. In the A-DSP, which consists of positive and negative propositions, pregnant women were asked to think about their mental state in the last 1 week and to mark how often they experienced/thought/felt the expression in each item. The scale questions designed as an ordinal scale and designed as a 4-point Likert type are answered as "never", "sometimes", "often" and "always". Responses to negative items were never=1, sometimes=2, often=3, always=4; responses to positive items were scored in the opposite way. It was accepted that the risk of depression increased as the total score that could be obtained from the scale increased.

Language Suitability and Content Validity: Expert opinion was sought to evaluate the language suitability and content validity of the A-DSP. The 53-item form was submitted to the

opinion of 22 experts. The items were reviewed in terms of language. It was checked whether the items were understandable, whether there were any errors in their meaning, whether the items expressed the desired thing correctly and clearly. Experts evaluated each item according to the options of "necessary and sufficient", "necessary but insufficient" and "unnecessary". After the expert evaluation, suggested corrections were made for the items evaluated as "necessary but insufficient". The content validity ratio (CVR) was calculated for each item and the content validity index (CVI) for the overall scale was calculated. 15 items with a CVR lower than 0.42 were removed from the scale. Thus, 40 questions remained in the form and the smallest recalculated CVR was 0.45, and the CVI was 0.789. The content validity of the scale was considered to be sufficient (13).

Preliminary Study: Incomprehensible and erroneous questions were corrected with the preliminary study applied to 21 pregnant women. As a result, 7 items were removed from the scale.

Pilot Study: At this stage of the study, since the number of items in the A-DSP was 33, it was aimed to apply a questionnaire to at least 165 pregnant women for the pilot study by taking 5 times the number of items (11). Draft scale was applied to 200 pregnant women. Internal consistency analyzes and item analysis were performed. 11 items with an item-total correlation coefficient below 0.30 were excluded from the scale (14). As a result of the pilot study, A-DSP decreased to 22 items.

Validity and Reliability Analysis with the Main Sample: After the pilot study, the main sample (n=369) was surveyed again. To test the construct validity, exploratory factor analysis (EFA, it was done in a different group of 308 pregnant women) and confirmatory factor analysis (CFA) was performed, and discriminant validity was tested with the help of differences between groups. Assumptions were checked before performing factor analysis. For factor analysis in the literature, reaching 5-10 times the number of items or a sample of 300 people is considered good (10, 11). In line with these recommendations, a sample of 369 pregnant women for EFA and 308 pregnant women for CFA is considered sufficient. In addition, the Kaiser-Meyer-Olkin (KMO) test and Bartlett test of sphericity were used to assess the adequacy of the sample size and suitability for factorization. The correlation matrix was examined to determine whether there was singularity and multiple collinearity. The anti-image correlation matrix and Measures of Sampling Adequacy (MSA) values were also examined for suitability for factor analysis. Direct oblimin rotation were used in EFA. Factors with an eigenvalue greater than 1 were taken into account when deciding on the number of factors (15, 16). The factor load limit value was 0.32 (15). According to the results of EFA made with the 22-item version of A-DSP, it was seen that an item fell into the same sub-dimension with questions that didn't measure the same feature as itself. For this reason, that item was removed from the scale. EFA was performed again with the remaining 21 items. Lavaan 0.6-7 package was used in R program for CFA. Before starting CFA, sample size, normality, singularity, and multiple collinearity assumptions were checked. The MVN 5.8 package was used to test the multivariate normal distribution. Evaluation was made with the Q-Q chart and Mardia's multivariate normality test. In the Q-Q plot of A-DSP, it was seen that many data points deviated from multivariate normality (Figure 1).

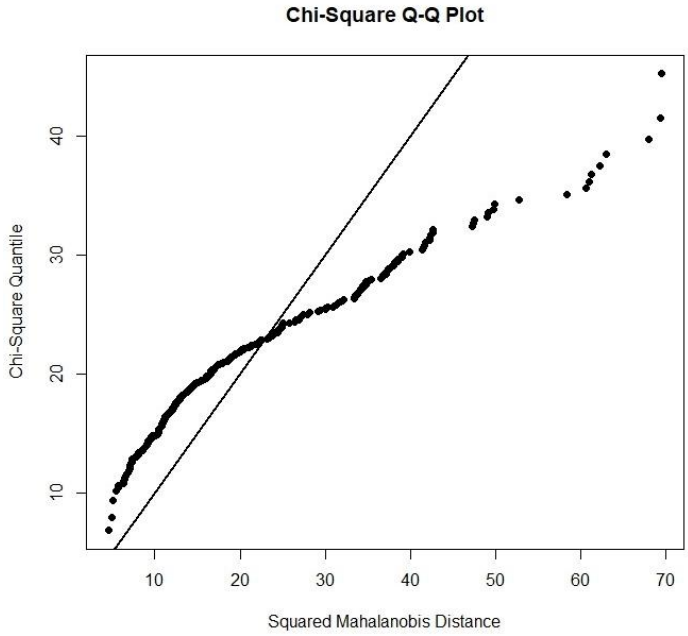

Figure 1. Q-Q Plot Obtained as a Result of Multivariate Normality Analysis

In addition, when Mardia's skewness and kurtosis values, the significance levels of these values and the statistical decision regarding the significance levels were examined, it was seen that multivariate normality was not achieved (Table 1).

Table 1. Mardia's Multivariate Normality Test Results

Test	Statistics	p value	Statistical decision
Skewness	4506.74	<0.001	No
Kurtosis	37.30	<0.001	No
MVN	-	-	No

Since the items of A-DSP were ordinal and multivariate normality could not be achieved, the diagonal weighted least squares method was used as the parameter estimation method. Robust versions of the goodness-of-fit indices obtained by adjusting for non-normal distributions were taken into account. Acceptable value for chi-square/degrees of freedom was considered as <5, for CFI and TLI (NNFI) as >0.95, for RMSEA and SRMR as <0.08 (17). The Spearman correlation coefficient between EDS and A-DSP was evaluated for concurrent criterion validity. Cronbach's Alpha reliability coefficient and item-total correlation coefficient were calculated for internal consistency. The item scores of the lower and upper 27% groups were compared for item discrimination. The test-retest method was used to evaluate stability, which is a component of reliability. The scale was applied to 26 pregnant women with an interval of 2 weeks, Pearson correlation coefficient and ICC value were calculated. ROC analysis was performed to determine the cut-off score corresponding to optimal sensitivity and optimal specificity of A-DSP. In the ROC analysis, groups with and without risk for depression were used according to the cut-off score of EDS, which is used as an equivalent criterion.

Data analysis was performed using SPSS (version 15.0) and R (version 4.0.3) statistical packages. Mann-Whitney U and Chi-square tests were used to compare the groups. Statistical significance level was accepted as $p<0.05$.

Ethical Approval: Ethics committee approval was obtained from Eskişehir Osmangazi University Non-Interventional Clinical Research Ethics Committee (25403353-050-99-E.38314, 26.03.2020).

RESULTS

Exploratory Factor Analysis: KMO test results were 0.934, and Bartlett's p-value was<0.001. The sample size was sufficient for factor analysis and the data were suitable for analysis. The correlation coefficient between all items was observed to be less than 0.8. Thus, it was concluded that there was no singularity and multiple collinearity in the data. The MSA values of the items in the anti-image correlation matrix were between 0.897 and 0.961. It was observed that the values outside the diagonal of the matrix were mostly small. Since all MSA values were above 0.5, it was concluded that the data could be factored.

As a result of factor analysis, it was seen that the scale consisted of 4 sub-dimensions and 21 items. Factor loadings ranged from 0.428 to 0.798. The total explained variance was 56.3% (Table 2).

Internal Consistency Reliability and Item Analysis: The Cronbach's Alpha coefficient of the 21-item final version of the A-DSP, was found to be 0.919. Item-total correlation coefficients ranged from 0.403 to 0.726. The Cronbach's Alpha value was found to be 0.860 for the first factor, 0.784 for the second factor, 0.698 for the third factor, and 0.765 for the fourth factor. All of the item-total correlation coefficients between each factor's own items were greater than 0.3 (Table 2). Internal consistency of the A-DSP and all factors was considered to be sufficient.

In order to evaluate item discrimination, the scores obtained from the A-DSP were ordered from high to low. The item medians of the lower and upper 27% groups were compared. A significant difference was found between the total scores of the upper 27% group and the lower 27% group from A-DSP and between the scores they received from each item ($p<0.001$ for each). It was accepted that each item of the A-DSP and the whole scale had item discrimination, and that it could distinguish pregnant women with and without depression risk.

Confirmatory Factor Analysis: CFA assumptions were checked and it was found that the sample size (308) was sufficient and there was no singularity and multicollinearity (all coefficients in the correlation matrix are less than 0.8).

The chi-square test value of the model is 475.099 ($p<0.001$). Among the goodness of fit indices obtained by CFA, the chi-square/degrees of freedom value of 2.56, CFI of 0.962 and TLI (NNFI) of 0.957 indicate a very good fit. The SRMR value of 0.060 and the RMSEA value of 0.071 indicate acceptable fit. It was found that sufficient model-data fit was achieved (17). It was determined that the standard regression coefficients (factor load) were sufficient (between 0.59-0.92, Figure 2). It was concluded that the A-DSP provided construct validity.

Criterion Validity: The total A-DSP scores of the pregnant women were between 21-70 (mean=36.15±10.11), and their EDS scores were between 0-26 (mean=8.40±5.54). A strong positive correlation was determined between EDS and A-DSP scores ($r=0.810$, $p<0.001$).

Discriminant Validity: It was found that those who scored 13 points (cut-off point) or higher on the EDS and who reported having a physician-diagnosed mental illness or depression scored higher on the A-DSP (Table 3). It was determined that the scale provided discriminant validity.

Table 2. The Final Factor Pattern of A-DSP

Factors	Items*	For All Scale		For Each Factor	
		Item-Total Correlation Coefficient	Factor Loading Values of Items	Cronbach Alpha if Item Deleted	Item-Total Correlation Coefficient
Factor 1: Low Energy **Initial Eigenvalue:** 8.239 **Common Factor Variance:** 39.234 **Cronbach's Alpha:** 0.860	1. It is physically and mentally difficult to devote myself to a job in my daily life.	0.566	0.798	0.841	0.621
	2. During pregnancy, my life energy decreased.	0.698	0.741	0.823	0.736
	3. I no longer enjoy the things I used to enjoy doing in my spare time before pregnancy.	0.603	0.652	0.842	0.615
	4. I feel that I do not have the strength to strive for something.	0.726	0.616	0.828	0.714
	5. I don't feel like doing anything.	0.590	0.602	0.845	0.589
	6. I have no desire to meet people.	0.472	0.581	0.859	0.494
	7. I'm not as cheerful as I used to be.	0.676	0.472	0.840	0.625
Factor 2: Pessimism **Initial Eigenvalue:** 1.257 **Common Factor Variance:** 5.986 **Cronbach's Alpha:** 0.784	8. My postpartum responsibilities scare me.	0.521	0.704	0.730	0.597
	9. I am hopeful for postpartum.	0.434	0.643	0.787	0.449
	10. I think that I will not be as productive as before in my life after birth.	0.653	0.624	0.704	0.676
	11. I'm afraid of not being able to take good care of my baby.	0.578	0.536	0.741	0.571
	12. I feel like everything will get worse as my pregnancy progresses.	0.646	0.447	0.752	0.536
Factor 3: Worthlessness-Guilt **Initial Eigenvalue:** 1.184 **Common Factor Variance:** 5.638 **Cronbach's Alpha:** 0.698	13. My life has no meaning.	0.444	0.708	0.661	0.451
	14. I think I have a negative impact on my baby's health.	0.523	0.684	0.607	0.523
	15. I think I am worthless from the perspective of the people around me.	0.533	0.669	0.606	0.524
	16. I feel guilty for things that went wrong during pregnancy.	0.484	0.541	0.653	0.456
Factor 4: Depressed Mood **Initial Eigenvalue:** 1.147 **Common Factor Variance:** 5.463 **Cronbach's Alpha:** 0.765	17. I feel like crying for no reason.	0.403	0.632	0.764	0.406
	18. I feel sad.	0.603	0.584	0.702	0.619
	19. I am having a happy pregnancy.	0.566	0.554	0.734	0.534
	20. I always think of bad possibilities regarding the pregnancy process.	0.605	0.536	0.713	0.563
	21. I think my mood is worse than other pregnant women.	0.659	0.428	0.698	0.609
Total Explained Variance: 56.321 **Total Cronbach's Alpha:** 0.919					

*Items 9 and 19 are reverse coded.

Test-Retest Reliability: The total scores from the first test ranged from 22 to 50 (mean=35.42±6.71). The total scores from the second test ranged from 21 to 55 (mean=34.53±8.11). The Pearson correlation coefficient between the total A-DSP scores obtained from the first test and the second test was 0.745 ($p<0.001$) and the ICC value was 0.845 (95% CI:0.655-0.931, $p<0.001$). It was concluded that A-DSP gave similar results in both measurements, had high reliability, was stable and didn't change over time.

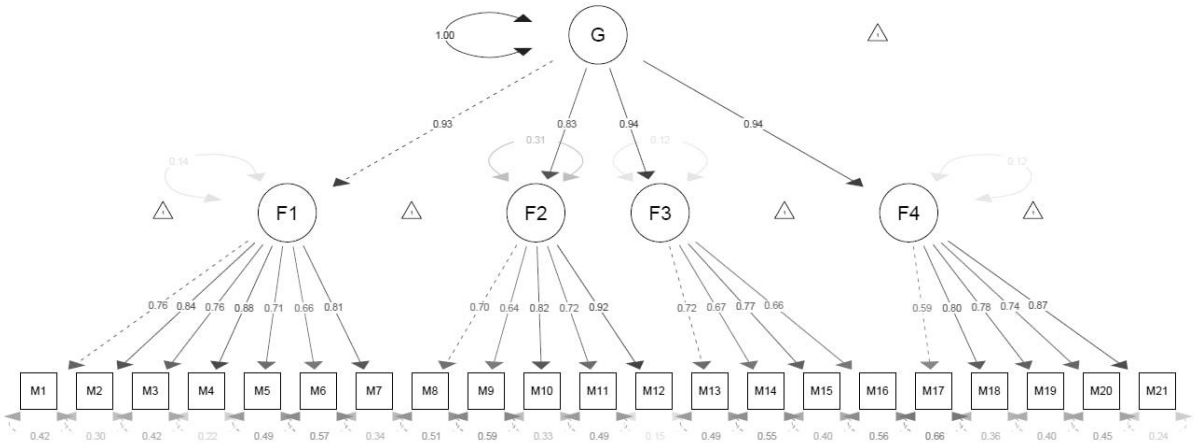

Figure 2. Path Diagram Showing the Model Structure and Standard Regression Coefficients

Table 3. Distribution of the A-DSP Scores of the Pregnant Women According to the EDS Scores and the Presence of Current Mental Illness

	n (%)	A-DSP Total Score		Test Statistic; p
		Mean±SD*	Median (minimum-maximum)	
EDS Score				
12 and below	282 (76.4)	32.33±6.99	31.50 (21.00-60.00)	**12.206;**
13 and above	87 (23.6)	48.55±8.63	47.00 (29.00-70.00)	**<0.001**
Current physician-diagnosed mental illness				
No	360 (97.6)	35.85±9.80	34.00 (21.00-70.00)	**3.046;**
Yes	9 (2.4)	49.44±13.68	53.00 (21.00-64.00)	**0.002**
Current physician-diagnosed depression				
No	365 (98.9)	36.02±10.04	34.00 (21.00-70.00)	**2.259;**
Yes	4 (1.1)	48.75±10.14	47.50 (39.00-61.00)	**0.024**
Total	369 (100.0)	36.1±10.11	34.0 (21.0-70.0)	

*Standard deviation

Cut-Off Score: In the ROC analysis, the area under the curve was found to be 0.932 (%95 CI: 0.905-0.959, p<0.001). The points where the sensitivity and specificity values were highest and closest to each other were examined. The optimal sensitivity (0.851) and specificity (0.876) values were found to be 40.5 cut-off points. In addition, the likelihood ratio (LR) value for this cut-off score was found to be 6.8.

When the EDS and A-DSP cut-off scores were examined, it was determined that 23.6% of the pregnant women according to the EDS and 29.5% according to the A-DSP were at risk of depression.

DISCUSSION

A-DSP was designed as a self-report scale based on the self-evaluation of pregnant women. Self-report scales are used to measure features that cannot be observed directly. Evaluation is made according to the person's statement. The fact that the individual answers the questions honestly affects the accuracy and reliability of the data obtained. A self-report-based screening test cannot replace clinical diagnosis, but it can show which pregnant women need further evaluation (18).

A comprehensive literature review was conducted to create the A-DSP. General information on depression, scales, and publications specific to depression during pregnancy were reviewed. Information on peripartum depression, where pregnancy depression was first defined, and information on postpartum depression were compiled. One of the scales used was the Pregnancy Depression Scale (PDS), which was created for the purpose of screening for depression in pregnant women. It was created by revising the Hamilton Depression Rating Scale using a structured clinical interview for DSM-4. It was determined that 7 items of the Hamilton Depression Rating Scale were associated with depression during pregnancy (depressed mood, feeling of guilt, decrease in work activities, psychomotor retardation, diurnal variation, fatigability, social withdrawal). It was stated that these 7 items forming the PDS predicted a major depressive episode during pregnancy (3). Items that question these symptoms are also found in A-DSP. Only diurnal variation is not included in the A-DSP. Unlike PDS and other depression scales used in

pregnancy, A-DSP questions were created using expressions specific to pregnancy.

It was seen that the A-DSP consisted of 4 factors. Because psychological characteristics have complex structures, it is generally not possible for scales measuring psychological characteristics to be unidimensional. Depression scales also measure the emotional, cognitive, somatic and perceptual symptoms of depression. Within the framework of these symptoms, it is expected that the scales will consist of sub-dimensions. It has been reported in many studies that depression scales consist of many sub-dimensions. Beck et al. defined the cognitive and somatic-affective dimensions of the Beck Depression Inventory. This structure was also confirmed in Turkey (19). It has been reported that the CES-D has a four-dimensional structure: depressive symptoms, positive affect, somatic symptoms, and difficulties in interpersonal relationships (20). Similarly, when the validity and reliability studies of depression scales on pregnant women are examined, it is seen that there are multidimensional structures. It has been reported that EDS, which was developed as one-dimensional, showed a three-factor structure including depression, anxiety and suicide in studies conducted in England and the Netherlands (21,22). In another study conducted in England, the existence of a two-factor structure, anxiety and depression, was mentioned in the first trimester (23). In a study conducted in France, it was reported that a two-factor structure was detected, including depression and other disorders including anxiety (24). In the validity and reliability study performed on Hungarian pregnant women, it was reported that EDS consisted of 3 factors (25). The multidimensional structure of A-DSP in our study is compatible with the literature.

There are different limits in the literature for the total variance explained by the scale. Having 50% or more of the total variance explained by a scale has been accepted as sufficient in many studies (15). It is considered sufficient that the total variance explained in the scales used in social areas is 50-60% (16). According to the data obtained as a result of EFA, the contribution of 21 items and 4 factors that make up A-DSP to the total variance is 56.3%. In other words, approximately 56.3% of the depression risks of pregnant women can be determined with the help of A-DSP. It can be said that the total variance level explained by A-DSP is sufficient.

The similarity between the measurement results of the newly developed test and the standard test, which is known to measure a feature correctly and proven validity and reliability, shows that the new scale provides criterion validity (26). In the hypothesis established in this direction, it was expected that the scores of the pregnant women in A-DSP and the scores they got in the EDS would show a positive and acceptable correlation. As expected, a strong positive correlation was found between the scores obtained from the two scales used in our study (r=0.810). The results obtained show that the A-DSP provided the criterion validity.

Reliability shows the ability of the scale to measure accurately and its invariance over time (14). In order to ensure reliability, the scale should be consistent, stable and sensitive. The Cronbach Alpha reliability coefficient is used to evaluate internal consistency. It shows the degree of consistency between the items of a scale and the whole scale. High values indicate that the scale items are self-consistent and that the scale measures a single feature. Although lower Cronbach Alpha values are accepted in scales with few questions, between 0.7-0.95 are generally accepted as reliable (14, 27). The Cronbach's Alpha coefficient of A-DSP was calculated as 0.860 for the first factor, 0.784 for the second factor, 0.698 for the third factor, 0.765 for the fourth factor, and 0.919 for the whole scale. A sufficient level of Cronbach's Alpha value for each factor and 21 questions that make up the whole scale shows that the questions are consistent. It can be interpreted that the internal consistency of the A-DSP is provided and it is quite reliable. The Cronbach's alpha value of the PDS, which was created to screen for depression in pregnant women using the Hamilton Depression Rating Scale, was found to be 0.81.

The reliability coefficient is affected by the sample size. In addition, the correlation coefficients between the items and the fact that the participants knew the purpose of the test before collecting data are among the factors affecting reliability (14). In our study, the sample size was sufficient and the correlation coefficients between the items were in the appropriate range. Before applying the questionnaire, the participants were informed about the purpose of the study, and the points that were curious or not understood by the participants were answered by the researcher and the participants were enlightened. For these reasons, it can be said that the reliability coefficient is calculated correctly without being affected by these factors.

Stability, which is a component of reliability, is evaluated with the test-retest method. In this method, which is based on applying the scale to the same people twice with a certain time interval and calculating the correlation coefficient between the two measurement results, the high correlation coefficient indicates that the measurement is stable (14, 26). Another coefficient calculated in the test-retest method is ICC. It is expected that the correlation coefficient and ICC value will be 0.70 and above (28). It was observed that there was a strong positive correlation between the scores obtained as a result of applying A-DSP to the same pregnant women at two-week intervals (r=0.745). In addition, the ICC value was calculated as 0.845. It was found that the scores obtained from A-DSP

did not change according to time, A-DSP was stable in repeated measurements, gave similar results, and provided test-retest reliability.

ROC analysis was performed to calculate the cut-off score of A-DSP. The fact that the area under the curve in the ROC analysis is close to 1 indicates that the test has high discrimination (29). In our study, the area under the curve was found to be 0.932. When deciding on the cut-off point, it is recommended to use the point where the sensitivity+selectivity value is the highest and the sensitivity and selectivity values are closest to each other (29). When evaluated according to these criteria, the cut-off score of A-DSP was 40.5. For this cut-off score, the sensitivity was 85.1% and the specificity was 87.6%. In other words, while the success of A-DSP to identify a pregnant woman at risk of depression is 85%, the success of identifying a pregnant woman without a risk of depression is 87%. In addition, the LR value for this cut-off score was found to be 6.8. The larger the LR value, the better distinguishing individuals who are truly at risk. It can be interpreted that A-DSP produced 6.8 true positive versus 1 false positive result (29).

Strengths and Limitations of the Study: This study is important because it's the first scale development study that includes pregnancy-specific questions designed only for the pregnancy period on depression. One of the strengths is that the study was carried out on a large sample.

The diagnostic criteria for depression in DSM and some normal symptoms during pregnancy are similar. For this reason, the diagnosis of depression can be made more than normal in the evaluations made according to the DSM criteria in pregnant women. Self-report scales developed according to DSM criteria may also indicate more cases than they actually are and produce erroneous results (30). In order to avoid such mistakes during the development of A-DSP, no questions were prepared that included somatic symptoms such as palpitations, weight gain, increased or decreased appetite, and decreased sexual desire, which overlapped with pregnancy symptoms. However, it was not possible to exclude all overlapping symptoms from the scale. Although changes in mood, symptoms of weakness and fatigue are expected symptoms of pregnancy, they are also among the most basic symptoms of depression. For this reason, items questioning these features are included in the A-DSP.

There were few cases of doctor-diagnosed depression in the study. EDS was used as the gold standard in the ROC analysis instead of clinical diagnosis. This limitation is one of the weaknesses of the study. In addition, this study was conducted on pregnant women who applied to only one medical school hospital. It would be useful to test the scale on larger groups for its generalizability.

CONCLUSION

As a result, it was seen that A-DSP is a valid and reliable scale that can be used to screen for depression in pregnant women in Turkish society. A-DSP consists of 21 items and 4 sub-dimensions. The total score that can be obtained varies between 21-84. It's accepted that the higher the score, the higher the risk of depression in pregnant. In addition, it's accepted that there is a suspicion of depression in pregnant women who score 41 and above. In line with this information, it was determined that 29.5% of the pregnant women had a suspicion of depression in this study.

It is thought that it would be beneficial to monitor the mental status of pregnant women using A-DSP, to perceive it as an early warning for depression when there is a change, to screen for depression using A-DSP, and to refer to a psychiatrist for further examination when necessary. Identifying pregnant women with suspected depression or an increased risk of depression during follow-up with the help of A-DSP will help to define the risk factors for pregnancy depression. It would be appropriate to study and test A-DSP in different parts of the society and in pregnant women with different characteristics.

REFERENCES

1. Woody C, Ferrari A, Siskind D, Whiteford H, Harris M. A systematic review and meta-regression of the prevalence and incidence of perinatal depression. Journal of Affective Disorders. 2017;219:86-92.
2. Breedlove G, Fryzelka D. Depression screening during pregnancy. Journal of Midwifery & Women's Health. 2011;56(1):18-25.
3. Altshuler LL, Cohen LS, Vitonis AF, Faraone SV, Harlow BL, Suri R, et al. The Pregnancy Depression Scale (PDS): a screening tool for depression in pregnancy. Archives of Women's Mental Health. 2008;11:277-85.
4. Campagne DM. The obstetrician and depression during pregnancy. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2004;116(2):125-30.
5. Su K-P, Chiu T-H, Huang C-L, Ho M, Lee C-C, Wu P-L, et al. Different cutoff points for different trimesters? The use of Edinburgh Postnatal Depression Scale and Beck Depression Inventory to screen for depression in pregnant Taiwanese women. General Hospital Psychiatry. 2007;29(5):436-41.
6. Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Depression during pregnancy: overview of clinical factors. Clinical Drug Investigation. 2004;24:157-79.
7. Tran TD, Tran T, La B, Lee D, Rosenthal D, Fisher J. Screening for perinatal common mental disorders in women in the north of Vietnam: a comparison of three psychometric instruments. Journal of affective disorders. 2011;133(1-2):281-93.

8. Ji S, Long Q, Newport DJ, Na H, Knight B, Zach EB, et al. Validity of depression rating scales during pregnancy and the postpartum period: impact of trimester and parity. Journal of Psychiatric Research. 2011;45(2):213-9.
9. Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Prevalence of depression during pregnancy: systematic review. Obstetrics & Gynecology. 2004;103(4):698-709.
10. Comrey AL, Lee HB. A first course in factor analysis. Psychology Press; 2013.
11. Osborne JW. Best practices in quantitative methods. Sage Publications; 2008.
12. Aydin N, Inandi T, Yigit A, Nalan Sahin Hodoglugil N. Validation of the Turkish version of the Edinburgh Postnatal Depression Scale among women within their first postpartum year. Social Psychiatry and Psychiatric Epidemiology. 2004;39:483-6.
13. Lawshe CH. A quantitative approach to content validity. Personnel Psychology. 1975;28(4):563-75.
14. Sürücü L, Maslakçı A. Validity and reliability in quantitative research. Business & Management Studies: An International Journal. 2020;8(3):2694-726.
15. Watson JC. Establishing evidence for internal structure using exploratory factor analysis. Measurement and Evaluation in Counseling and Development. 2017;50(4):232-8.
16. Williams B, Onsman A, Brown T. Exploratory factor analysis: A five-step guide for novices. Journal of Emergency Primary Health Care (JEPHC). 2010;8:1-13.
17. Hooper D, Coughlan J, Mullen MR. Structural equation modelling: guidelines for determining model fit. Electronic Journal of Business Research Methods. 2008;6(1):53-60.
18. Yonkers KA, Wisner KL, Stewart DE, Oberlander TF, Dell DL, Stotland N, et al. The management of depression during pregnancy: a report from the American Psychiatric Association and the American College of Obstetricians and Gynecologists. General Hospital Psychiatry. 2009;31(5):403-413.
19. Arkar H, Şafak C. Exploring dimensions of the Beck Depression Inventory in a clinical sample. Turkish Journal of Psychology. 2004;19(53):117-127.
20. Tatar A, Saltukoglu G. The adaptation of the CES-Depression Scale into Turkish through the use of confirmatory factor analysis and item response theory and the examination of psychometric characteristics. Bulletin of Clinical Psychopharmacology. 2010;20(3):213-227.
21. Jomeen J, Martin C. Replicability and stability of the multidimensional model of the Edinburgh Postnatal Depression Scale in late pregnancy. Journal of Psychiatric and Mental Health Nursing. 2007;14(3):319-24.
22. Brouwers EP, van Baar AL, Pop VJ. Does the Edinburgh Postnatal Depression Scale measure anxiety? Journal of Psychosomatic Research. 2001; 51(5):659-663.
23. Jomeen J, Martin CR. Confirmation of an occluded anxiety component within the Edinburgh Postnatal Depression Scale (EPDS) during early pregnancy. Journal of Reproductive and Infant Psychology. 2005;23(2):143-54.
24. Adouard F, Glangeaud-Freudenthal N, Golse B. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in a sample of women with high-risk pregnancies in France. Archives of Women's Mental Health. 2005;8:89-95.
25. Töreki A, Andó B, Keresztúri A, Sikovanyecz J, Dudas RB, Janka Z, et al. The Edinburgh Postnatal Depression Scale: translation and antepartum validation for a Hungarian sample. Midwifery. 2013;29(4):308-15.
26. Heale R, Twycross A. Validity and reliability in quantitative studies. Evidence-Based Nursing. 2015;18(3):66-67.
27. Tavakol M, Dennick R. Making sense of Cronbach's alpha. International Journal of Medical Education. 2011;2:53-55.
28. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Archives of Orofacial Science. 2017;12(1):1-11.
29. Hajian-Tilaki K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian Journal of Internal Medicine. 2013;4(2):627-635.
30. Matthey S, Ross-Hamid C. (2011). The validity of DSM symptoms for depression and anxiety disorders during pregnancy. Journal of Affective Disorders. 2011;133(3):546-552.