Mediterranean Journal of Clinical Psychology

ISSN 2282-1619

MJCP

Volume 13, n 1, 2025

Articles

Development of a short version of the Spielberger state and trait anxiety inventory (STAI) for pregnant women

*Alberto Stefana [1] *, Marie Furuta [2], Loredana Cena [3], Alice Trainini [3], Gabriella Palumbo [4], Fiorino Mirabella [4] §, Antonella Gigantesco [4] §*

Abstract

Background: The Spielberger state-trait anxiety inventory (STAI) is a widely recognized tool for assessing state and trait anxiety components in the general population. While reliable and valid, each scale in the STAI has 20 items, which limits its usability in certain real-world clinical settings. Additionally, certain items can inaccurately increase anxiety scores among expectant women. This study aimed to develop a concise yet dependable and valid version of these scales tailored for pregnant individuals.

Methods: We engaged 1,158 expectant women who completed the STAI and other assessments to determine criterion validity. Our methodology incorporated item response theory, confirmatory factor analysis, and k-fold cross-validation. Additionally, Bland-Altman regressions and plots assisted in gauging score accuracy, and receiver operating characteristic analyses determined discriminative validity. We also set clinical change benchmarks.

Results: Confirmatory factor analysis yielded the following fit indices: $X^2(5) = 32.08$, CFI = .99, TLI = .97, RMSEA = .07 (90% CI [.05, .09]), and SRMR = .02 for the state anxiety scale (omega total = .85), and $X^2(5) = 23.58$, CFI = .99, TLI = .97, RMSEA = .06 (90% CI [.04, .08]), and SRMR = .02 for the trait anxiety scale (omega total = .79). Furthermore, both scales showed meaningful correlations with measures depression.

Conclusion: The resulting five-item STAI short forms manifested robust psychometric attributes and reliability, providing a theoretically grounded and reliable anxiety assessment during pregnancy. This can be effectively applied in obstetric and gynecological settings, offering a streamlined response experience for expectant women.

[1] Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

[2] Department of Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Kyoto, Japan

[3] Department of Clinical and Experimental Sciences, Section of Neuroscience, Observatory of Perinatal Clinical Psychology, University of Brescia, Brescia, Italy

[4] Center for Behavioural Sciences and Mental Health, National Institute of Health, Rome, Italy

§ These authors equally contributed to this work and should be considered co-last authors.

E-mail corresponding author: alberto.stefana@gmail.com

Keywords:
Perinatal anxiety; Pregnancy; Psychometric properties; Scale development; Self-report.

Received: 15 January 2025
Accepted: 22 April 2025
Published: 30 April 2025

Citation: Stefana, A., Furuta, M., Cena, L., Trainini, A., Palumbo, G., Mirabella, F., & Gigantesco, A. (2025). Development of a short version of the Spielberger state and trait anxiety inventory (STAI) for pregnant women. *Mediterranean Journal of Clinical Psychology 13*(1). https://doi.org/10.13129/2282-1619/mjcp-4505

1. Introduction

Pregnancy is often perceived as a period of joy and happiness for expectant mothers (Dunkel Schetter, 2011) with some theories suggesting hormonal safeguards against the development of mental disorders (Bennett et al., 2004). However, empirical evidence contradicts these beliefs, indicating that pregnancy can be a period of vulnerability to mental disorders, particularly anxiety and depression (Caffieri & Margherita, 2021; Cena et al., 2021b ; Dennis et al., 2017; Howard & Khalifeh, 2020; Yin et al., 2021), and that these potential implications are sometimes considered only superficially by healthcare professionals and pregnant women alike (Dimitriu et al., 2020).

Anxiety disorders, often comorbid with depression (Cena et al., 2021a), have received attention for their increased prevalence during pregnancy, potentially exceeding that of the general population (Val & Míguez, 2023). A systematic review of 102 studies from 34 countries, published in 2017, found that the prevalence of self-reported anxiety symptoms and a clinical diagnosis of any anxiety disorder during pregnancy was 23% and 15%, respectively (Dennis et al., 2017). A meta-analysis of 19 studies from 10 countries examining the psychological impact of the COVID-19 pandemic on pregnant women suggested an increase in the prevalence of self-reported anxiety symptoms of up to 42%, partly due to reduced family, social and professional support during pregnancy following the pandemic (Fan et al., 2021).

A comprehensive concept analysis of pregnancy-related anxiety, based on 38 studies, identified three key attributes: affective responses (e.g., fear and panic), cognitive processes (e.g., excessive worry), and somatic symptoms (e.g., sleep problems and fatigue). These attributes collectively encompass nine dimensions of pregnancy-related anxiety: concerns about fetal health, fear of fetal loss, anxiety about childbirth, worries related to parenting and newborn care, the mother's well-being, body image issues, healthcare-related concerns, financial pressures, and family and social support (Bayrampour et al., 2016). Various psychosocial stressors, such as societal expectations, cultural norms, and socioeconomic conditions, play a crucial role in shaping this anxiety—for example, societal pressures often exacerbate concerns about infant care roles.

The impact of pregnancy-related anxiety is significant, affecting both maternal and child health (Dennis et al., 2017). Women with this anxiety may opt for elective cesarean sections due to childbirth fears (Hall et al., 2009; Rubertsson et al., 2014) adopt less effective coping strategies (George et al., 2013), and are at a higher risk for postpartum depression and suicide (Farias et al., 2013). Moreover, antenatal anxiety can adversely affect fetal and infant health, leading to more preterm births (Ibanez et al., 2012), impaired maternal-child bonding, and poor biological and behavioral development in children (Farré-Sender et al., 2018; Field, 2017; Grigoriadis et al., 2019). These associations, indicating that pregnancy-related anxiety is a distinct type of

anxiety, are not consistently observed with other forms of anxieties (Brunton et al., 2019), emphasizing the importance of its early identification and management.

Currently, several clinical guidelines and professional organizations, including the American College of Obstetricians and Gynecologists (American College of Obstetricians and Gynecologists, 2018), the American Psychiatric Association (American Psychiatric Association, 2022) and the UK National Institute for Health and Care Excellence (NICE, 2014) recommend screening for psychosocial problems for all women in the perinatal period. However, reports suggest that implementing such recommendations in practice is uncommon. The lack of time and resources are considered the main obstacles for clinicians that hinder the assessment of anxiety and depression, especially in busy obstetric settings (Coleman et al., 2008; Dennis et al., 2017). The lengthy nature of multiple questionnaires on anxiety and depression may also lead to disinterest and incomplete responses among pregnant women (Ellis et al., 2022). To enhance the feasibility of screening for perinatal mental health issues, there is a need for shorter, less demanding tools that do not require specialized training for interpretation, which may be favored over longer tools (Gigantesco et al., 2022).

The Spielberger state-trait anxiety inventory (STAI), translated into more than 70 languages and dialects, is the most well-known and cited measure of state and trait anxiety (Knowles & Olatunji, 2020). Its strong psychometric properties (Barnes et al., 2002) make it particularly valuable in distinguishing between people with and without anxiety disorders (Curtiss & Klemanski, 2015). It also has been widely used in perinatal populations (Ayers et al., 2015). However, its comprehensive nature can be cumbersome in extensive evaluations. Additionally, some STAI items, such as "I feel comfortable" and "I feel at ease," can overlap with symptoms affected by normal physiological changes during pregnancy, potentially leading to an inappropriate inflation of anxiety scores among pregnant women (Adhikari et al., 2021; Davies et al., 2021).

Previous efforts to develop a short version of the STAI have primarily targeted the general population (Adhikari et al., 2021) and faced several methodological challenges such as small sample sizes, the inclusion of reverse-scored items that diminish reliability and validity, and unclear statistical analyses (Zsido et al., 2020). These challenges were also present in a 1992 UK study by Marteau & Bekker, the only study focusing on creating a short version of the STAI based on pregnant women (Marteau & Bekker, 1992). Marteau and Bekker reported high correlation coefficients ($r = .95$) for their six-item short form with the full 20-item STAI for the state anxiety scale, suggesting that the short form maintains the psychometric properties of the full form (Marteau & Bekker, 1992). However, they did not create a short form for the trait anxiety scale. These limitations highlight the urgent need for a tailored, validated, and reliable

short version of the STAI that addresses both state and trait anxiety specifically for pregnant women (Hadfield et al., 2022).

In this study, we aim to create a brief version of the STAI state and trait anxiety scales specifically for pregnant women, preserving its psychometric strength. Using the item response theory, we focus on the items most indicative of varied anxiety levels to provide standard scores and cutoff points, benefiting future research and clinical practice on antepartum anxiety.

2. Materials and Methods

2.1 Study design and participants

This report represents secondary analysis of a study that evaluated the prevalence of antepartum and postpartum depression and anxiety among women attending one of eleven publicly funded primary or obstetrics-gynecology centers in Italy. The original study protocol provided a comprehensive account of the rationale and methodology employed (Cena et al., 2020). Inclusion criteria for participants were proficiency in spoken and written Italian, and either pregnancy or having a biological infant aged 6 months or younger. Exclusion criteria encompassed psychotic symptoms or substance use. This report exclusively analyzes data from the antepartum sample (N = 1158). Table 1 displays sociodemographic information of the participants.

Table 1. Sociodemographics and reproductive characteristics of the sample

Characteristics	% (*n*)
Age	
18–29	24% (274)
30–35	47% (543)
>35	29% (339)
Nationality	
Italian	93% (1076)
Other	7% (82)
Marital status	
Married or cohabiting	92% (1053)
Single, separated, divorced or widowed	8% (96)
Educational level	
University	51% (590)
Secondary	36% (415)
Primary or Illiterate	13% (144)
Working status	
Permanent employee	72% (825)
Temporary employee	10% (115)
Student, homemaker or unemployed	18% (204)
Economic Status	
Average high status	6% (75)
A few problems without specific difficulties	48% (544)
Same or many problems	46% (525)
Previous pregnancies	

Yes	75% (868)
No	25% (290)
Living children	
Yes	17% (195)
No	83% (963)
EPDS, *Mean (SD)*	4.91 (3.92)
PHQ-9, *Mean (SD)*	4.44 (3.10)

Note: EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire-9.

The research received approval from the Ethics Committee of the Healthcare Centre at Bologna Hospital (Reg. no. 0077805; June 27, 2017). All participants were given detailed oral and written information about the study's content and implications and provided informed consent by signing the consent form. Recruitment began on January 7, 2017, and ended on June 27, 2018, while data collection concluded on November 30, 2019.

2.3 Procedures

Women who agreed to participate underwent a preliminary clinical interview conducted by a peripartum clinical assessment-trained psychologist. The interview aimed to gather information on maternal psychiatric history, including symptoms of anxiety, depression, psychosis, self-harm tendencies, suicidal ideation, and substance abuse. Participants with significant symptoms were referred for further psychiatric assessment and were excluded from the study. Interviews took place individually in private rooms. Subsequently, participants completed self-report questionnaires on anxiety and depression.

2.4 Measures

The Italian version of the *State-Trait Anxiety Inventory* (STAI Form Y-1) (Spielberger, 1983) is a 40-item self-report measure of anxiety, with responses evaluated on a four-point Likert scale ranging from 'Not at all' to 'Very much so.' It incorporates two 20-item scales: the State Anxiety Scale (SAS) and the Trait Anxiety Scale (TAS), each measuring different dimensions of anxiety. The SAS gauges situational anxiety, capturing feelings experienced in the present moment. On the other hand, the TAS evaluates the propensity for anxiety as a stable trait over time. Ten items on the SAS and seven on the TAS are reverse-scored. Participants completed the validated Italian version of STAI Form Y-1 (Spielberger et al., 1989). The inventory exhibited good internal consistency with Cronbach's alpha ranging from .86 to .95 (Spielberger, 1983, 1989; Spielberger et al., 1989).

The Italian version of the *Edinburgh Postnatal Depression Scale* (EPDS) (Benvenuti et al., 1999; Cox et al., 1987) is a 10-item self-report screening measure of perinatal depressive symptoms. Each item is rated on a four-point Likert scale, allowing mothers to indicate how frequently they

experienced the corresponding symptom over the previous seven days. In our study, the EPDS demonstrated a reliability coefficient of Cronbach's alpha = .80 and average inter-item r = .28.

The Italian version of the *Patient Health Questionnaire-9* (PHQ-9) (Kroenke et al., 2001; Picardi et al., 2004) is 9-item self-report screening measure of depression severity based on DSM-5 (American Psychiatric Association, 2013) diagnostic criteria. Each item on the scale is rated using a four-point Likert scale, through which mothers can express the frequency of their experience with the corresponding symptom over the preceding 14-day period. The internal consistency of this measure in our sample was Cronbach's alpha = .74 and average inter-item r = .24. It should be noted that a comparison of the factor structures and reliabilities of the PHQ-9 compared to the EPDS showed that the second should be preferred for screening antepartum. It should be noted that comparisons of factor structures and reliability coefficients between the PHQ-9 and the EPDS indicate that the EPDS is preferable for antepartum screening (Stefana et al., 2023, 2024).

2.5 Statistical analyses

The development and validation of the short-form scales followed the steps recommended in literature (Stefana et al., 2025a, 2025b; Youngstrom et al., 2020). In the initial stage of item selection, reverse-scored items were eliminated from both the SAS (items 1, 2, 5, 8, 10, 11, 15, 16, 19, 20) and the TAS (items 21, 23, 26, 27, 30, 33, 34, 36, 39). This decision was grounded on evidence suggesting that these items pose greater difficulty in responding and may not assess the identical construct as the straight-scored items (Zhang & Savalei, 2016; Zsido et al., 2020), thus lowering the internal consistency, reliability, and validity of the scale (Lindwall et al., 2012). Principal component analysis (PCA) of the standardized residuals from an item response theory (IRT) model was performed to examine the unidimensionality (Smith, 2002) in both the SAS and the TAS. Unidimensionality was assessed using *t*-tests that compared pairs of ability estimates from separate IRT calibrations of the two sets of items, either loading positively or negatively on the first component of PCA. To achieve strict unidimensionality, the proportion of significant *t*-tests must be less than 5%. The lower bound of the binomial confidence interval for proportions <5% would then be acceptable. The IRT graded response model (GRM) (Samejima, 2010) was separately calculated for the SAS and TAS to determine the parameters of item discrimination (a) and difficulty (β). It was also used to quantify reliability, information (a measure of reliability that examines the precision of score estimates across the underlying trait range), and item characteristics. Items that provided high information ($a > 1.35$) (Baker & Kim, 2017) across a broad range of theta (θ) levels were selected. Confirmatory factor analysis (CFA) with maximum likelihood estimation was utilized to evaluate the fit of the factor solution for the refined short form of each scale. *K*-fold cross-validation was employed to verify the

robustness of the final models. The formulas from Smith et al. (Smith et al., 2000) were implemented to predict potential McDonald's omega (ω) total and content coverage. Average item correlation was used as an estimate of internal consistency not dependent on scale length (Streiner et al., 2015), while McDonald's omega total was employed as a secondary estimate. In addition, marginal reliability across a range of theta levels was estimated using IRT (Feuerstahler et al., 2020). Bland-Altman plots (Bland & Altman, 1986) were utilized to examine the accuracy of scores based on the short forms as compared to the full-length 20-item versions. These plots provide estimates of score bias and "limits of agreement." Receiver operating characteristic (ROC) analyses (DeLong et al., 1988) were conducted to evaluate the discriminative validity of both short-form scales. Lastly, criterion validity was assessed by examining the correlation between, on the one hand, the short and full-length SAI and TAI scores, and on the other hand, sociodemographics, EPDS, and PHQ-9 scores. Cohen's q test (Cohen, 1992) was used to compare differences in correlation coefficients.

Analyses were performed using the R packages *caret* v6.0-94 (twice *k*-fold cv), *kfa* v0.2.0 (*k*-fold cv), *lavaan* v0.6-11 (CFAs), *mirt* v1.36.1 (IRT analyses), *pROC* v1.18.0 (ROC analyses), *psych* v2.2.9 (scoring, classical test theory reliability estimates), and *ggplot2* v3.4.2 and *semPlot* v1.1.6 (additional visualizations).

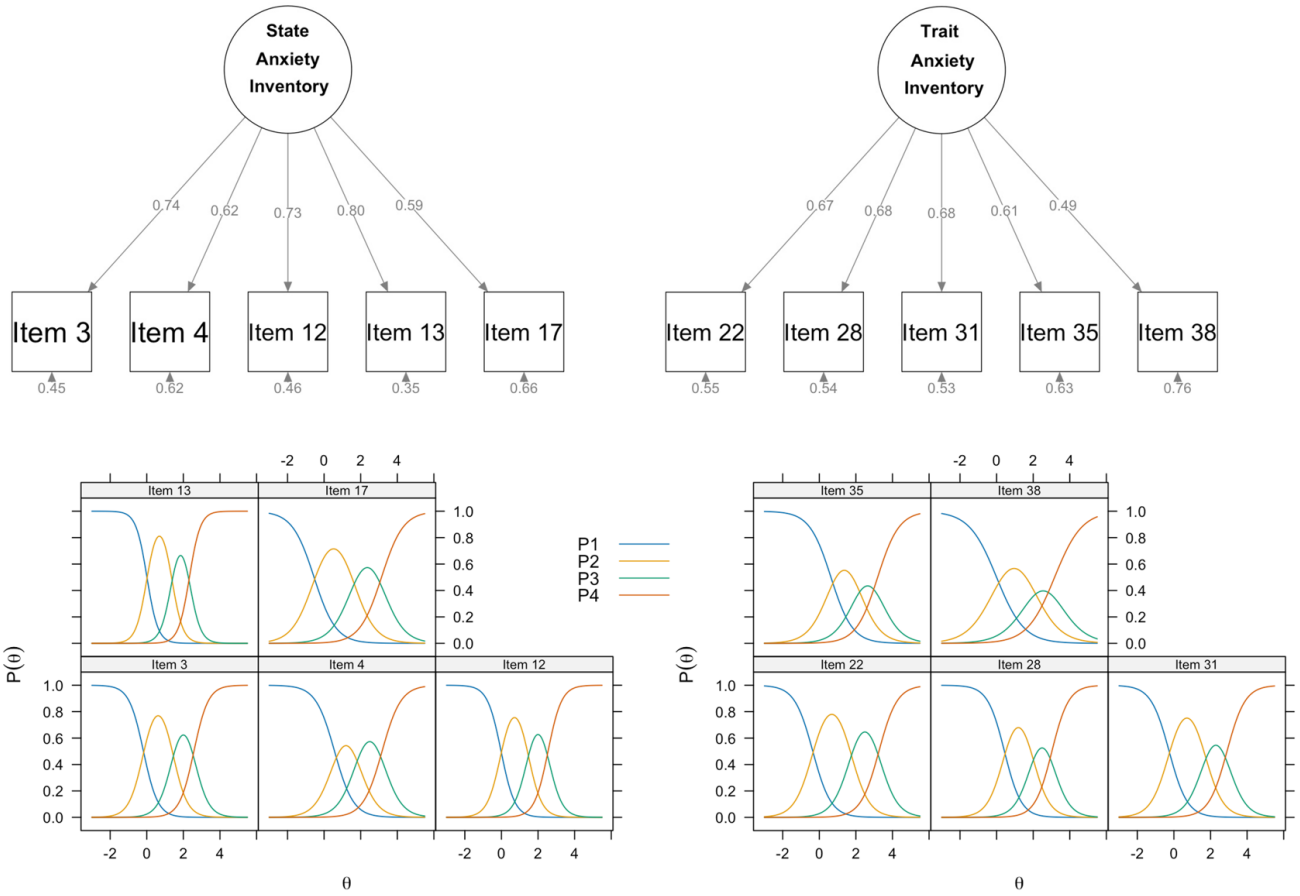
3. Results

3.1. Unidimensionality of the scales

PCA of standardized residuals from the IRT model was conducted separately on both the SAS and the TAS, following the exclusion of reverse-scored items. In both instances, the analyses showed that the first principal component captured sufficient variance, and paired t-tests of ability estimates from items loading positively and negatively on this component showed no significant differences. These results confirm the unidimensionality of both scales (see Supplementary Online Material 1 for detailed statistics).

3.2 Item response modeling

Graded response model (GRM) item response theory (IRT) was used to analyze the 10-item SAS and the 11-item TAS. Table 2 provides discrimination (a) and difficulty (β) parameters values. Six items for each scale showed high to very high discrimination ability. The five items with highest information across a wide range of theta (θ) levels were selected for the abbreviated trait and state scales. Figure 2a shows the item option characteristic curves for the short version of both scales.



Note. P1, P2, P3, and P4 denote the probabilities of picking each of the four multiple-choice answers, expressed as functions of the latent trait (θ). These probability curves, which collectively total to 1, show the growing likelihood of selecting higher response options as θ increases. This highlights how each option is associated with different levels of the underlying ability or trait.

Figure 2. Item option characteristic curves and measurement model visualization of the short forms of State and Trait Anxiety Scales

Table 2. Discrimination (α) and difficulty (β) parameters for the State and Trait Anxiety Scales based on IRT models

	n.	Item content	α	β1	β2	β3
State Anxiety Scale	3	I am tense	2.51	-0.19	1.42	2.58
	4	I feel strained	2.03	0.56	1.79	3.09
	6	I feel upset	2.45	0.73	1.98	3.04
	7	I am presently worrying over possible misfortunes	1.31	0.32	2.11	3.40
	9	I feel frightened	1.23	0.38	2.47	4.30
	12	I feel nervous	2.24	-0.03	1.53	2.71
	13	I am jittery	3.27	-0.00	1.37	2.35
	14	I feel indecisive	1.54	0.45	1.96	3.09
	17	I am worried	1.80	-0.53	1.52	3.00
	18	I feel confused	1.60	0.90	2.46	3.64

Trait Anxiety Scale	22	I feel nervous and restless	2.09	-0.35	1.71	3.17
	24	I wish I could be as happy as others seem to be	1.03	0.65	2.19	3.21
	25	I feel like a failure	1.95	1.44	2.49	3.15
	28	I feel that difficulties are piling up so that I cannot overcome them	1.92	0.45	2.07	3.17
	29	I worry too much over something that really doesn't matter	1.40	-0.57	1.61	2.88
	31	I have disturbing thoughts	1.99	-0.27	1.69	2.89
	32	I lack self-confidence	1.42	0.05	1.94	3.35
	35	I feel inadequate	1.83	0.62	2.03	3.04
	37	Some unimportant thought runs through my mind and bothers me	1.74	-0.12	1.99	3.12
	38	I take disappointments so keenly that I can't put them out of my mind	1.56	-0.00	1.75	2.90
	40	I get in a state of tension or turmoil as I think over my recent concerns and interests	1.91	-0.61	1.24	2.24

Note: The reversed scored items have been removed.

3.3 Confirmatory factor analysis

CFA was run for each scale separately to test the one-factor model. The 5-item SAS showed an excellent fit for the data: $X^2(5) = 32.08$, CFI = .99, TLI = .97, RMSEA = .07 (90% CI [.05, .09]), and SRMR = .02. Similarly, the 5-item TAS showed an $X^2(5) = 23.58$, CFI = .99, TLI = .97, RMSEA = .06 (90% CI [.04, .08]), and SRMR = .02. Furthermore, *k*-fold cross-validation setting 4 subsets was performed to test the robustness of the unidimensional models and provided the following fit indices: $X^2(5) = 15.60$, CFI = .98, TLI = .96, RMSEA = .08 (90% CI [.06, .13]), and SRMR = .03 for the SAI, and $X^2(5) = 10.50$, CFI = .98, TLI = .97, RMSEA = .05 (90% CI [.00, .09]), and SRMR = .03 for the TAS. Figure 2c shows a measurement model visualization of the short forms of both TAS and SAS.

3.4 Reliability and Precision

Projected internal consistency (Smith et al., 2000) McDonald's ω total α was .74 for the 5-item SAS and .72 for the 5-item TAS. The observed internal consistency results were McDonald's ω total = .83 and average inter-item r = .49 for the SAS, and McDonald's ω total = .76 and average inter-item r = .39 for the TAS. IRT analyses showed that the SAS had reliability >.80 from θ of -0.6 to +3.2, whereas the TAS had reliability >.80 from θ of +0.2 to +3.4 (see Figure 2b). Table 3 reports key statistical parameters used to evaluate the reliability and sensitivity of our short forms.

Table 3. Descriptive statistics and precision of the short and full lengths State and Trait Anxiety Scales

	Short forms		Full length forms	
	SAS	TAS	SAS	TAS
Descriptive statistics				
Potential Range	5 to 20	5 to 20	20 to 80	20 to 80
Observed Range	5 to 20	5 to 20	20 to 77	20 to 71
Mean, *SD*	8.22 (2.73)	7.90 (2.47)	34.70 (8.96)	34.98 (9.04)
Skew	.85	1.08	.88	.88
Kurtosis	.69	1.13	.86	.62
Standard Error of Measurement (SE_m)	1.13	1.21	2.53	2.71
Standard Error of Difference (SE_d)	1.59	1.71	3.58	3.84
Internal consistency reliability				
Average inter-item *r*	.49	.39	.36	.34
Cronbach's alpha	.83	.76	.92	.91
McDonald's omega total	.83	.76	.92	.91
Clinical change benchmarks				
90% Critical Change	1.85	1.99	4.17	4.46
95% Critical Change	2.21	2.37	4.97	5.32
Minimally Important Difference (MID)	1.36	1.24	4.48	4.52
Minimum Change for a Reliable Change	3.12	3.35	7.02	7.52

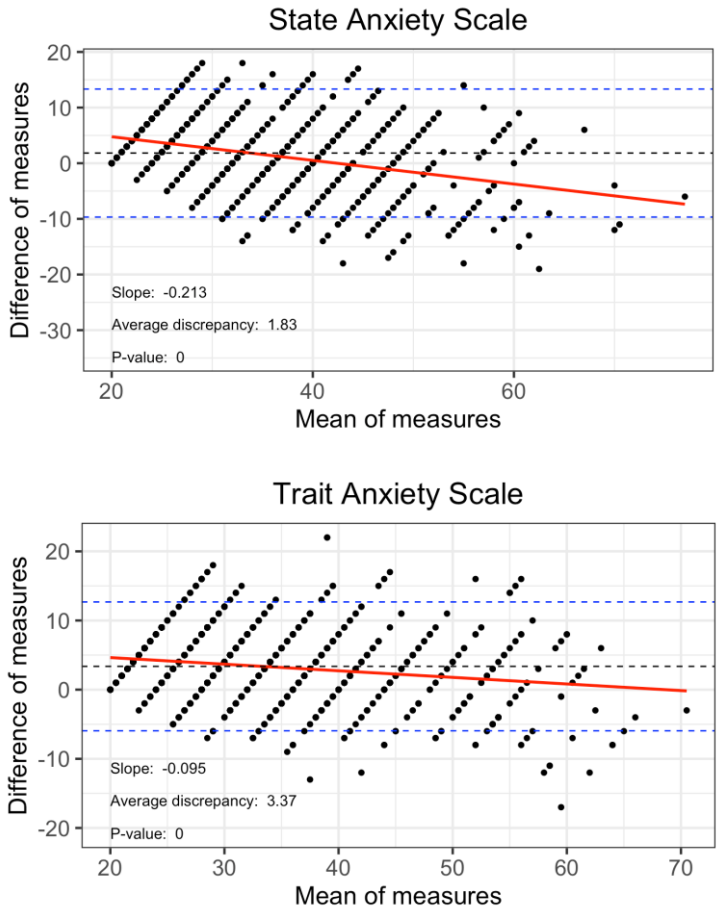
Note: SAS = State Anxiety Scale; TAS = Trait Anxiety Scale.

More specifically, the standard error of measurement is a representation of the average amount that observed scores deviate from true scores, providing an indication of the precision of the individual scores on our test (Streiner et al., 2015). The standard error of difference provides an estimate of the standard deviation of differences between scores for individuals who take the test twice under identical conditions (Streiner et al., 2015). This allows for an understanding of the precision of difference scores and the likely range within which an individual's true score difference lies. The 90% and 95% Critical Change values represent the smallest change in scores that can be considered statistically significant with 90% and 95% confidence respectively, thus helping to ascertain whether an observed change in a person's score over time is due to measurement error or a real change (Jacobson & Truax, 1991; Kazis et al., 1989). The minimally important difference is an estimate of the smallest change in score that would be perceived as beneficial by the individual or would lead to a change in the person's management (Jaeschke et al., 1989; Revicki et al., 2008). Finally, the minimum change for a reliable change is the smallest change in a score that can be considered reliable, not due to measurement error, and therefore reflects a true change in the individual's status (Jacobson & Truax, 1991; Wise, 2004).

3.5 Retained Content Coverage, Agreement, and Bias

Projected correlations (Smith et al., 2000) between the short form (k = 5) and full-length version (k = 20) were r = .68 for the SAS and r = .65 for the TAS. The observed correlations were r = .84 for the SAS and r = .88 for the TAS.

The Bland-Altman plots and regressions were used to measure the agreement between the short and long forms of the scales (Figure 1).



Note. Red lines are the regression of the difference in scores on the average of the scores, and blue dashed lines represent the "limits of agreement" or 95% CI around the differences.

Figure 1. Bland-Altman Plots comparing accuracy of short and full-length version of State and Trait Anxiety Scales.

For a valid comparison, the short forms were pro-rated to have the same scale as the full-length forms. Upon scaling, the SAS total score showed an average discrepancy of 1.83 points higher on one scale, with a statistically significant and clinically not relevant slope of -0.21. A negative slope indicates that the difference between the scores from the two versions of the scale tend to decrease as the average anxiety score increases (the discrepancy between the two forms is less for individuals with higher anxiety levels). Within the range where most participant scores fell, discrepancies were relatively minor.

These results indicate a good degree of agreement with minimal evidence of bias between the short and full-length versions of the SAS. Regarding the TAS, the total score exhibited an average discrepancy of 3.37 points higher using the full item set, with a statistically significant but clinical negligible slope of -0.10. Within the scope where most participants' scores were located, the discrepancy was near zero, implying that the scores from both versions were largely similar. Nevertheless, there was a tendency for the full item set to yield higher scores at the lower end of the possible range. The results signal a strong agreement with minimal bias between the short and full-length versions of the TAS.

3.6 ROC Analysis

We utilized pre-defined threshold scores for high state anxiety (>40) and high trait anxiety (>40) in perinatal samples (Dennis et al., 2013; Grant et al., 2008) to perform receiver operating characteristic (ROC) analysis.

The 5-item SAS exhibited the highest Youden index (.683), indicating its optimal performance, when utilizing a cut-off score above 9.5. This threshold yielded a sensitivity of 0.809 and a specificity of .874. Thus, individuals scoring equal to or higher than 10 on the SAS can be identified as potentially clinically anxious. With an AUC of .914 (95% CI: .896–.933) and an accuracy of .858, the SAS demonstrated remarkable discriminatory power.

Similarly, the 5-item TAS demonstrated its highest Youden index (.753) at a cut-off score exceeding 8.5. This threshold resulted in a sensitivity of .906 and a specificity of .846, indicating that individuals scoring equal to or higher than 9 on the TAI should be considered potentially clinically anxious. The TAS exhibited an AUC of .947 (95% CI: .935–.959) and an accuracy of .862, highlighting its excellent discriminatory power. Hence, both the SAS and the TAS can be considered as clinically useful.

Table 4 reports performance metrics at various cutoff scores for high state trait anxiety.

Table 4. Performance metrics of the short forms of State and Trait Anxiety Scales at various threshold points

	Threshold	Cutoff	Sensitivity	Specificity	LR+	LR-	Yi	TP	FN	FP	TN	Accuracy	AUC
State Anxiety Scale	38	8.5	.878	.794	4.26	.153	.672	339	47	159	613	.822	.916
	39	9.5	.767	.901	7.77	.259	.668	266	81	80	731	.861	.917
	40	9.5	.809	.874	6.42	.219	.683	237	56	109	756	.858	.914
	41	9.5	.814	.861	5.85	.216	.675	223	51	123	761	.85	.909
	42	9.5	.827	.846	5.37	.204	.673	206	43	140	769	.842	.907

	43	9.5	.844	.832	5.02	.188	.676	189	35	157	777	.834	.909
	44	9.5	.859	.818	4.71	.172	.677	171	28	175	784	.825	.913
	45	9.5	.875	.799	4.35	.156	.674	147	21	199	791	.81	.915
	46	10.5	.781	.904	8.11	.242	.685	118	33	97	910	.888	.921
	47	10.5	.803	.894	7.56	.220	.697	106	26	109	917	.883	.926
	48	10.5	.809	.880	6.73	.217	.689	89	21	126	922	.873	.929
	49	10.5	.874	.870	6.73	.145	.744	76	11	139	932	.87	.94
	50	10.5	.873	.865	6.45	.146	.738	69	10	146	933	.865	.94
	51	10.5	.859	.858	6.06	.164	.717	61	10	154	933	.858	.938
	52	11.5	.831	.918	10.14	.185	.749	49	10	90	1009	.914	.945
	53	11.5	.882	.915	10.39	.129	.797	45	6	94	1013	.914	.953
Trait Anxiety Scale	38	8.5	.844	.893	7.92	.174	.738	320	59	83	696	.877	.939
	39	8.5	.879	.864	6.44	.140	.742	290	40	113	715	.868	.941
	40	8.5	.906	.846	5.90	.111	.753	271	28	132	727	.862	.947
	41	8.5	.923	.818	5.09	.094	.742	240	20	163	735	.842	.948
	42	8.5	.932	.800	4.65	.086	.731	218	16	185	739	.826	.95
	43	9.5	.810	.925	10.81	.206	.735	170	40	71	877	.904	.952
	44	9.5	.884	.917	10.66	.127	.801	160	21	81	896	.912	.965
	45	9.5	.892	.907	9.61	.119	.799	149	18	92	899	.905	.964
	46	9.5	.905	.894	8.55	.106	.799	134	14	107	903	.896	.965
	47	10.5	.857	.960	21.43	.149	.817	114	19	41	984	.948	.966
	48	10.5	.916	.956	20.69	.088	.872	109	10	46	993	.952	.974
	49	10.5	.927	.949	18.00	.077	.875	101	8	54	995	.946	.973
	50	10.5	.929	.940	15.38	.076	.868	91	7	64	996	.939	.971
	51	10.5	.943	.933	14.02	.061	.876	83	5	72	998	.934	.98
	52	10.5	.973	.923	12.56	.030	.895	71	2	84	1001	.926	.979
	53	10.5	.982	.910	10.93	.019	.893	56	1	99	1002	.914	.978

Note: FP = false positives, FN = false negatives, LR+ = likelihood ratio positive, LR- = likelihood ratio negative, TP = true positives, TN = true negatives, Yi = Youdens index.

3.7 Criterion Validity

Table 4 presents all computed correlation coefficients. Both the SAS and TAS showed very weak correlations with all sociodemographic variables, and moderate correlations with both the EPDS and the PHQ-9. When comparing the correlation coefficients between the SAS short-form and full-form with the depression scores, statistically significant Cohen's *q* values were observed for both EPDS (r = .12, p < .001) and PHQ (r = .13, p < .001). A similar pattern emerged for the TAS short-form in comparison to its full-form, demonstrating a significant difference in correlation with EPDS (r = .13, p < .001) and PHQ (r = .09, p < .05). However, despite the statistically significant *p-values*, the difference between the correlations of the short-form and full-form measures with the depression scales were found to be very small, suggesting that the short versions of these inventories can capture nearly the same information as their full-length counterparts.

Table 5. Criterion validity correlations

	Short forms		Full length forms		Cohen's *q*	
	SAS	TAS	SAS	TAS	SASs	TASs
Age	-.04	-.02	-.05	-.02	.01	.00
Nationality	-.04	-.03	-.05	-.01	.01	-.02
Marital status	-.00	-.02	-.04	-.06	.04	.04
Education	-.13***	-.01	-.15***	-.08**	.02	.07
Working status	-.09**	-.07*	-.12***	-.09**	.03	.02
Economic status	-.05	-.07*	-.08**	-.11***	.03	.04
Previous pregnancies	.05	-.00	.05	.04	.00	-.04
Living children	.08*	.00	.08*	.04	.00	-.04
EPDS	.46***	.57***	.55***	.65***	.12***	.13***
PHQ-9	.37***	.40***	.48***	.47***	.13***	.09*
5-item SAS	--					
5-item TAS	.47***	--				
20-item SAS	.84***	.61***	--			
20-item TAS	.49***	.88***	.69***	--		

Note: EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire-9.

Coefficients are point-biserial correlations for dichotomized variables, point-biserial correlations for dummy-coded categorical variables, Spearman correlations for ordinal variables, and Pearson correlations for continuous variables.

*p<.05, **p<.01, ***p<.001

4. Discussion

4.1 Summary of main findings

Previous research has consistently validated the reliability and utility of the Spielberger State-Trait Anxiety Inventory (STAI) for both research and clinical contexts. However, a version of the State Anxiety Scale (SAS) and the Trait Anxiety Scale (TAS) tailored specifically for pregnant individuals and compact enough for routine use in clinical settings was needed. Through this study, we aimed to devise a condensed form of SAS and TAS, using item response theory to pinpoint items with the best difficulty and discrimination parameters.

The final 5-item SAS and 5-item TAS exhibited strong internal consistency and very high correlation ($r > .80$) with the respective full-length scales. The absence of statistically significant discrepancies in the correlations of the criteria emphasizes the viability of these shortened versions as alternatives to their longer counterparts. The correlation coefficients between the scores of the short and full versions of both SAS and TAS exceeded mathematical expectations, reinforcing the validity of the new forms. Bland-Altman plots and regression analyses further confirmed a high concordance between the versions, revealing minimal bias, particularly for the TAS. Although the full-length scales tended to produce slightly higher scores at the lowest end of the range, this disparity seems clinically insignificant. Furthermore, receiver operating characteristic analyses affirmed the excellent discriminatory prowess of the shortened SAS and TAS.

It is also important to note that the new abbreviated scales mirrored the original versions in their correlation patterns with sociodemographic and clinical variables. As expected, the SAS and the TAS demonstrated a moderate correlation with both the EPDS and the PHQ-9. The greater correlations with the EPDS, are theoretically consistent with the fact that, unlike the PHQ-9, the EPDS includes items addressing anxiety. Taken together, the very weak correlations between each of the STAI scales and the sociodemographic variables of the pregnant women, in addition to moderate correlations with two validated measures of depressive symptomatology (there is a high comorbidity of anxiety and depressive disorders during the perinatal period), signified sound criterion validity.

To strengthen the utility of the scales in clinical practice, we calculated clinical change benchmarks for both the short and full versions. Regarding the benchmarks for the abbreviated forms, to be 90% confident of genuine change, the scores need to shift by 1.85 on the SAS and 1.99 on the TAS. At a 95% confidence level, these changes rise slightly to 2.21 and 2.37 for the SAS and TAS, respectively. The minimally important difference, representing the smallest change perceived as beneficial by the patients, was 1.36 for the SAS and 1.24 for the TAS. These values underscore the sensitivity of the tools in detecting clinically significant changes.

Furthermore, to establish a reliable change, shifts of 3.12 for the SAS and 3.35 for the TAS are required. Such benchmarks help distinguish genuine symptomatic changes from mere test-retest variations. In essence, these benchmarks provide tangible criteria to assess the effectiveness of interventions, facilitating the distinction between true shifts in anxiety states and traits *versus* measurement variability.

4.2 Comparison with existing studies

In 1992, Marteau and Bekker (1992) introduced a six-item short form of the SAS specifically designed to assess state anxiety in a sample of 200 pregnant women. This instrument, referred to as the STAI-6, has since been used extensively in antenatal and postnatal research, influencing a wide range of studies over the decades (Adhikari et al., 2021; Bond et al., 2015; Davies et al., 2021; Fraser et al., 2023; Goldman et al., 2019; Grech et al., 2024; Hewison et al., 2007). However, the STAI-6 focuses solely on state anxiety and integrates both anxiety-present ('tense', 'upset', and 'worried') and anxiety-absent ('calm', 'relaxed', and 'content') items.

Our research methodology differs significantly from that of Marteau and Bekker (1992). We have developed five-item versions of the state and trait anxiety scales based on a large sample of pregnant women to address recognized shortcomings in previous short-form anxiety scale developments (Zsido et al., 2020). We excluded reverse-scored (anxiety-absent) items, which have previously been shown to reduce the reliability and validity of scores because of participant confusion (Zhang & Savalei, 2016). Additionally, unlike the STAI-6, which relies on Cronbach's alpha, assuming tau equivalence between the items, we used McDonald's omega total to improve reliability assessments. This adjustment better captures the different contributions of each item within the complex structure of the STAI, thereby improving measurement precision.

Regarding construct validity, we observed a correlation coefficient of $r = .84$ for the five-item SAS. This strong correlation, though slightly lower than the $r = .95$ reported by Marteau and Bekker (1992) for their six-item short form with the full 20-item SAS, supports the validity of our tool. The lower correlation in our study may reflect the challenges of capturing the full range of anxiety symptoms with fewer items or could be due to differences in sample characteristics and methodological approaches. For the five-item TAS, we observed a correlation of $r = .88$, again supporting validity. However, comparison with results from Marteau and Bekker (1992) is not possible, because they did not develop a short form for trait anxiety.

4.3 Implications for clinical practices and future research

Early identification of pregnancy-related anxiety is crucial to ensure that women receive the necessary support. However, existing studies indicate that only about half of perinatal mental health cases are identified. The primary obstacles to effective screening include clinicians' time

constraints and maternal factors such as stigma and misconceptions regarding 'normal' anxiety levels during pregnancy (Royal College of General Practitioners, 2023).

Given the scarcity of validated tools specifically designed to assess pregnancy-related anxiety, current clinical guidelines recommend adapting items from general anxiety disorder scales for screening purposes (Bayrampour et al., 2016; NICE, 2014). The five-item versions of the SAS and TAS we developed aim to simplify the assessment process. This streamlined approach reduces the time burden on healthcare providers and is expected to improve response accuracy and reduce errors (Marteau & Bekker, 1992). Consequently, these improvements will likely enhance the robustness of screening results, thereby improving early identification and intervention efforts in perinatal mental health.

Although our findings provide promising preliminary evidence for the validity and reliability of the short-form screening tool for pregnancy-related anxiety in our sample, it is important to interpret these results with caution. The tool's validation in different clinical settings and among various cultural and socioeconomic groups requires further investigation. Additionally, integrating this tool into routine antenatal care should be approached stepwise, ensuring ongoing evaluation and adaptation based on real-world use.

4.4 Strengths and limitations

Our study significantly advances previous methodologies by providing more reliable and valid tools tailored for pregnant women. We utilized McDonald's omega to determine reliability since it offers a better estimate of reliability than Cronbach's alpha (Revelle & Condon, 2019). Our sample size of over 1000 participants is considered optimal for conducting factor analyses (Comrey & Lee, 2013) and strengthens the validity of our findings.

Despite these strengths, our study is not without limitations. A primary concern is the potential for cultural variation in how anxiety disorders present, which might limit the applicability of our results beyond the Italian context to other cultural settings. Lastly, the psychometrics of the five items extracted for each scale should be confirmed in a sample where these items are not embedded in the original 20-item sets, ensuring that the performances are similar without context effects.

5. Conclusion

This study offers two short-form instruments that are both theoretically meaningful and statistically sound, with preliminary evidence supporting their reliability and construct validity in assessing antenatal-specific anxiety. These short forms of the SAS and TAS can be integrated as part of an evidence-based psychological assessment battery in a prenatal care setting with a minimal response burden for pregnant women. The ability to assess maternal antenatal anxiety

quickly and accurately will allow clinicians to effectively distinguish between anxious and non-anxious pregnant individuals and encourage researchers to promote deeper insights into the development and maintenance of anxiety during pregnancy.

Ethical approval

The research received approval from the Ethics Committee of the Healthcare Centre at Bologna Hospital (Reg. no. 0077805; June 27, 2017).

Informed Consent Statement

All participants were given detailed oral and written information about the study's content and implications and provided informed consent by signing the consent form.

Data Availability Statement

Requests to access the datasets should be directed to the corresponding author.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any potential conflict of interest.

Authors' Contribution

LC and GP contributed equally to the general study design. LC and AT from the Observatory of Perinatal Clinical Psychology coordinate and manage the implementation of the study in each healthcare centre. AS and FM designed the plan of statistical analysis of the study. AG served primarily as supervisor. AS and MF participated in the writing of the manuscript. All authors have critically reviewed and agreed this final version of the article.

References

1. Adhikari, K., Patten, S. B., Williamson, T., Patel, A. B., Premji, S., Tough, S., Letourneau, N., Giesbrecht, G., & Metcalfe, A. (2021). Assessment of anxiety during pregnancy: are existing multiple anxiety scales suitable and comparable in measuring anxiety during pregnancy?. *Journal of Psychosomatic Obstetrics & Gynecology, 42*(2), 140-146. https://doi.org/10.1080/0167482x.2020.1725462
2. American College of Obstetricians and Gynecologists. (2018). ACOG Committee Opinion No. 757: screening for perinatal depression. *Obstetrics & Gynecology, 132*(5), e208-e212. https://doi.org/10.1097/aog.0000000000002927
3. American Psychiatric Association. (2013). *The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5).* American Psychiatric Association Publishing.
4. American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders: DSM-5-TR. Fifth edition, text revision* (5th ed.) American Psychiatric Association Publishing.
5. Ayers, S., Coates, R., & Matthey, S. (2015). Identifying perinatal anxiety. In J. Milgrom & A. W. Gemmill (Eds.), *Identifying Perinatal Depression and Anxiety: Evidence: Based Practice in Screening, Psychosocial Assessment, and Management* (pp. 93-107). Wiley.
6. Baker, F. B., & Kim S-H. (2017). *The Basics of Item Response Theory Using R* (1st ed.) Springer International Publishing.
7. Barnes, L. L., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger state-trait anxiety inventory. *Educational and Psychological Measurement, 62*(4), 603-618. https://doi.org/10.1177/0013164402062004005
8. Bayrampour, H., Ali, E., McNeil, D. A., Benzies, K., MacQueen, G., & Tough, S. (2016). Pregnancy-related anxiety: A concept analysis. *International Journal of Nursing Studies, 55*, 115-130. https://doi.org/10.1016/j.ijnurstu.2015.10.023
9. Bennett, H. A., Einarson, A., Taddio, A., Koren, G., & Einarson, T. R. (2004). Prevalence of depression during pregnancy: systematic review. *Obstetrics & Gynecology, 103*(4), 698-709. https://doi.org/10.1097/01.aog.0000116689.75396.5f
10. Benvenuti, P., Ferrara, M., Niccolai, C., Valoriani, V., & Cox, J. L. (1999). The Edinburgh postnatal depression scale: validation for an Italian sample. *Journal of Affective Disorders, 53*(2), 137-141. https://doi.org/10.1016/s0165-0327(98)00102-5
11. Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 327*(8476), 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8
12. Bond, D., Raynes-Greenow, C., & Gordon, A. (2015). Anxiety and acceptability related to participation in stillbirth research. *Midwifery, 31*(11), 1067-1072. https://doi.org/10.1016/j.midw.2015.07.005
13. Brunton, R., Dryer, R., Saliba, A., & Kohlhoff, J. (2019). Re-examining pregnancy-related anxiety: A replication study. *Women and Birth, 32*(1), e131-e137. https://doi.org/10.1016/j.wombi.2018.04.013
14. Caffieri, A., & Margherita, G. (2021). The psychological impact of COVID-19 on women's wellbeing during pregnancy and postpartum one year after pandemic outbreak in Italy. A Systematic review. *Mediterranean Journal of Clinical Psychology, 9*(2). http://dx.doi.org/10.13129/2282-1619/mjcp-3026

15. Cena, L., Gigantesco, A., Mirabella, F., Palumbo, G., Camoni, L., Trainini, A., & Stefana, A. (2021a). Prevalence of comorbid anxiety and depressive symptomatology in the third trimester of pregnancy: Analysing its association with sociodemographic, obstetric, and mental health features. *Journal of Affective Disorders, 295*, 1398-1406. https://doi.org/10.1016/j.jad.2021.09.015
16. Cena, L., Mirabella, F., Palumbo, G., Gigantesco, A., Trainini, A., & Stefana, A. (2021b). Prevalence of maternal antenatal and postnatal depression and their association with sociodemographic and socioeconomic factors: A multicentre study in Italy. *Journal of Affective Disorders, 279*, 217-221. https://doi.org/10.1016/j.jad.2020.09.136
17. Cena, L., Palumbo, G., Mirabella, F., Gigantesco, A., Stefana, A., Trainini, A., Tralli, N., Imbasciati, A., & Observatory of Perinatal Clinical Psychology Group. (2020). Perspectives on early screening and prompt intervention to identify and treat maternal perinatal mental health. Protocol for a prospective multicenter study in Italy. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.00365
18. Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159. https://doi.org/10.1037//0033-2909.112.1.155
19. Coleman, V. H., Carter, M. M., Morgan, M. A., & Schulkin, J. (2008). Obstetrician-gynecologists' screening patterns for anxiety during pregnancy. *Depression and Anxiety, 25*(2), 114-123. https://doi.org/10.1002/da.20278
20. Comrey, A. L., & Lee, H. B. (2013). A *first course in factor analysis*. Psychology press. https://doi.org/10.4324/9781315827506
21. Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry, 150*(6), 782-786. https://doi.org/10.1192/bjp.150.6.782
22. Curtiss, J., & Klemanski, D. H. (2015). Identifying individuals with generalised anxiety disorder: A receiver operator characteristic analysis of theoretically relevant measures. *Behaviour Change, 32*(4), 255-272. https://psycnet.apa.org/doi/10.1017/bec.2015.15
23. Davies, S. M., Christiansen, P., Harrold, J. A., Silverio, S. A., & Fallon, V. (2021). Creation and validation of the postpartum specific anxiety scale research short-form (PSAS-RSF). *Archives of Women's Mental Health, 24*(6), 957-969. https://doi.org/10.1007/s00737-021-01114-7
24. DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 837-845. https://doi.org/10.2307/2531595
25. Dennis, C. L., Coghlan, M., & Vigod, S. (2013). Can we identify mothers at-risk for postpartum anxiety in the immediate postpartum period using the State-Trait Anxiety Inventory?. *Journal of Affective Disorders, 150*(3), 1217-1220. https://doi.org/10.1016/j.jad.2013.05.049
26. Dennis, C. L., Falah-Hassani, K., & Shiri, R. (2017). Prevalence of antenatal and postnatal anxiety: systematic review and meta-analysis. *British Journal of Psychiatry, 210*(5), 315-323. https://doi.org/10.1192/bjp.bp.116.187179

27. Dimitriu, M., Stoian, A. M. P., Roşu, G., Dan, A., Caimacan, A., Zygouropoulos, N., Pacu, I., Socea, B., Ionescu, C., Gheorghiu, D. C., Bacalbasa, N., Cîrstoveanu, C., Haradja, H., Gheorghiu, N., Al-Azawi, A., Măcău, A-M-. Ples, L., Merlo, E. M., Settineri, S., & Jacotă–Alexe, F. (2020). Postpartum Depression-a serious, dangerous, disabling, extremely frequent condition that is almost completely ignored in Romania. *Mediterranean Journal of Clinical Psychology, 8*(1). https://doi.org/10.6092/2282-1619/mjcp-2413
28. Dunkel Schetter, C. (2011). Psychological science on pregnancy: stress processes, biopsychosocial models, and emerging research issues. *Annual Review of Psychology, 62*(1), 531-558. https://doi.org/10.1146/annurev.psych.031809.130727
29. Ellis, L. A., Pomare, C., Churruca, K., Carrigan, A., Meulenbroeks, I., Saba, M., & Braithwaite, J. (2022). Predictors of response rates of safety culture questionnaires in healthcare: a systematic review and analysis. *BMJ Open, 12*(9), e065320. https://doi.org/10.1136/bmjopen-2022-065320
30. Fan, S., Guan, J., Cao, L., Wang, M., Zhao, H., Chen, L., & Yan, L. (2021). Psychological effects caused by COVID-19 pandemic on pregnant women: a systematic review with meta-analysis. *Asian Journal of Psychiatry, 56*, 102533. https://doi.org/10.1016/j.ajp.2020.102533
31. Farias, D. R., Pinto, T. D. J. P., Teofilo, M. M. A., Vilela, A. A. F., dos Santos Vaz, J., Nardi, A. E., & Kac, G. (2013). Prevalence of psychiatric disorders in the first trimester of pregnancy and factors associated with current suicide risk. *Psychiatry Research, 210*(3), 962-968. https://doi.org/10.1016/j.psychres.2013.08.053
32. Farré-Sender, B., Torres, A., Gelabert, E., Andrés, S., Roca, A., Lasheras, G., Valdés, M., & Garcia-Esteve, L. (2018). Mother–infant bonding in the postpartum period: Assessment of the impact of pre-delivery factors in a clinical sample. *Archives of Women's Mental Health, 21*, 287-297. https://doi.org/10.1007/s00737-017-0785-y
33. Feuerstahler, L. M., Waller, N., & MacDonald III, A. (2020). Improving measurement precision in experimental psychopathology using item response theory. *Educational and Psychological Measurement, 80*(4), 695-725. https://doi.org/10.1177/0013164419892049
34. Field, T. (2017). Prenatal anxiety effects: a review. *Infant Behavior and Development, 49*, 120-128. https://doi.org/10.1016/j.infbeh.2017.08.008
35. Fraser, E. E., Ogden, K. J., Radford, A., Ingram, E. R., Campbell, J. E., Dennis, A., & Corbould, A. M. (2023). Exploring the psychological wellbeing of women with gestational diabetes mellitus (GDM): increased risk of anxiety in women requiring insulin. A Prospective Longitudinal Observational Pilot Study. *Health Psychology and Behavioral Medicine, 11*(1), 2170378. https://doi.org/10.1080/21642850.2023.2170378
36. George, A., Luz, R. F., De Tychey, C., Thilly, N., & Spitz, E. (2013). Anxiety symptoms and coping strategies in the perinatal period. *BMC Pregnancy and Childbirth, 13*, 1-6. https://doi.org/10.1186/1471-2393-13-233
37. Gigantesco, A., Palumbo, G., Cena, L., Camoni, L., Trainini, A., Stefana, A., & Mirabella, F. (2022). A Brief Depression Screening Tool for Perinatal Clinical Practice: The Performance of the PHQ-2 Compared with the PHQ-9. *Journal of Midwifery & Women's Health, 67*(5), 586-592. https://doi.org/10.1111/jmwh.13389
38. Goldman, K. N., Blakemore, J., Kramer, Y., McCulloh, D. H., Lawson, A., & Grifo, J. A. (2019). Beyond the biopsy: predictors of decision regret and anxiety following preimplantation genetic testing for aneuploidy. *Human Reproduction, 34*(7), 1260-1269. https://doi.org/10.1093/humrep/dez080

39. Grant, K. A., McMahon, C., & Austin, M. P. (2008). Maternal anxiety during the transition to parenthood: a prospective study. *Journal of Affective Disorders, 108*(1-2), 101-111. https://doi.org/10.1016/j.jad.2007.10.002
40. Grech, A. M., Sharma, S., Kizirian, N., & Gordon, A. (2024). Impact of the COVID-19 pandemic on new parents enrolled in the 'BABY1000'birth cohort study in Sydney, Australia: A mixed-methods study. *Australian and New Zealand Journal of Public Health, 48*(1), 100127. https://doi.org/10.1016/j.anzjph.2024.100127
41. Grigoriadis, S., Graves, L., Peer, M., Mamisashvili, L., Tomlinson, G., Vigod, S. N., Dennis, C-L., Steiner, M., Brown, C., Cheung, A., Dawson, H., Rector, N. A., Guenette, M., & Richter, M. (2019). A systematic review and meta-analysis of the effects of antenatal anxiety on postpartum outcomes. *Archives of Women's Mental Health, 22*, 543-556. https://doi.org/10.1007/s00737-018-0930-2
42. Hadfield, K., Akyirem, S., Sartori, L., Abdul-Latif, A. M., Akaateba, D., Bayrampour, H., Daly, A., Hadfield, K., & Abiiro, G. A. (2022). Measurement of pregnancy-related anxiety worldwide: a systematic review. *BMC Pregnancy and Childbirth, 22*(1), 331. https://doi.org/10.1186/s12884-022-04661-8
43. Hall, W. A., Hauck, Y. L., Carty, E. M., Hutton, E. K., Fenwick, J., & Stoll, K. (2009). Childbirth fear, anxiety, fatigue, and sleep deprivation in pregnant women. *Journal of Obstetric, Gynecologic & Neonatal Nursing, 38*(5), 567-576. https://doi.org/10.1111/j.1552-6909.2009.01054.x
44. Hewison, J., Nixon, J., Fountain, J., Hawkins, K., Jones, C. R., Mason, G., Morley, S., & Thornton, J. G. (2007). A randomised trial of two methods of issuing prenatal test results: the ARIA (Amniocentesis Results: Investigation of Anxiety) trial. *BJOG, 114*(4), 462-468. https://doi.org/10.1111/j.1471-0528.2007.01276.x
45. Howard, L. M., & Khalifeh, H. (2020). Perinatal mental health: a review of progress and challenges. *World Psychiatry, 19*(3), 313-327. https://doi.org/10.1002/wps.20769
46. Ibanez, G., Charles, M. A., Forhan, A., Magnin, G., Thiebaugeorges, O., Kaminski, M., Saurel-CubizollesM-J., & EDEN Mother–Child Cohort Study Group. (2012). Depression and anxiety in women during pregnancy and neonatal outcome: data from the EDEN mother–child cohort. *Early Human Development, 88*(8), 643-649. https://doi.org/10.1016/j.earlhumdev.2012.01.014
47. Jacobson, N. S. & Truax P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19. https://psycnet.apa.org/doi/10.1037/0022-006X.59.1.12
48. Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status: ascertaining the minimal clinically important difference. *Controlled Clinical Trials, 10*(4), 407-415. https://doi.org/10.1016/0197-2456(89)90005-6
49. Kazis, L. E., Anderson, J. J., & Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical Care, 27*(3), S178-S189. https://doi.org/10.1097/00005650-198903001-00015
50. Knowles, K. A., & Olatunji, B. O. (2020). Specificity of trait anxiety in anxiety and depression: Meta-analysis of the State-Trait Anxiety Inventory. *Clinical Psychology Review, 82*, 101928. https://psycnet.apa.org/doi/10.1016/j.cpr.2020.101928

51. Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x
52. Lindwall, M., Barkoukis, V., Grano, C., Lucidi, F., Raudsepp, L., Liukkonen, J., & Thøgersen-Ntoumani, C. (2012). Method effects: The problem with negatively versus positively keyed items. *Journal of Personality Assessment, 94*(2), 196-204. https://doi.org/10.1080/00223891.2011.645936
53. Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology, 31*(3), 301-306. https://doi.org/10.1111/j.2044-8260.1992.tb00997.x
54. NICE. (2014). Antenatal and postnatal mental health: clinical management and service guidance. National Institute for Health and Care Excellence (NICE), London. Available: https://www.ncbi.nlm.nih.gov/books/NBK553127/
55. Picardi, A., Abeni, D., Mazzotti, E., Fassone, G., Lega, I., Ramieri, L., Sagoni, E., Tiago, A., & Pasquini, P. (2004). Screening for psychiatric disorders in patients with skin diseases: a performance study of the 12-item General Health Questionnaire. *Journal of Psychosomatic Research, 57*(3), 219-223. https://doi.org/10.1016/s0022-3999(03)00619-6
56. Revelle, W., & Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological Assessment, 31*(12), 1395. https://psycnet.apa.org/doi/10.1037/pas0000754
57. Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology, 61*(2), 102-109. https://doi.org/10.1016/j.jclinepi.2007.03.012
58. Royal College of General Practitioners. Antenatal and postnatal mental health NICE guideline CG192: Practical implications for GPs. 2023. Available: https://elearning.rcgp.org.uk/pluginfile.php/175878/mod_book/chapter/606/Perinatalmentalhealth2023.pdf
59. Rubertsson, C., Hellström, J., Cross, M., & Sydsjö, G. (2014). Anxiety in early pregnancy: prevalence and contributing factors. *Archives of Women's Mental Health, 17*, 221-228. https://doi.org/10.1007/s00737-013-0409-0
60. Samejima, F. (2010). The general graded response model. In M. L. Nering & R. Ostini (Eds.), *Handbook of polytomous item response theory models* (pp. 77–107). Routledge.
61. Smith, G. T., McCarthy D.M., & Anderson K.G. (2000). On the sins of short-form development. *Psychological Assessment, 12*, 102–111. https://doi.org/10.1037//1040-3590.12.1.102
62. Smith Jr, E. V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement, 3*(2), 205-231.
63. Spielberger, C. D. (1983). *Manual for the State-Trait Inventory STAI (Form Y)*. Mind Garden.
64. Spielberger, C. D. (1989). *State-trait anxiety inventory* (2nd ed). Palo Alto.

65. Spielberger, C. D., Pedrabissi, L., & Santinello, M. (1989). *Inventario per l'ansia di stato e di tratto: nuova versione italiana dello STAI, forma Y: manuale*. Organizzazioni speciali.
66. Stefana, A., Cena, L., Trainini, A., Palumbo, G., Gigantesco, A., & Mirabella, F. (2024). A Screening for antenatal maternal depression: comparative performance of the Edinburgh Postnatal Depression Scale and Patient Health Questionnaire-9. *Annali dell'Istituto Superiore di Sanità, 60*(1), 55-63. https://doi.org/10.4415/ANN_24_01_08
67. Stefana, A., Damiani, S., Granziol, U., Provenzani, U., Solmi, M., Youngstrom, E. A., & Fusar-Poli, P. (2025a). Psychological, psychiatric, and behavioral sciences measurement scales: best practice guidelines for their development and validation. *Frontiers in Psychology, 15*, 1494261. https://doi.org/10.3389/fpsyg.2024.1494261
68. Stefana, A., Fusar-Poli, P., Vieta, E., Gelso, C. J., & Youngstrom, E. A. (2025b). Development and validation of an 8-item version of the Real Relationship Inventory-Client form. *Psychotherapy Research, 35*(3), 395–411. https://doi.org/10.1080/10503307.2024.2320331
69. Stefana, A., Langfus, J., Palumbo, G., Cena, L., Trainini, A., Gigantesco, A., & Mirabella, F. (2023). Comparing the factor structures and reliabilities of the EPDS and the PHQ-9 for screening antepartum and postpartum depression: a multigroup confirmatory factor analysis. *Archives of Women's Mental Health, 26*(5):659-668. https://doi.org/10.1007/s00737-023-01337-w
70. Streiner, D. L., Norman G.R., & Cairney J. (2015). *Health measurement scales: A practical guide to their development and use.* (5th ed.) Oxford, New York: Oxford University Press. https://doi.org/10.1093/med/9780199685219.001.0001
71. Val, A., & Míguez, M. C. (2023). Prevalence of antenatal anxiety in European women: a literature review. *International Journal of Environmental Research and Public Health, 20*(2), 1098. https://doi.org/10.3390/ijerph20021098
72. Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment, 82*(1), 50-59. https://doi.org/10.1207/s15327752jpa8201_10
73. Yin, X., Sun, N., Jiang, N., Xu, X., Gan, Y., Zhang, J., Qiu, L., Yang, C., Shi, X., Chang, J., & Gong, Y. (2021). Prevalence and associated factors of antenatal depression: Systematic reviews and meta-analyses. *Clinical Psychology Review, 83*, 101932. https://doi.org/10.1016/j.cpr.2020.101932
74. Youngstrom, E. A., Van Meter, A., Frazier, T. W., Youngstrom, J. K., & Findling, R. L. (2020). Developing and validating short forms of the parent general behavior inventory mania and depression scales for rating youth mood symptoms. *Journal of Clinical Child & Adolescent Psychology, 49*(2), 162–177. https://doi.org/10.1080/15374416.2018.1491006
75. Zhang, X., & Savalei, V. (2016). Improving the factor structure of psychological scales: The Expanded format as an alternative to the Likert scale format. *Educational and Psychological Measurement, 76*(3), 357-386. https://doi.org/10.1177/0013164415596421

76. Zsido, A. N., Teleki, S. A., Csokasi, K., Rozsa, S., & Bandi, S. A. (2020). Development of the short version of the spielberger state—trait anxiety inventory. *Psychiatry Research, 291*, 113223. https://doi.org/10.1016/j.psychres.2020.113223

©2025 by the Author(s); licensee Mediterranean Journal of Clinical Psychology, Messina, Italy. This article is an open access article, licensed under a Creative Commons Attribution 4.0 Unported License. Mediterranean Journal of Clinical Psychology, Vol. 13, No. 1 (2025). International License (https://creativecommons.org/licenses/by/4.0/).
DOI: 10.13129/2282-1619/mjcp-4505

Appendix A

State Anxiety Scale–Short Form

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you feel right now, that is, *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

Item		Not at All	Moderately so	Somewhat	Not at all
1	I am tense.	1	2	3	4
2	I feal strained.	1	2	3	4
3	I feel nervous.	1	2	3	4
4	I am jittery.	1	2	3	4
5	I am worried.	1	2	3	4

Trait Anxiety Scale–Short Form

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you *generally* feel.

Item		Almost never	Sometimes	Often	Almost always
1	I feel nervous and restless.	1	2	3	4
2	I feel that difficulties are piling up so that I cannot overcome them.	1	2	3	4
3	I have disturbing thoughts.	1	2	3	4
4	I feel inadequate.	1	2	3	4
5	I take disappointments so keenly that I can't put them out of my mind.	1	2	3	4